Priming of lowpass-filtered speech affects response bias, not sensitivity, in a bandwidth discrimination task

Richard L. Freyman^{a)} and Amanda M. Griffin

Department of Communication Disorders, University of Massachusetts, 358 North Pleasant Street, Amherst, Massachusetts 01003

Neil A. Macmillan

Department of Psychology, University of Massachusetts, Tobin Hall, Amherst, Massachusetts 01003

(Received 26 February 2013; revised 1 May 2013; accepted 9 May 2013)

Priming is demonstrated when prior information about the content of a distorted, filtered, or masked auditory message improves its clarity. The current experiment attempted to quantify aspects of priming by determining its effects on performance and bias in a lowpass-filter-cutoff frequency discrimination task. Nonsense sentences recorded by a female talker were sharply lowpass filtered at a nominal cutoff frequency (F) of 0.5 or 0.75 kHz or at a higher cutoff frequency ($F + \Delta F$). The listeners' task was to determine which interval of a two-interval-forced-choice trial contained the nonsense sentence filtered with $F + \Delta F$. On priming trials, the interval 1 sentence was displayed on a computer screen prior to the auditory portion of the trial. The prime markedly affected bias, increasing the number of correct and incorrect interval 1 responses but did not affect overall discrimination performance substantially. These findings were supported through a second experiment that required listeners to make confidence judgments. The paradigm has the potential to help quantify the limits of speech perception when uncertainty about the auditory message is removed.

© 2013 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4807824]

PACS number(s): 43.66.Fe, 43.66.Lj, 43.71.Gv [EB]

Pages: 1183–1192

I. INTRODUCTION

When a listener hears speech that has been severely degraded, a striking change in perception takes place once he/she knows the content of the message. A compelling demonstration related to this experience is to listen to a sentence that has been processed through a noise vocoder with a small number of channels. Depending on the precise stimulus conditions, listeners will hear the signal as a noisy distorted sound that may or may not be interpreted as speech, and, if it is, is likely to have limited intelligibility. However, after listeners are informed of the content of the message and the presentation of the vocoded sample is repeated, they are often impressed about how easy it is to understand the message.

Listeners of course rarely know the exact content of a sentence before they hear it. However, they often benefit from knowledge of the context of the message (e.g., Bilger *et al.*, 1984; Nittrouer and Boothroyd, 1990; Pichora-Fuller *et al.*, 1995; Dubno *et al.*, 2000; Most and Adi-Bensaid, 2001; Fallon *et al.*, 2002; Helfer and Freyman, 2008; Sheldon *et al.*, 2008). For example, Bilger *et al.* (1984) measured word recognition in adult listeners with sensorineural hearing loss using low- and high-predictability sentences in noise. Their results showed better speech recognition performance for high-context sentences than low-context sentences.

Even when they are not explicitly memorized, words listeners are exposed to a few minutes before they hear them or

see them in a test environment are identified more readily than words to which they were not exposed (e.g., Jacoby and Dallas, 1981; Roediger, 1990; Tulving and Schacter 1990; Schacter and Church 1992; Church and Schacter 1994; Schacter *et al.* 1994; Ratcliff *et al.* 1997; Ratcliff and McKoon 1997; Pilotti *et al.* 2000). This is often known as “priming.” These studies usually include relatively long lists of words provided before test trials begin and so do not eliminate uncertainty about what will be presented on each trial.

The current experiments explore the case of priming where all uncertainty about what will be heard on each trial is removed. Knowing the exact content of an auditory message before hearing it is the limiting case of providing context or previous exposure. It could serve as an upper bound on the effectiveness of providing context and as such a landmark against which listeners' ability to use context can be compared. Unfortunately, measurement in this boundary condition is difficult because once the message is known, the listener could simply remember it and repeat it whether or not the degraded auditory signal was correctly perceived.

A few different approaches have been used to try to get around this problem. Freyman *et al.* (2004) investigated the priming of sentences that were presented in a highly confusable competing speech background. The first two key words of a three-key-word nonsense sentence were primed. The listener's task was to identify a third key word that was not primed. To the extent that priming helped listeners perceptually extract the target message from the background, the improved perception was predicted to continue as the sentence continued via processes related to auditory streaming, improving recognition of the third key word. Using this

^{a)}Author to whom correspondence should be addressed. Electronic mail: rlf@comdis.umass.edu

partial priming technique, Freyman *et al.* (2004) and Ezzatian *et al.* (2011) (for English speech) as well as Yang *et al.* (2007) and Wu *et al.* (2011, 2012) (for Chinese speech) all found substantial benefits of priming in the presence of competing speech maskers. Sheldon *et al.* (2008) used a similar technique but included higher-context, noise-vocoded sentences. Their results showed that both linguistic context and priming increased performance and that the two effects facilitated one another when used in combination.

A second approach was employed by Jones and Freyman (2012) in which the entire nonsense sentence was provided as written text on half the trials while on the other half one of the three key words was replaced with a foil. Listeners were asked if what they heard was the same sentence as what they read. In two separate conditions, the written text was provided before or instead after the auditory presentation of the sentence. The effectiveness of priming was measured by comparing listeners' performance in the two orders of presentation. The results demonstrated large benefits of priming in the presence of interfering speech, steady noise, and speech-envelope modulated noise. Priming appears to help listeners make better use of barely audible speech information.

Neither the partial priming technique nor the same-different method described in the preceding text completely removes uncertainty about what will be heard during a trial. Partial priming does not provide a direct prime for the main key word to be tested. The same-different task, by definition, includes trials in which the priming sentence does not match the target sentence, 50% of the trials in the case of Jones and Freyman (2012). The experiments reported in the current paper were an attempt to further understand how priming affects the auditory perception of speech by asking listeners to make judgments about what they heard for sentences that were primed completely. That is, on every priming trial, they were told exactly, in printed text, the content of the spectrally degraded sentence they were about to hear. If subjects were asked to report the sentence they heard, it would be too easy for them to simply report what they read. Instead we asked how priming affects the ability to discriminate between sentences that had been degraded to slightly different degrees. The reasoning behind the approach was that to the extent that primed speech sounds clearer than unprimed speech, it could make it difficult to judge whether an acoustic signal is more or less distorted than another acoustic signal that is not primed. If successful, this approach could help quantify how speech perception is affected by the removal of uncertainty about what will be heard.

The current study specifically investigated how priming affects listeners' ability to discriminate differences in the cutoff frequency of lowpass-filtered speech. A sample of speech filtered with a substantially higher cutoff frequency than another sample of speech would presumably sound different in a number of ways. The speech with the increased high-frequency content might be noticeably more intelligible. The inclusion of higher-frequency energy would also change the timbre and possibly the loudness. This paper considers the case where one of two speech samples is primed

so that the listener knows the content of one of the messages. The task was to select which of the two samples was filtered with the higher cutoff frequency and so spanned the greater bandwidth. If the speech sample with the lower cutoff frequency is primed, it might sound at least as intelligible as the unprimed sample with the higher cutoff frequency, increasing the likelihood that subjects would select it as the incorrect answer. If the sample with the higher cutoff frequency is primed, an already more intelligible sentence will sound even clearer, enhancing the likelihood that the sample will be chosen as the one with the higher cutoff. Thus it could be predicted that the primed sentence will be selected as the one with the higher cutoff frequency regardless of whether it is true or not.

Other outcomes are also possible. For example, listeners may be able to focus on some other aspect of the signal, such as timbre or loudness, that is perhaps not altered by priming. In that case, a bias toward selecting the primed interval may not be observed. It is also possible that by making speech recognition easier, priming would free up resources that would allow subjects to focus on acoustic differences between the utterances and improve discrimination performance. Finally, it is possible that listeners could learn through correct-answer feedback that priming was biasing their responses and attempt to compensate for this. Specifically, if the perceived clarity of a primed and unprimed speech sample was approximately equal, and subjects understood that priming was increasing the clarity of the primed sentence, they may well decide to choose the unprimed sentence as having the higher frequency content. This perceptual compensation would never be expected to be perfect, leading to the expectation of decreased discrimination performance.

II. PRELIMINARY INTELLIGIBILITY STUDY

The purpose of this preliminary study was to verify the premise upon which the design of the main experiments was based, specifically that having prior knowledge of the content of a lowpass-filtered sentence improves intelligibility. The remainder of the studies examined the effect of this presumed increased intelligibility on listeners' responses when asked to compare sentences filtered with different cutoff frequencies. The task employed in the preliminary study was identical to that used by Jones and Freyman (2012), although that study used several types of interfering sounds as a means to limit intelligibility, while the current study used lowpass filtering. Many of the details of the procedure can be found in that paper. Briefly, listeners were asked whether the sentences they heard were the same as the ones they saw printed on a computer monitor. On half the trials, the typed and acoustically presented sentences matched exactly (same trials), whereas on the other half, one of the three key words in the printed sentence did not match the acoustically presented sentence (foil trials). On 50% of the same and foil trials, the printed sentence was delivered 3 s in advance of the auditory sentence and ended immediately before the auditory sentence was presented (prime trials). On the other 50% of trials, the printed sentence followed the onset of the

presentation of the auditory sentence by three seconds (control trials). The experimental question was the extent to which the order of acoustic and printed presentations affected performance on the same-different task. A block consisted of 90 trials in which 45 were foil trials and 45 were same trials. Within the 45 trials, each of the five low-pass cutoff frequencies, which covered the range from 500 to 900 Hz in 100-Hz steps, was delivered nine times in a randomly interspersed order. For foil trials, each of the three possible foil positions was used three times each. Signals were presented at 63 dBA. Six young normal-hearing subjects participated.

The results are shown in Fig. 1. Consistent with the data of Jones and Freyman (2012) for masked signals, performance was better when the printed text was delivered before the lowpass-filtered auditory presentation than after [$F(1,5) = 71.248, p < 0.001$]. The effect of priming on d' increased as the cutoff frequency and performance in the control condition increased. The improvement in performance in the priming condition relative to the control was interpreted as an enhancement of intelligibility of the low-pass filtered speech, which facilitated an improved ability to compare the written and auditory signals. This preliminary result thus supports the premise that intelligibility of lowpass-filtered speech is altered by prior knowledge of message content.

III. EXPERIMENT 1. FILTER CUTOFF FREQUENCY DISCRIMINATION

This experiment examined listeners' ability to discriminate differences in the high-frequency cutoff of lowpass-filtered speech in a two-interval forced-choice (2IFC) paradigm. On half the trials, subjects were asked to read the sentence to be presented in the first of the two intervals. The purpose was to determine how this information would affect sensitivity and bias in the bandwidth discrimination task.

FIG. 1. Discrimination performance across listeners in the preliminary experiment. Performance in d' is plotted as a function of cutoff frequency for the prime and control conditions. Each data point represents 108 trials, 18 from each of 6 listeners. Error bars represent ± 1 standard error of the mean.

A. Methods

1. Subjects

Forty normal-hearing listeners (39 females, 1 male) with audiometric thresholds ≤ 20 dB hearing level (HL) at octave frequencies between 500 and 8000 Hz participated in this experiment. Their ages ranged between 20 and 47 yr. Subjects were recruited from the undergraduate Communication Disorders Program at the University of Massachusetts Amherst and received extra course credit for their participation.

2. Stimuli and procedures

The target stimuli were nonsense sentences spoken by a female talker. These sentences were syntactically but not semantically correct and contained three key words, e.g., "The ocean could shadow our peak." Details of the recording methodology are found in Helfer (1997). The recorded waveforms were sampled at 22 050 Hz and stored on hard drive of a personal computer. The nonsense sentences were sharply lowpass filtered at a nominal cutoff frequency (F) with a ± 50 -Hz rove (from a rectangular distribution) or at a higher cutoff frequency ($F + \Delta F$). Two nominal standard cutoff frequencies were utilized: 500 and 750 Hz. A 256th-order finite impulse response (FIR) filter was constructed using a Hanning-window design and was applied to all speech signals. For the 500-Hz cutoff frequency, the filter's attenuation at one octave (1000 Hz) was approximately 66 dB; for the 750-Hz cutoff, the filter's attenuation at one octave was approximately 78 dB. One group of 20 subjects listened to the 500-Hz standard, and the other group of 20 to the 750-Hz standard. Four ΔF 's were utilized for each of the two nominal standard frequencies: 25, 50, 75, and 100 Hz for the 500-Hz standard and 62.5, 125, 187.5, and 250 Hz for the 750-Hz standard. The listeners' task was to determine which interval of a 2IFC trial was presented with cutoff frequency ($F + \Delta F$). The paradigm is shown schematically in Fig. 2.

Each subject listened to a total of 160 trials, divided into two blocks of 80 trials. Condition (primed or unprimed) was fixed within each block with the order of blocks

FIG. 2. Schematic illustration of experimental paradigm used in experiment 1.

counterbalanced across listeners. On primed trials, a written message containing the interval 1 nonsense sentence was displayed on a computer monitor prior to the auditory portion of the trial.¹ For unprimed trials, the phrase “Ready Now” appeared on the computer monitor instead of the priming sentence. The printed sentence remained on the monitor for 2 s and was followed immediately by the acoustic presentation of the first and then the second nonsense sentence. The time between the onset of the first and second sentences was 3 s. Within each of the two blocks, there were eight different trials types. Twenty of the 80 trials were devoted to each of the four ΔF ’s. Within those 20, 10 had $F + \Delta F$ in interval 1 and 10 had $F + \Delta F$ in interval 2. The interval containing $F + \Delta F$ was randomly interleaved within a block of 80 trials, and different nonsense sentences were selected without replacement for each trial.

Subjects were instructed to pick the sentence interval (1 or 2) that was filtered with the higher cutoff frequency. They were told to use whatever cues they could to get the answer right but were informed that the sentence filtered with the higher cutoff frequency might sound clearer, brighter, or more intelligible. They were instructed to read the sentence on the computer monitor on each trial and were told that on priming trials (when the written text was a nonsense sentence), it would be the same sentence as the first sentence they would hear.

Prior to test sessions, subjects completed a short practice session to familiarize them with the experimental procedures. TVM sentences were used during the practice trials (Helfer and Freyman, 2009), which were not employed in the main experiment. There were 10 practice trials in total, five with priming and five unprimed. Subjects were exposed to several different ΔF ’s during the practice trials. Correct-answer feedback was given during the practice sessions.

Listening sessions were conducted in a double-walled sound-treated booth (IAC 1604). The stimuli were generated on a personal computer with a 24-bit sound card at a 22.05 kHz digital-to-analog conversion rate, attenuated (TDT PA4), and delivered monaurally through a TDH 39 headphone via a headphone amplifier (TDT HB5) and a passive attenuator. Custom MATLAB software (Mathworks, Natick, MA), executed on a desktop computer outside the test booth, was used to present the sentences to the subject,

provide correct-answer feedback on each trial, and record the subjects’ responses. Subjects entered the number 1 or 2 on a keyboard to indicate that sentence 1 or 2, respectively, had the higher cutoff frequency.

Presentation level was calibrated using a speech spectrum noise having the same rms level as unfiltered nonsense sentences, which had all been scaled for equal rms. The level of the calibration noise was 63 dBA as measured through a 6 cm³ coupler and sound level meter. No further scaling or adjustments were made as a result of the lowpass filtering during the experimental presentation.

B. Results

Figure 3 shows discrimination performance across listeners in d' as a function of the change in cutoff frequency for the 500-Hz (A) and 750-Hz (B) standards. Each data point in the figure represents 400 trials, 20 from each of 20 listeners. For both standard frequencies, performance generally increased as ΔF increased with the exception of the 250-Hz ΔF for the 750-Hz standard. With that one exception, the effect of priming on overall performance was fairly small and not consistent. For the 500-Hz standard, no significant effect of priming was found [$F(1,19) = 0.001$, $p = 0.979$]. The same was true for the 750-Hz standard [$F(1,19) = 2.119$, $p = 0.155$]. Performance on both primed and unprimed conditions reached a d' of 1 at about 40–50 Hz for the standard frequency of 500 Hz and about 125 Hz for the standard frequency of 750 Hz.

More about the effect of priming is revealed in the four panels of Fig. 4. There, the percentage of interval 1 responses is plotted separately for the cases where the higher cutoff frequency was in interval 1 [Figs. 4(A) and 4(B)] or in interval 2 [Figs. 4(C) and 4(D)]. When $F + \Delta F$ was in interval 1, priming increased the number of trials in which interval 1 was correctly selected. The effect was greatest when ΔF was small (16–17 percentage points) and decreased to about 7 percentage points at the largest ΔF . The change was largely due to a consistently increasing percentage in the no-prime condition while performance in the prime condition reached an asymptote. It is perhaps understandable that the effect of priming on the number of interval 1 responses for the sentence with the already higher cutoff frequency would

FIG. 3. Discrimination performance across listeners in d' as a function of the change in cutoff frequency for the 500-Hz (A) and 750-Hz (B) standards. Each data point represents 400 trials, 20 from each of 20 listeners. Error bars represent ± 1 standard error of the mean.

FIG. 4. The mean percentage of interval 1 responses is plotted separately for the cases where the higher cutoff frequency was in interval 1 [(A) and (B)] or in interval 2 [(C) and (D)]. Each data point represents 200 trials, 10 from each of 20 listeners. Error bars represent ± 1 standard error of the mean.

be greatest when it would exaggerate a small difference in intelligibility. Priming also increased the number of interval 1 responses when $F + \Delta F$ was in interval 2. Thus the overall effect of priming was to increase the number of both correct and incorrect interval 1 responses. For the 500-Hz cutoff frequency, approximately 71% of the responses were interval 1 in the priming condition, while approximately half (49.1%) of the responses were interval 1 in the no-prime condition.

For the 750-Hz nominal cutoff frequency, the percentage of interval 1 responses for the prime condition was 62.3% compared with 46.5% for the no-prime condition. This bias toward selecting the primed interval suggests strongly that clarity was increased in a way that caused listeners to think that the primed sentence contained the greater bandwidth.

IV. EXPERIMENT 2. CONFIDENCE RATINGS

The results of experiment 1 suggested that for each ΔF , primed and unprimed conditions produced differences in response bias but not in sensitivity. A simple indicator of bias is the percentage of interval 1 responses, which we have just seen was much higher in the prime than in the no-prime condition. Figure 3 suggests that the effect on sensitivity, as measured by d' , was slight at best. However, d' is a flawed sensitivity index, especially when response bias is substantial (as is the case here). For d' to be a pure measure of sensitivity, the underlying receiver operating characteristic (ROC) must be symmetric and consistent with equal-variance underlying Gaussian distributions. To see the problem, consider the hypothetical ROCs in Fig. 5. The curves plot proportions of “interval 1” responses when that interval contains the higher value of ΔF (hits) against the proportions of “interval 1” responses when interval 2 is correct (false alarms). The solid curve is constructed to have constant d' and is symmetric around the minor diagonal. The prime and no-prime points are at different locations on the curve, reflecting the difference in bias. But suppose that in both the prime and the no-prime conditions, the $F + \Delta F$ interval produces an effect that is more variable than the F interval (see Macmillan and Creelman, 2005, p. 74–77 for a discussion of

FIG. 5. Hypothetical ROCs illustrating two different loci for the priming effect. If the true ROC is symmetric (solid curve), then the prime condition (circle) and the no-prime condition (square) lie on the same curve, differing only in response bias. But if the true ROC is asymmetric (dashed curves), then the prime condition leads to greater accuracy as well as a change in bias.

this general topic). The resulting ROCs will be asymmetric as illustrated by the dashed lines in the figure. The prime and no-prime points are now seen to differ in both bias (overall “interval 1” rate) and sensitivity (height of the appropriate ROC curve). The results of experiment 1 cannot distinguish between two interpretations of the priming effect: Bias alone (if underlying distributions are equal-variance) and bias plus sensitivity (if those distributions differ in variance).

The purpose of the second experiment was to determine whether or not the points for the prime and no-prime conditions fell on the same or different ROCs. We obtained confidence-rating judgments on each trial instead of only binary judgments to establish 5-point ROCs for each condition (see Macmillan and Creelman, 2005, Chap. 3, for details of ROC construction and interpretation).

A. Methods

1. Subjects

Twenty normal-hearing adult listeners (18 females, 2 males) with audiometric thresholds of less than 20 dB HL at octave frequencies between 500 and 8000 Hz participated in experiment 2. Their ages ranged between 20 and 37 yr. Subjects were recruited from the undergraduate Communication Disorders Program at the University of Massachusetts Amherst and received extra course credit for their participation.

2. Stimuli and procedures

The stimuli and procedures were nearly identical to those utilized in experiment 1, including practice, stimulus levels, and calibration procedures. However, in experiment

2, only one standard cutoff frequency (750 Hz) was utilized with the four associated ΔF 's: 62.5, 125, 187.5, and 250 Hz. Unlike experiment 1, a confidence rating scale was employed. The subjects' task was to again choose which sentence had the higher cutoff frequency, but this time additionally indicate how confident they were in their decision. Subjects could choose between three different confidence indicators: “Sure,” “probably,” or “maybe.” Subjects entered the number 1, 2, or 3 if they felt that sentence 1 had the higher cutoff frequency and numbers 4, 5, or 6 if they felt that sentence 2 had the higher cutoff frequency. Numbers 1 and 6 represented “sure,” 2 and 5 represented “probably,” and 3 and 4 represented “maybe.” Feedback was provided after every trial. When $F + \Delta F$ was in interval 1, correct responses were 1, 2, and 3. When $F + \Delta F$ was in interval 2, correct responses were 4, 5, and 6.

B. Results

As in experiment 1, subjects selected interval 1 more frequently in the primed condition (68.4%, or 1095 of the 1600 trials across subjects) than in the unprimed condition (47.3%). Figure 6 displays the aggregate results. It is of interest to examine the frequency with which subjects selected a confidence value of 1, indicating highest confidence that the interval 1 sentence contained the higher bandwidth. As shown in Fig. 6(A), the number of highest confidence responses for interval 1 was 336 of 1600 total trials for prime and only 125 responses for no-prime (left-most pair of bars). Of those 336, 211 (of 800) were correct responses on interval 1 trials [Fig. 6(B)] and 125 were incorrect responses on

FIG. 6. Number of responses across all trials and subjects for each of the six confidence ratings in experiment 2 (A). The total number of trials for both prime and no-prime conditions is 1600. (B) and (C) show the number of trials for each of the six ratings separated into $F + \Delta F$ =interval 1 and $F + \Delta F$ =interval 2, respectively. 1 = “sure interval 1,” 2 = “probably interval 1,” 3 = “maybe interval 1,” 4 = “maybe interval 2,” 5 = “probably interval 2,” and 6 = “sure interval 2.”

interval 2 trials [Fig. 6(C)]. The latter value indicates that 125 times listeners reported highest confidence that interval 1 had the higher bandwidth even though the greater bandwidth was actually in interval 2. Not shown in the figure is that these 125 responses were approximately uniformly spread across the ΔF 's, including 28 responses (of 200 trials) for the largest ΔF of 250 Hz. Thus even when the cutoff frequency in interval 2 was 250 Hz higher than in interval 1, subjects sometimes reported with the highest confidence that the greater bandwidth was in interval 1. For the no-prime condition, the number of incorrect "1's" was only 26 of 800 trials across all ΔF 's [left-most open bar in Fig. 6(C)].

ROCs were constructed as described by Macmillan and Creelman (2005, Chap. 3) and are displayed in Fig. 7. In Fig. 7(A), the lowest point on each curve represents the proportion of incorrect "sure interval 1" responses when F Hz was the cutoff frequency in interval 1 (false alarms, horizontal axis) and correct "sure interval 1" responses when $F + 62.5$ Hz was the cutoff frequency in interval 1 (hits, vertical axis). The second point is the proportion of "sure interval 1" plus "probably interval 1" responses and represents a more liberal criterion than the first point. The other points are constructed similarly.

Figure 7(A) shows that priming shifted response bias: All points on the prime ROC were to the right of those for the no-prime ROC. There was also a small sensitivity difference in favor of the no-prime condition. A simple non-parametric measure of sensitivity is the area under the ROC (A_g ; see Macmillan and Creelman, 2005, p. 64), defined by connecting the ROC points with straight lines. The value of

A_g was 0.59 in the prime condition and 0.65 in the no-prime condition. The symmetry of the ROCs can be assessed by replotting them on inverse-normal coordinates (Macmillan and Creelman, 2005, Chap. 3) and evaluating the slopes. When this is done, the resulting ROCs are found to be symmetric (have slopes of approximately unity), so the potential difficulty of interpretation illustrated in Fig. 5 does not arise.

The results for the other values of ΔF can be summarized as follows. The response bias effect was replicated throughout: At every value of confidence, the prime condition produced more "interval 1" responses than the no-prime condition. The relative sensitivity of the two conditions was inconsistent: At $\Delta F = 125$ Hz, $A_g = 0.69$ in the prime condition and 0.65 in the no-prime condition; for $\Delta F = 187.5$ Hz, the values were 0.69 and 0.76; and for $\Delta F = 250$ Hz the values were 0.72 and 0.75. The no-prime condition was better in three of four cases, and the average advantage was 0.03. It seems safe to say that the sensitivity effect was small; perhaps more important, whatever effect does occur is in favor of non-priming. The priming manipulation does not improve accuracy.

V. EXPERIMENT III. NORMAL VERSUS REVERSED SPEECH

The results of the first two experiments suggest that the listeners were biased toward choosing interval 1 when the sentence presented during that interval was primed. Our interpretation is that priming increased intelligibility (see

FIG. 7. ROCs constructed from confidence ratings in experiment 2 for each of the four values of ΔF .

preliminary experiment), one of the cues that subjects may have used to determine which of two sentences was filtered with the higher cutoff frequency. It appears that intelligibility was a factor in subjects' judgments, but the question remains how well listeners could discriminate cutoff frequencies if intelligibility was not available as a cue. The current experiment compared discrimination performance in the cutoff frequency discrimination task for normal and time-reversed speech. The purpose of the time-reversed condition was to remove intelligibility from among the cues listeners could use to solve the discrimination task, because the speech was unintelligible at any cutoff frequency.

A. Methods

1. Subjects

Twenty normal-hearing listeners (8 females, 12 males) with audiometric thresholds ≤ 20 dB HL at octave frequencies between 500 and 8000 Hz participated in this experiment. Their ages ranged between 18 and 40 yr. Subjects were students at the University of Massachusetts Amherst and were paid for their participation.

2. Stimuli and procedures

The target stimuli and procedures were identical to those used in experiment 1 except that on half the blocks the low-pass filtered stimuli were played in reverse. Only the 500-Hz standard cutoff frequency, with the same roving as before, was used in this study. Only the unprimed condition was tested in two 80-trial blocks, one with forward speech, one with reversed speech, in counterbalanced orders as in experiment 1. All other details were identical to what was used in experiment 1, including counterbalancing, practice, etc.

B. Results

The results of this experiment, shown in Fig. 8, reveal relatively small and inconsistent differences between filter cutoff frequency discrimination in forward and reversed speech. A one-way analysis of variance showed the main effect of direction (forward versus backward) was not significant [$F(1,19) = 1.438$, $p = 0.245$]. The leveling off in performance for the reversed speech between delta cutoffs of 75 and 100 Hz is not easily explained. Performance would presumably increase further with larger cutoff frequency differences. The change in timbre may have been small in just that region. With forward speech, where intelligibility could also be used as a cue, performance increased steadily. From the results across the conditions, it must be concluded that with the possible exception of $\Delta f = 100$ Hz, cutoff frequency discrimination in the tested frequency region was no more difficult when intelligibility cues were available as when they were not available.

VI. DISCUSSION

Subjects were asked to compare two sentences to determine which of the two was lowpass filtered with the higher cutoff frequency. On experimental trials, they were presented with the exact content of the first of the two sentences

FIG. 8. Discrimination performance across listeners in experiment 3. Performance in d' is plotted as a function of cutoff frequency for the forward and backward conditions. Each data point represents 400 trials, 20 from each of 20 listeners. Error bars represent ± 1 standard error of the mean.

just before the trial. Our question was how this knowledge would affect their ability to correctly judge which speech signal had the greater bandwidth. The results were fairly clear. On priming trials, subjects selected the first of the two sentences as the one with the higher cutoff frequency on approximately two-thirds of trials averaged across conditions and experiments compared with just under half for non-primed trials. As a result, on priming trials, the number of correct responses increased when the correct answer was interval 1 and decreased when the correct answer was interval 2. Whereas this bias effect was strong, the two opposing results produced little net effect in overall discrimination performance as measured in d' . An experiment using confidence ratings and ROC curves extended this conclusion to multiple criterion settings.

Listeners presumably had several cues at their disposal to determine which of the two sentences presented on each trial included the more extended frequency content. The signal attenuation was not adjusted to compensate for the presumably greater sound intensity of the higher bandwidth signal, so increased loudness could have been one cue. The quality or timbre would also be expected to change as higher frequencies were added. Finally, the literature is replete with examples of how intelligibility is increased as the upper cutoff frequency of lowpass-filtered speech is increased, especially in the derivation of frequency importance functions (e.g., French and Steinberg, 1947; Fletcher and Galt, 1950; Studebaker *et al.*, 1987; Henry *et al.*, 1998; Whitmal and DeRoy, 2011). Here the effects of these cues (except in the preliminary same-different task) were not independently measured, but the judgments of any of these differences across the two intervals of the trial would be expected to contribute to subjects' decisions about which sentence had the higher cutoff frequency. The fact that the results of experiment 3 showed no overall advantage in the task for listening to normal speech versus time-reversed speech

suggests that timbre and/or loudness cues were sufficient, i.e., changes in intelligibility with increased cutoff frequency were not required to solve the discrimination task.

In this regard, it may be puzzling that priming biased responses toward the primed interval when tested in experiments 1 and 2. The simplest explanation for this effect, supported by the results of the preliminary experiment, is that priming increased the intelligibility of the sentence presented in interval 1, and subjects believed the more intelligible sentence was the one filtered with the higher cutoff frequency. Thus even though experiment 3 showed that intelligibility was generally not a necessary cue, it may well have been used when it was available in the normal (forward) conditions. Several aspects of the experimental design may well have strengthened listeners' apparent emphasis on intelligibility cues when they were available and perhaps deemphasized cues other than intelligibility. As suggested by the preliminary experiment, the effect of priming on intelligibility could be large, and subjects were explicitly told that increased intelligibility was one of the potential cues that would indicate the interval filtered with the higher cutoff frequency. The use of correct-answer feedback on every trial gave listeners the opportunity to learn to use any and all cues to solve the task, but mitigating this somewhat was the fact that the standard cutoff frequency was roved and ΔF varied from trial to trial. The use of blocks of trials with fixed ΔF might have provided a greater opportunity for listeners to learn to focus on subtle loudness or timbre differences when intelligibility differences were also present. The individual studies were conducted with 20 listeners, each of whom contributed 160 responses. An alternative design using fewer subjects who each listened to many more trials could have provided greater potential for learning that attending to intelligibility differences did not improve performance.

A priori there was some possibility that subjects would gain the intuition that reading the content of the first sentence altered the auditory perception of that sentence when it was heard immediately afterward. However, if they had such intuitions, it is not obvious in the experimental results. The large majority of '1' responses suggests that priming made the interval 1 sentence appear more intelligible and subjects simply selected that sentence rather than trying to compensate for whatever they perceived to be the effect of knowing the sentence content if any.

There are few examples we are aware of in the literature in which listeners' estimation of intelligibility is affected by knowledge of the speech presented prior to or simultaneous with the stimulus trial. Rankovic and Levy (1997) investigated listeners' ability to estimate percent correct intelligibility using orthographic representations of nonsense syllables displayed on a computer monitor either simultaneously with the auditory presentation or 500 ms after the auditory presentation. Most subjects overestimated actual scores in the simultaneous presentation, suggesting that they may have been hearing the speech sounds more clearly in this condition. Similarly Sohoglu *et al.* (2012) found that visual presentation of monosyllabic words presented prior to a degraded auditory presentation of those same words produced higher subjective ratings of intelligibility than neutral

or mismatched visual displays. In a related study, Jacoby *et al.* (1988) asked subjects to judge the loudness of a masking noise during a speech recognition task. The noise was judged to be less loud when presented with sentences that the subjects had been exposed to previously. The relationship to the current study is that in both examples priming altered the listening experience, such that masking, in the case of Jacoby *et al.* (1988), and the degree of degradation, in our case, appeared to be reduced.

As mentioned in Sec. I, measuring the effects of priming objectively has been challenging. The bias effects observed in the current study could potentially be useful in this way, specifically if these effects were expressed in terms of equivalent added bandwidth. Figure 4(C) perhaps gives the most straightforward way of viewing the data for this kind of analysis. Here, interval 1 responses are incorrect because the $F + \Delta F$ stimulus was in interval 2. These incorrect responses decrease with increasing ΔF for both primed and unprimed conditions, roughly in parallel. The function for the primed condition is shifted to the right by about 50 Hz. For example, the result of 30% interval 1 responses was obtained at approximately 100-Hz ΔF for the primed condition compared with 50 Hz for the unprimed condition. This could be interpreted as suggesting that for the 500-Hz nominal standard cutoff frequency, priming caused the sentences to appear as intelligible as if the bandwidth was increased by 50 Hz. This does not mean, of course, that the central processing of the primed stimulus is the same as what is produced by the less degraded unprimed stimulus. For example, using MEG recordings, Sohoglu *et al.* (2012) showed that priming produced specific neural activity patterns that were different from when test words were preceded by incorrect or neutral stimuli.

A similar analysis could be applied to the bias toward correct interval 1 responses in Fig. 3(B), although the absence of parallel functions makes the expression of the effect in terms of equivalent cutoff frequency more dependent on the criterion performance chosen for comparison. Complicating this approach to quantifying priming is that the listeners had access to other cues besides intelligibility, and it is possible that the effect of priming on intelligibility could have been underestimated because of this. An alternative approach would be to study the effect of priming when listeners are asked to compare the intelligibilities of utterances that are degraded in completely different ways. This would force the judgments to be based on intelligibility alone. Conducting this as an objective task would require assigning a correct answer on each trial and therefore would necessitate that data on intelligibility growth functions for both types of degradations be available prior to the experiment.

In summary, the results show that knowing the content of an utterance ahead of time changes auditory perception of lowpass-filtered speech in a way that is measurable in a 2IFC discrimination experiment. The same technique could be applied to speech with varying levels of other types of distortion, such as degraded spectral envelope or temporal fine structure information, or to speech presented in varying levels of noise or competing speech. The current design,

with a large number of subjects listening to a limited number of trials each, does not lend itself to an examination of individual differences in the perceptual changes caused by prior knowledge of message content. However, further explorations of the effects of priming using this technique could reveal which listeners might benefit most from aural rehabilitation therapy that emphasizes the use of context to improve the intelligibility of upcoming messages.

ACKNOWLEDGMENTS

The authors are grateful for the assistance of Jamie Chevalier, Kaley Gray, Lindsay Woods, Yihao Lu, and Charlotte Morse-Fortier and for the support of the National Institute on Deafness and other Communication Disorders (R01 DC01625).

¹Note that a symmetrical design in which interval 2 would be primed could not be achieved without undesirable consequences. Specifically, if an interval 2 prime was delivered before both acoustic intervals (just like the interval 1 prime), the listener's memory for the prime might be affected by the intervening sentence and additional delay time. The discrimination task might be altered by the insertion of the prime between the two intervals; this could disrupt memory of what was perceived from the first and the second intervals.

- Bilger, R. C., Nuetzel, M. J., Rabinowitz, W. M., and Rzeczkowski, C. (1984). "Standardization of a test of speech perception in noise," *J. Speech Hear. Res.* **27**, 32–48.
- Church, B. A., and Schacter, D. L. (1994). "Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency," *J. Exp. Psychol. Learn. Mem. Cogn.* **20**, 521–533.
- Dubno, J. R., Ahlstrom, J. B., and Horwitz, A. R. (2000). "Use of context by young and aged adults with normal hearing," *J. Acoust. Soc. Am.* **107**, 538–546.
- Ezzatian, P., Li, L., Pichora-Fuller, K., and Schneider, B. (2011). "The effect of priming on release from informational masking is equivalent for younger and older adults," *Ear Hear.* **32**, 84–96.
- Fallon, M., Trehub, S. E., and Schneider, B. A. (2002). "Children's use of semantic cues in degraded listening environments," *J. Acoust. Soc. Am.* **111**, 2242–2249.
- Fletcher, H., and Galt, R. H. (1950). "The perception of speech and its relation to telephony," *J. Acoust. Soc. Am.* **22**, 89–151.
- French, N. R., and Steinberg, J. C. (1947). "Factors governing the intelligibility of speech sounds," *J. Acoust. Soc. Am.* **19**, 90–119.
- Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2004). "Effect of number of masking talkers and auditory priming on informational masking in speech recognition," *J. Acoust. Soc. Am.* **115**, 2246–2256.
- Helfer, K. S. (1997). "Auditory and auditory-visual perception of clear and conversational speech," *J. Speech Lang. Hear. Res.* **40**, 432–443.
- Helfer, K. S., and Freyman, R. L. (2008). "Aging and speech-on-speech masking," *Ear Hear.* **29**, 87–98.
- Helfer, K. S., and Freyman, R. L. (2009). "Lexical and indexical cues in maskin by competing speech," *J. Acoust. Soc. Am.* **125**, 447–456.
- Henry, B. A., McDermott, H. J., McKay, C. M., James, C. J., and Clark, G. M. (1998). "A frequency importance function for a new monosyllabic word test," *Aust. J. Audiol.* **20**, 70–86.
- Jacoby, L. L., Allan, L. G., Collins, J. C., and Larville, L. K. (1988). "Memory influences subjective experience: Noise judgement," *J. Exp. Psychol. Learn. Mem. Cogn.* **14**, 240–247.
- Jacoby, L. L., and Dallas, M. (1981). "On the relationship between autobiographical memory and perceptual learning," *J. Exp. Psychol. Gen.* **3**, 306–340.
- Jones, A. J., and Freyman, R. L. (2012). "Effect of priming on energetic and informational masking in a same-different task," *Ear Hear.* **33**, 124–133.
- Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*, 2nd ed. (Erlbaum, Mahwah, NJ), Chap. 3, pp. 64, 74–77.
- Most, T., and Adi-Bensaid, L. (2001). "The influence of contextual information on the perception of speech by postlingually and prelingually profoundly hearing-impaired Hebrew-speaking adolescents and adults," *Ear Hear.* **22**, 252–263.
- Nittrouer, S., and Boothroyd, A. (1990). "Context effects in phoneme and word recognition by young children and older adults," *J. Acoust. Soc. Am.* **87**, 2705–2715.
- Pichora-Fuller, M. K., Schneider, B. A., Daneman, M. (1995). "How young and old adults listen to and remember speech in noise," *J. Acoust. Soc. Am.* **97**(1), 593–608.
- Pilotti, M., Bergman, E. T., Gallo, D. A., Sommers, M., Roediger, H. L., III (2000). "Direct comparison of auditory implicit memory tests," *Psychon. Bull. Rev.* **7**, 347–353.
- Rankovic, C. M., and Levy, R. M. (1997). "Estimating articulation scores," *J. Acoust. Soc. Am.* **102**, 3754–3761.
- Ratcliff, R., Allbritton, D., and McKoon, G. (1997). "Bias in auditory priming," *J. Exp. Psychol. Learn. Mem. Cogn.* **23**, 143–152.
- Ratcliff, R., and McKoon, G. (1997). "A counter model for implicit priming in perceptual word identification," *Psychol. Rev.* **104**, 319–343.
- Roediger, H. L., III. (1990). "Implicit memory: Retention without remembering," *Am. Psychol.* **45**, 1043–1056.
- Schacter, D. L., and Church, B. A. (1992). "Auditory priming: Implicit and explicit memory for words and voices," *J. Exp. Psychol. Learn. Mem. Cogn.* **18**, 915–930.
- Schacter, D. L., Church, B. A., and Treadwell, J. (1994). "Implicit memory in amnesic patients: Evidence for spared auditory priming," *Psychol. Sci.* **5**, 20–25.
- Sheldon, S., Pichora-Fuller, M. K., and Schneider, B. A. (2008). "Priming and sentence context support listening to noise-vocoded speech by younger and older adults," *J. Acoust. Soc. Am.* **123**, 489–499.
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (2012). "Predictive top-down integration of prior knowledge during speech perception," *J. Neurosci.* **32**(25), 8443–8454.
- Studebaker, G. A., Pavlovic, C. V., and Sherbecoe, R. L. (1987). "A frequency importance function for continuous discourse," *J. Acoust. Soc. Am.* **81**, 1130–1138.
- Tulving, E., and Schacter, D. L. (1990). "Priming and human memory systems," *Science*. **247**, 301–306.
- Whitmal, N. A., III, and DeRoy, K. (2011). "Adaptive bandwidth measurements of importance function for speech intelligibility prediction," *J. Acoust. Soc. Am.* **130**, 4032–4043.
- Wu, M., Li, H., Gao, Y., Lei, M., Teng, X., Wu, X., and Li, L. (2011). "Adding irrelevant information to the content prime reduces the prime-induced unmasking effect on speech recognition," *Hear. Res.* **283**, 136–143.
- Wu, M., Li, H., Hong, Z., Xian, X., Li, J., Wu, X., and Li, L. (2012). "Effects of aging on the ability to benefit from prior knowledge of message content in masked speech recognition," *Speech Commun.* **54**, 529–542.
- Yang, Z., Chen, J., Huang, Q., Wu, X., Wu, Y., Schneider, B. A., and Li, L. (2007). "The effect of voice cuing on releasing Chinese speech from informational masking," *Speech Commun.* **49**, 892–904.