
The Effectiveness of Clear Speech as a Masker

Lauren Calandruccio
Kristin Van Engen
Sumitrajit Dhar
Ann R. Bradlow

Northwestern University, Evanston, IL

Purpose: It is established that speaking clearly is an effective means of enhancing intelligibility. Because any signal-processing scheme modeled after known acoustic-phonetic features of clear speech will likely affect both target and competing speech, it is important to understand how speech recognition is affected when a competing speech signal is also spoken clearly. In 2 experiments, the authors investigated whether listeners would experience improved intelligibility when both target and nontarget speech were spoken clearly.

Method: Listeners' recognition of sentences in competing sounds was examined in 2 experiments. For both experiments, the target speech was spoken in conversational and clear styles. The competing sounds in Experiment 1 included 2-talker maskers spoken in conversational and clear styles of English or Croatian. The competing sounds in Experiment 2 included 1-talker maskers spoken in clear or conversational styles and temporally modulated white noise maskers shaped to mimic the 1-talker maskers.

Results: Performance increased for clear versus conversational targets. No significant differences were found between conversational and clear maskers.

Conclusions: If it were possible to implement clear speech through a listening device, it appears that listeners would still receive a clear-speech benefit, even if all sounds (including competing sounds) were (inadvertently) processed to be more clear.

KEY WORDS: clear speech, speech perception, informational maskers

When a talker is asked to speak clearly, the acoustical changes in the subsequent speech production help improve listeners' recognition of speech (see, e.g., Helfer, 1997; Payton, Uchanski, & Braida, 1994; Picheny, Durlach, & Braida, 1985; Schum, 1996). The superior intelligibility of clear speech has been repeatedly documented in the literature, with improvements in recognition as high as 26 percentage points between clear and conversational speech targets (Payton et al., 1994). This clear-speech advantage is robust in both quiet and adverse listening conditions, including increased reverberation and competing steady-state or speech noise at various signal-to-noise ratios (SNRs; for a comprehensive review of clear-speech production and perception, see Smiljanic & Bradlow, 2009). The advantage is available to listeners even when talkers are trained to speak clearly at a conversational speaking rate (Krause & Braida, 2002). This clear-speech benefit has also been reported for many different listener groups, including listeners with hearing loss (e.g., Picheny et al., 1985), children with learning disabilities (Bradlow, Kraus, & Hayes, 2003), older listeners (e.g. Schum, 1996), and listeners who are not native speakers of English (Bradlow & Alexander, 2007; Bradlow & Bent, 2002).

One long-term aim of clear-speech research is to incorporate some of the acoustical properties of clear speech into signal-processing strategies (in hearing aids, cochlear implants, and other listening devices) to

improve listeners' recognition of speech in both quiet and noise. However, the efficacy of transforming the real-world speech environment into clear speech is largely unknown. For example, one rarely encounters speech in the real world without competition, be it competition from other talkers or nonspeech environmental sounds. Therefore, it is necessary to understand what happens to listeners' recognition of speech not only when the target or intended speech signal becomes more clear but also when all of the sounds in the environment (received by the listening device) undergo the same signal processing. In this article, we report the results of a series of experiments that are a first attempt to systematically investigate the speech recognition of listeners with normal hearing when both the target (intended) and the nontarget (competing) speech are spoken clearly.

We propose two alternative hypotheses for how clear-speech maskers could affect listeners' speech understanding. The first hypothesis is that clear speech would be a more effective masker. By this, we mean that it would be more difficult to recognize the target speech signal when the competing background speech is in a clear-speech style than when it is in a conversational speech style. This prediction is based on the well-established greater intelligibility of clear speech relative to conversational speech. This greater intelligibility may cause specific lexical items in the masker speech to "pop out" from the background, thereby creating a greater distraction from the target speech for the listener. In this scenario, the clear-speech masker would have greater linguistic informational masking influences. This hypothesis should hold true only when the masker is in a language that is familiar to the listener (e.g., English, a language that is known to our listener population). Therefore, as an additional means of contrasting these two hypotheses, in Experiment 1 we included both English and Croatian speech maskers. Croatian was selected as the unfamiliar language because clear speech spoken in Croatian has been shown to produce similar perceptual benefits for Croatian listeners and to be characterized by global acoustic properties similar to English clear speech (Smiljanic & Bradlow, 2005). In addition, few listeners from our subject demographic area would have knowledge of Croatian. Thus, if clear speech is a more effective masker because of greater lexical salience, the increased masking should occur only in the presence of the English clear-speech maskers. Croatian conversational and clear-speech maskers, in contrast, should be equally effective. However, if clear speech is a more effective masker in both English and Croatian, this could indicate that other complex acoustic-phonetic differences that are present between clear and conversational speech—such as slower speech rate, longer and more frequent pauses, greater pitch range, greater vowel space area, and greater vowel space dispersion reported for

both languages (Smiljanic & Bradlow, 2005)—could be responsible for the effectiveness of the clear-speech masker.

The alternative hypothesis is that clear speech is a less effective masker. In other words, it would be easier to recognize the target speech signal with competing clear speech compared with conversational speech in the background. This prediction is based on the greater temporal modulations present within a clear-speech signal (Krause & Braida, 2004, 2009). We predicted that listeners could potentially "listen within the dips" (Festen & Plomp, 1990) of these exaggerated temporal modulations to gain glimpses of the target speech signal and improve their speech recognition relative to speech embedded in a conversational masker. In this case, the clear-speech masker would have less of an energetic influence compared with the conversationally spoken speech masker. To probe this idea further, in Experiment 2 we included white noise maskers that were temporally modulated to match the temporal envelopes of the clear and conversational speech maskers also used in Experiment 2. If listeners were able to take greater advantage of the temporal modulations in the clear maskers in comparison to the conversational maskers used in Experiment 2, then noise maskers modulated to similar clear-speech envelope patterns should also prove to be less effective than their conversational speech counterparts. Comparing temporally modulated noise maskers with those composed of actual speech also provides a means of isolating the contributions of energetic versus informational masking from the various speech maskers.

We tested all of the masker conditions across two SNRs, with the easier SNR in the first half of the experiment and the harder SNR in the second half of the experiment. This was done for two reasons: (a) as a means of avoiding ceiling or floor effects that may be expected on the basis of previous informational masking experiments (e.g., Freyman, Helfer, & Balakrishnan, 2007; Kidd, Mason, Deliwala, Woods, & Colburn, 1994; Van Engen & Bradlow, 2007) and (b) to ensure that any practice effects observed within the experiment would be counterbalanced by the increased difficulty of the second SNR condition presented in the second half of the experiments.

Experiment 1: Two-Talker Maskers

Method

Subjects. Thirty adult listeners with normal hearing (17 women and 13 men; mean age: 21;7 [years;months]) participated in Experiment 1. All listeners were native monolingual speakers of American English with no knowledge of Croatian. The institutional review board

at Northwestern University approved all procedures. Listeners were paid for their participation and provided written informed consent.

Otoscopic evaluations were performed to ensure clear ear canals for all participants. Listeners' audiometric thresholds were tested using a Maico M26 clinical audiometer. All listeners had air conduction thresholds equal to or less than 20 dB HL between 250 and 8000 Hz, bilaterally (American National Standards Institute, 2004).

Stimuli. We used the Bamford–Kowal–Bench (BKB; Bench, Kowal, & Bamford, 1979) Revised Sentence Lists 1–16 for target sentences in all experiments. Each BKB list has 16 sentences and a total of 50 key words. The 16 sentences in each list contain either three or four scoreable key words. An example of a three-key word BKB sentence is “The orange is very sweet” (key words are underlined). Two versions of these sentence lists were recorded by a female talker at Northwestern University. The first version was recorded with the talker instructed to speak in a “typical” or “conversational” style of speech. For the second version of the sentences, the talker was instructed to speak “clearly,” or how she may naturally speak if she knew the listener was having a difficult time understanding her. Each sentence from the two versions of the BKB lists was root-mean-square (RMS) normalized to the same pressure. We did all RMS normalization for the stimuli in this project using Praat (Boersma & Weenink, 2009).

Four distinct two-talker female babble strings were used throughout testing as competing maskers. Two of the two-talker female babble strings were spoken in English by native English speakers, and two were spoken in Croatian by native Croatian speakers. The competing English masker speech consisted of each talker speaking 20 meaningful English sentences taken from the list of sentences in Nazzi, Bertoncini, and Mehler's (1998) multilanguage corpus. The sentences had an average of 10 words per sentence (range: 9–14 words). An example

of one of the sentences is “The next local elections will take place during the winter.” Two female talkers (different from the English target talker) spoke these sentences using both a conversational style and a clear style of speech. The sentences spoken by these two females have been shown to produce a clear-speech advantage; that is, when listeners with normal hearing recognized the sentences spoken by these two talkers, performance significantly increased for the clear versus the conversational sentences (see Smiljanic & Bradlow, 2008, Talkers F6 and F2). Each sentence used to construct the babble strings was RMS normalized to the same pressure. Sentences were then concatenated in different orders for each talker without any silent intervals between sentences. Because both talkers spoke the same sentences, we carefully planned the order in which the sentences were concatenated to ensure that the two talkers never spoke the same sentence at the same time. Once each talker's sentence string was concatenated, we mixed together the two talkers' strings into one single audio file using Audacity software (<http://audacity.sourceforge.net/>).

The same methods used to create the English two-talker babble were used to create the Croatian two-talker babble. The same sentences used in the two-talker English babble described earlier were translated into Croatian (initially for the purposes of a separate study but made available to this study by a native Croatian-speaking colleague). The Croatian translations were specifically matched to the sentences used for the English two-talker babble for the number of expected syllables (15–18/sentence). All four competing babble maskers were at least 48 s in length (the overall duration varied depending on the individual talkers and clear vs. conversational styles; see Table 1). After concatenation of the sentences, all maskers were again RMS normalized to the same pressure level. Only the first 48 s of each babble string were used as competing maskers.

Table 1. Speaking rate, number of pauses, and pause duration of the speech produced by the five talkers used in Experiments 1 and 2.

| Talker | Gender | Speaking rate (wpm) | | Average # of pauses/sentence | | Average pause duration(s) | |
|--------|--------|---------------------|--------|------------------------------|-------|---------------------------|-------|
| | | Conv. | Clear | Conv. | Clear | Conv. | Clear |
| T1 | F | 192.58 | 127.95 | 1.98 | 3.45 | 0.07 | 0.12 |
| EM1 | F | 274.68 | 222.04 | 5.30 | 5.70 | 0.06 | 0.08 |
| EM2 | F | 289.99 | 269.98 | 2.74 | 7.00 | 0.08 | 0.09 |
| CM1 | F | 198.67 | 160.02 | 3.00 | 4.60 | 0.06 | 0.05 |
| CM2 | F | 177.22 | 147.10 | 4.55 | 6.10 | 0.07 | 0.09 |

Note. wpm = words per minute; Conv. = conversational; T1 = target talker, EM1 and EM2 = English Masker Talker 1 and 2, respectively; CM1 and CM2 = Croatian Masker Talker 1 and 2, respectively.

The speaking rate, number of pauses, and pause duration are reported in Table 1 for the speech produced by the five talkers used in these experiments. Rate was defined as the number of words spoken per minute, excluding pauses (defined as silent periods 10 ms or longer; see Krause & Braida, 2002, for a similar analysis). Picheny, Durlach, and Braida (1986) reported that a characteristic of clear speech is that it contains more and longer pauses (or gaps between words) compared with conversational speech. Although the insertion or deletion of these pauses alone has not been shown to account for intelligibility differences between clear and conversational speech (Uchanski, Choi, Braida, Reed, & Durlach, 1996), these pauses (or lack thereof) are a prominent difference between these two speaking styles. Therefore, also reported in Table 1 are the average number of pauses per sentence produced by each talker, and the average pause duration. It should be noted that the sentences used for the masker conditions were significantly longer in length compared with the target BKB sentences (Nazzi et al., 1998).

Procedure. Speech signals (the target and masker stimuli) were mixed in real time and output through custom software created using MaxMSP (Cycling '74, 2008) running on an Apple Macintosh computer. A MOTU 828 MkII input/output firewire output device was used for digital-to-analog conversion (24 bit). Generated signals were passed through a Behringer Pro XL headphone amplifier and passed to MB Quart 13.01HX drivers. The signals from these drivers were delivered to both ear canals using an assembly similar to the one used in clinical audiometry (i.e., via plastic tubes and disposable 13-mm foam eartips). For each trial, one BKB sentence was played. A random portion of the appropriate babble masker was presented 1 s longer than the target sentence (500 ms prior to the beginning of the sentence and 500 ms at the end of the sentence). The overall RMS level of the target speech stimuli was held constant at 70 dB SPL, and the masker speech varied in level around the target speech, depending on the desired SNR.

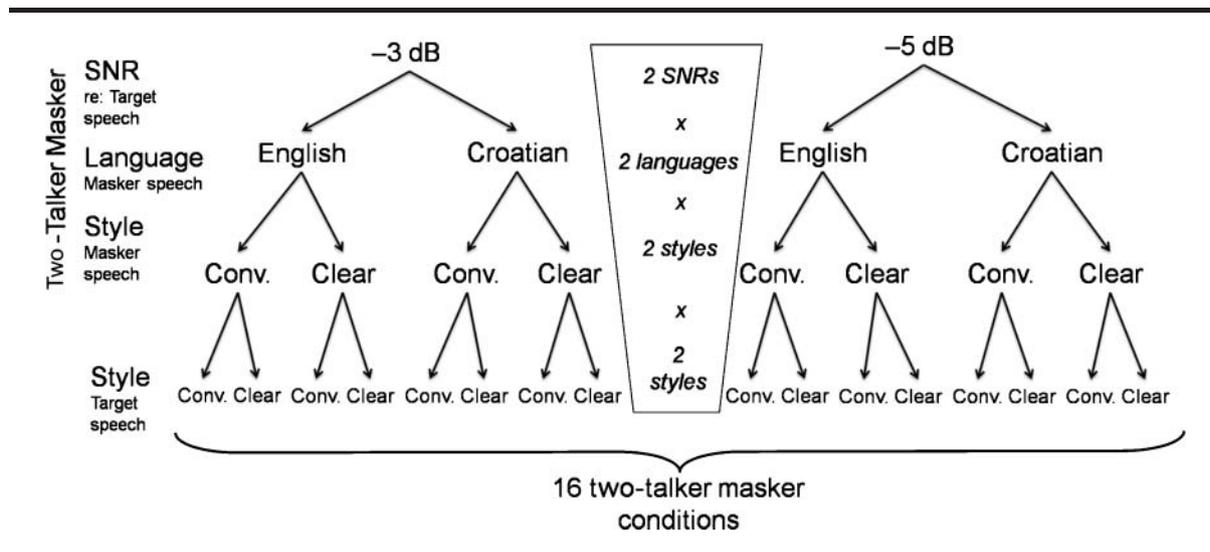
Listeners were instructed to listen for the target female talker and repeat back what she said. Listeners' responses were scored online for the number of key words correct and digitally recorded using an Olympus digital voice recorder. All listener responses were rescored offline by an independent second examiner for reliability measures. Items on which online and offline scores disagreed were replayed to the two examiners, who then scored those items by consensus. Such disagreement and subsequent resolution happened in 1% of the total trials scored.

Listeners were given an opportunity to familiarize themselves with the task and the target talker's voice prior to the beginning of each experiment. Each listener

was presented with 16 sentences (50 key words) from BKB List 21. During familiarization, a favorable SNR of 10 dB was used for the first eight sentences, and an SNR of 5 dB was used for the second eight sentences. Eight of the sentences were presented in the presence of Croatian maskers, and the other eight sentences were presented in the presence of English maskers. The style of the target and the style of the masker were randomly varied across the 16 sentences. All combinations of target style and masker style were presented during familiarization (i.e., with respect to target-masker combinations for both languages and both SNRs, the following were played one time each: conversational-clear, conversational-conversational, clear-conversational, clear-clear).

During Experiment 1, 16 conditions were tested on all subjects: 2 SNRs (-3 and -5 dB) × 2 masker languages (Croatian and English) × 2 styles of masker speech (conversational and clear) × 2 styles of target speech (clear and conversational; see Figure 1 for a schematic illustration of Experiment 1). One list of the BKB sentences (randomly chosen from Lists 1-16) was used for each test condition. We quasi-randomized the presentation of these conditions across subjects. The easier of the two SNR conditions (-3 dB; the first eight conditions) was always presented before the more difficult SNR condition (-5 dB; the second eight conditions). These two SNRs were chosen on the basis of pilot data analyzed to determine the SNRs needed to obtain usable data from most of our listeners (without reaching ceiling or floor effects) in all of the listening conditions (because many of the variables changed performance levels significantly, e.g., clear vs. conversational speech; Croatian vs. English speech maskers). Croatian maskers were always presented prior (Conditions 1-4 and 9-12) to the English maskers for each SNR (Conditions 5-8 and 13-16). This ordering was based on the results of previous literature indicating that, in comparison to their native language, monolingual listeners enjoy a release in masking when an unfamiliar language is competing in the background (see Freyman, Balakrishnan, & Helfer, 2001; Garcia Lecumberri & Cooke, 2006; Tun, O'Kane, & Wingfield, 2002; and Van Engen & Bradlow, 2007). We also had conducted pilot testing in our laboratory to ensure that, on average, the Croatian masker was easier in comparison to the English masker for native English monolingual listeners. The listeners used in the pilot testing were not included in this study and thus are not discussed in this article. Because of the ordering of the SNR conditions and masker languages, there were always four blocks presented in a specific order within the 16 conditions (Croatian-3, English-3, Croatian-5, and English-5). With this ordering, the more difficult conditions (lower SNR, English masker) always occurred after the easier conditions (higher SNR, Croatian masker) as a means of counteracting any potential

Figure 1. Schematic illustration of the design of Experiment 1. Listeners were tested across 16 different listening conditions that included Croatian and English two-talker masker conditions, spoken in clear or conversational (Conv.) styles. The target speech also varied between clear and conversational styles. Stimuli were presented to the listeners at two signal-to-noise ratios (SNRs; -3 and -5 dB).



learning effects and to avoid ceiling effects for the easier listening conditions (see also Van Engen & Bradlow, 2007). Within those four blocks, the order of presentation of the target and masker styles were randomized for each listener. The four conditions (target style–masker style) included conversational–conversational, conversational–clear, clear–conversational, and clear–clear.

Results

Throughout this article, we have transformed listeners' performance scores for all conditions to rationalized arcsine units to stabilize error variance (Studebaker, 1985) prior to statistical analyses. A $2 \times 2 \times 2 \times 2$ repeated measures analysis of variance (ANOVA) that examined all within-subject effects (SNR, language of masker speech, style of masker speech, style of target speech) indicated a significant main effect of SNR, language of the masker speech, and style of the target speech: $F(1, 29) = 42.08$, $p < .001$; $F(1, 29) = 14.69$, $p = .001$; and $F(1, 29) = 108.54$, $p < .001$, respectively. In other words, listeners always performed significantly better when the SNR was -3 dB compared with -5 dB, when the masker speech was spoken in Croatian compared with English, and when the target speech was spoken clearly compared with conversationally (see Figure 2). The main effect of style of the masker speech was not significant, $F(1, 29) = 3.18$, $p = .085$.

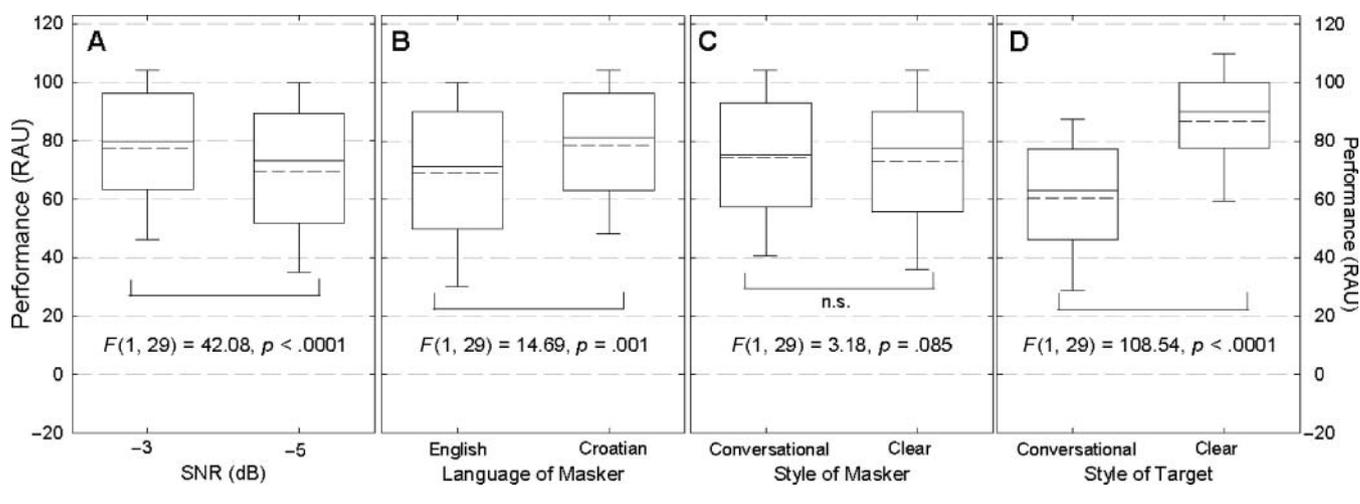
We also observed two significant interactions in the four-way repeated measures ANOVA. We noted a significant interaction between the language of the masker

and the style of the target speech, $F(1, 29) = 5.46$, $p = .027$. This interaction revealed that, overall, listeners benefited significantly more from a clear speech target when the competing speech was spoken in English compared with when it was spoken in Croatian.

The second significant interaction we observed was for the language of the masker and the style of the masker speech, $F(1, 29) = 34.14$, $p < .001$. This interaction indicated that, overall, listeners identifying speech in the presence of English maskers performed significantly better in the presence of a clear-speech masker than when the masker was spoken conversationally. When the competing speech was Croatian, however, the opposite effect was observed; that is, listeners' performance in the presence of the Croatian masker significantly decreased when the competing Croatian speech was spoken clearly.

To probe the interaction of language of the masker and style of the masker speech further, we conducted post hoc pairwise comparisons with a Bonferroni multiple comparison correction ($\alpha = .05/28 = .0018$), which indicated no significant differences in masker styles within individual masker language conditions (as indicated in Figure 3 by the "n.s." brackets), $ts(29)$ ranging between 3.28 and 0.10, ps ranging between .921 and .007; that is, no significant differences were found for style of the masker speech when all other within-subject factors were held constant (e.g., those conditions tested with the same SNR, the same style of the target speech, and the same language of the masker speech).

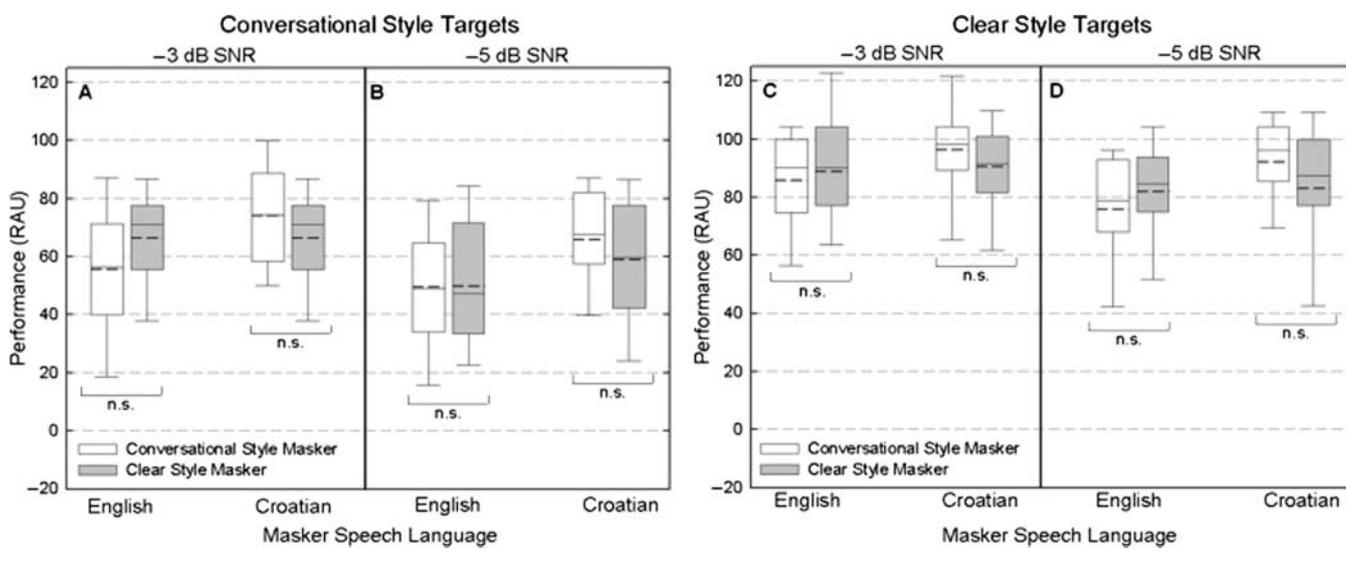
Figure 2. Rationalized arcsine unit (RAU) performance results for the 30 listeners who participated in Experiment 1. Panel A: RAU performance for all -3- and -5-dB SNR test conditions. Panel B: RAU performance for all conditions in which the masker was spoken in English versus Croatian. Panel C: RAU performance for all conditions in which the masker was spoken in a conversational style versus a clear style. Panel D: RAU performance for all conditions in which the target speech was spoken in a conversational versus a clear style. The dashed lines within the box plots in all panels represent the mean, and the solid lines indicate the median.



In summary, listeners performed better when the target speech was spoken clearly, regardless of the masker condition or the SNR. This clear-speech benefit was greater when the competing speech was spoken in English compared with Croatian. It should be noted that listeners' performances decreased when the masker speech was English compared with Croatian; therefore, the increased clear-speech benefit in the English

masker conditions may have to do with significantly poorer recognition in the conversational target listening conditions, allowing more room for improvement when provided with a clear target signal. There was no difference in performance between the clear and conversational style maskers within each test condition (i.e., within each of the four panels of Figure 3 in the same masker language).

Figure 3. Data from the 16 listening conditions in Experiment 1. Post hoc analyses using a Bonferroni correction indicated no significant differences in performance between conversational style maskers (open box plots) and clear style maskers (gray box plots; as indicated by the n.s. [not significant] brackets).



Experiment 2: One-Talker and Modulated-Noise Maskers

Method

The purpose of Experiment 2 was to further examine the effect of clear speech as a masker and to determine whether any of the results we noted in Experiment 1 were due to two talkers speaking simultaneously in the masking conditions. Specifically, we wanted to ensure that we were not losing any of the modulation differences between the clear and conversational masker speech, and hence the clear-speech masker advantage or disadvantage, by combining two talkers. We potentially could be losing some of the greater modulation depths commonly observed in clear speech by overlaying the two speaker's voices. Therefore, in this experiment, we included two different one-talker maskers: one clear and one conversational. Also, to eliminate linguistic influences but continue to evaluate differences in modulation between the two types of maskers, we included two temporally modulated white noise maskers shaped to match the one-talker clear and conversational masker speech signals.

Subjects. Twenty adult listeners with normal hearing (13 women and 7 men; mean age: 23;4) who did not take part in Experiment 1 participated in Experiment 2. Again, all listeners were native monolingual speakers of American English, had normal hearing, were paid for their participation, and provided written informed consent. Otoscopic and audiometric evaluations were performed as described in Experiment 1.

Stimuli. The same versions and lists of the BKB sentences used in Experiment 1 were used in Experiment 2. Four new masker conditions were tested in Experiment 2. The first two masker conditions consisted of a single female native English talker speaking the same meaningful sentences described in Experiment 1 (Talker

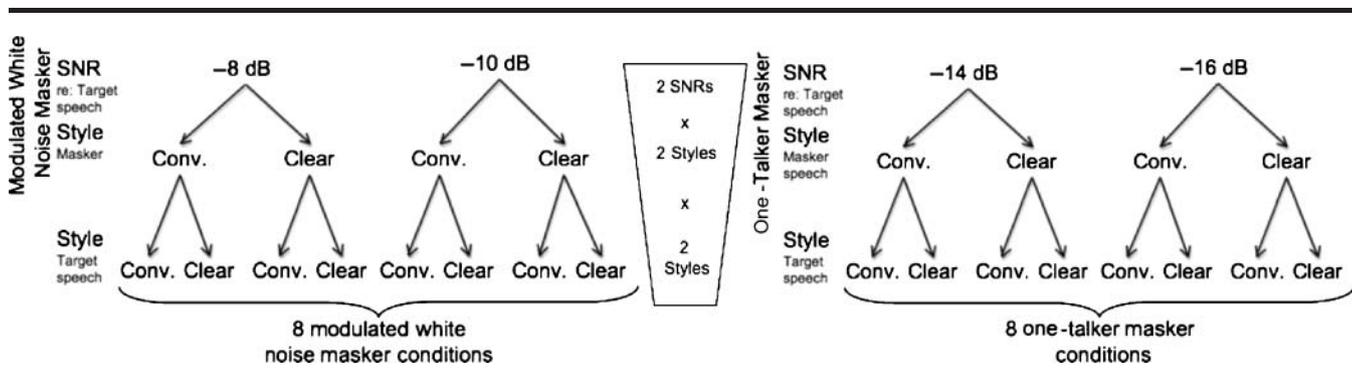
EM1). The first masker condition was composed of concatenated sentences spoken in a clear style, and the second condition contained the same sentences spoken in a conversational style.

We created the second two masker conditions using the temporal envelopes of the one-talker masker strings (i.e., the concatenated sentences); specifically, the envelopes of the clear one-talker masker and the conversational one-talker masker were computed in MATLAB. We applied a full-wave rectification Hilbert transform to the stimuli, which were then low-pass filtered using a rectangular filter with a cutoff frequency equal to 50 Hz and a sampling frequency of 22.1 kHz (see Davidson, Gilkey, Colburn, & Carney, 2006). A Gaussian white noise, generated in MATLAB, with a sampling frequency of 22.1 kHz, was multiplied by the two envelopes to create temporally modulated one-talker shaped noise maskers (one for the clear masker and one for the conversational masker).

Because of the slower speaking rate of the clear-speech stimuli, the clear style maskers were trimmed in length so that all four maskers were 40 s long.

Procedure. The procedures used to output the stimuli, the instructions given to the listeners, and the scoring technique used in Experiment 2 were identical to those in Experiment 1. Listeners were tested on a total of 16 different listening conditions. The first eight listening conditions included modulated white noise maskers, and the second eight conditions included one-talker maskers (2 SNRs \times 2 styles of maskers \times 2 styles of target speech for 2 different types of maskers; see Figure 4 for a schematic illustration of the overall design). The SNRs used for the modulated-noise masker conditions and the one-talker masker conditions varied from each other (-8 and -10 dB vs. -14 and -16 dB, respectively) and from those used in Experiment 1 (-3 and -5 dB). We did this in an attempt to approximate equal performance across the different masker conditions because some of the maskers used are inherently less effective (easier) than others (see Simpson & Cooke, 2005).

Figure 4. Schematic illustration of the design of Experiment 2. All modulated white noise maskers were presented prior to the one-talker masker listening conditions. For both maskers, the easier of the two SNRs was always presented prior to more difficult SNR conditions.



The two masker types (modulated white noise and one talker) were presented in separate tasks. All test conditions with the two modulated-noise maskers (eight conditions) were conducted first, followed by all of the eight one-talker masker conditions, to allow the listeners to become familiar with the target voice. To familiarize listeners with the recognition task in the presence of the modulated-noise maskers, we presented them eight sentences (25 key words) from List 21 of the BKB sentence list. The first four sentences were presented at 5 dB SNR, and the last four sentences were presented at 0 dB SNR. The style of the masker conditions (clear or conversational) and the style of the target speech (clear or conversational) were randomly varied. Listeners were tested on one random BKB list/condition (chosen from Lists 1–16, 50 key words/condition). Similar to Experiment 1, we quasi-randomized the eight listening conditions across subjects; that is, the easier of the two SNR conditions (–8 dB) was always presented first, and the more difficult SNR (–10 dB) condition was always presented last. The styles of the modulated-noise maskers and the style of the target speech were randomized across all subjects. Listeners were scored on the number of key words correct.

To familiarize listeners with the one-talker masker condition task, we presented them with the other eight sentences (25 key words) from List 21 of the BKB sentence list (the other eight sentences were used to familiarize listeners with the modulated-noise masker task). The first four sentences were presented at 5 dB SNR, and the last four sentences were presented at –5 dB SNR. The style of the masker conditions (clear or conversational) and the style of the target speech (clear or conversational)

were randomly varied. Again, the eight listening conditions were quasi-randomized across subjects (the easier of the two SNR conditions [–14 dB] was always presented first, and the more difficult SNR [–16 dB] condition was always presented last). Listeners were tested on one random BKB list/condition (50 key words/condition). The styles (clear or conversational) of the one-talker maskers and the style of the target speech were randomized across all subjects.

Results

Modulated-noise maskers. A $2 \times 2 \times 2$ repeated measures ANOVA with all within-subject factors, including SNR (–8 and –10 dB), style of the target speech (clear and conversational), and style of the masker speech (clear and conversational), indicated a significant main effect of both SNR, $F(1, 19) = 37.69, p < .001$, and style of the target speech, $F(1, 19) = 312.46, p < .001$, for the modulated-noise masker conditions. No significant main effect of the style of the masker speech or any significant interactions were observed (see Figure 5).

One-talker maskers. We used a $2 \times 2 \times 2$ repeated measures ANOVA to examine all within-subject factors—including SNR (–14 and –16 dB SNR), style of the target speech (conversational and clear), and style of the masker speech (conversational and clear)—for the one-talker masker conditions. The analysis indicated a main effect of the style of the target speech, $F(19) = 240.72, p < .001$. No significant main effects of SNR or the style of the masker speech, or any significant interactions, were observed (see Figure 6).

Figure 5. Overall RAU performance for 20 listeners for the modulated-noise maskers used in Experiment 2. Panel A: RAU performance for all –8- and –10-dB SNR test conditions. Panel B: RAU performance for all conditions in which the target speech was spoken conversationally versus clearly. Panel C: RAU performance for all conditions in which the masker speech was modulated to mimic conversational versus clear speech. The dashed lines within the box plots represent the mean data, and the solid lines represent the median.

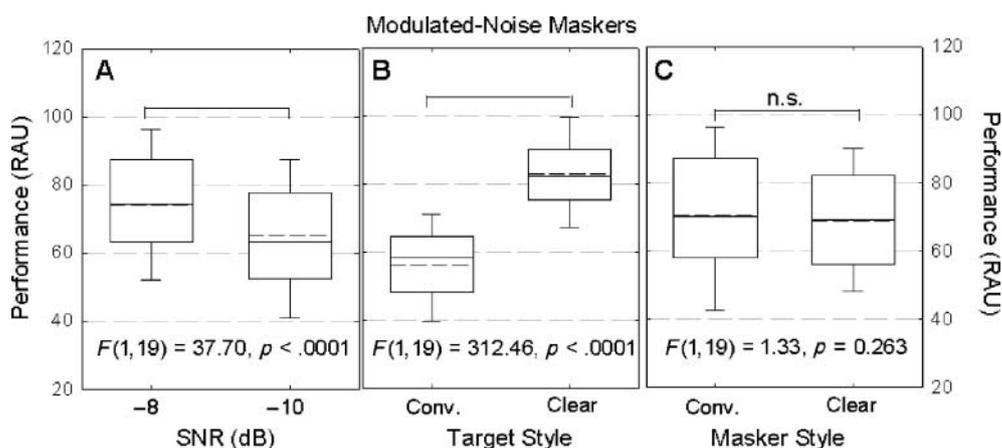
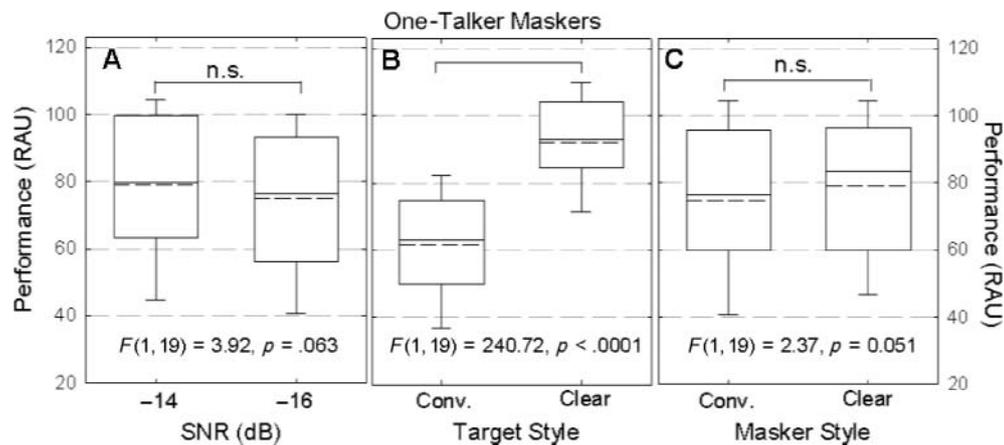


Figure 6. Overall sentence recognition performance (RAU) for 20 listeners for the one-talker maskers used in Experiment 2. Panel A: RAU performance for all of the -14- and -16-dB SNR conditions. Panel B: RAU performance for all of the conditions in which the target speech was spoken conversationally versus clearly. Panel C: RAU performance for all of the conditions in which the masker speech was spoken in a conversational versus clear style. The dashed lines within the box plots represent the mean data, and the solid lines depict the median.



In summary, listeners' performance was significantly improved when the target speech was spoken clearly regardless of masker type or SNR. For the modulated white noise maskers, listeners' performance significantly decreased in the harder SNR condition, but there was no significant difference in performance across the two SNRs for the one-talker masker conditions. Performance was not significantly affected by the masker style (conversational vs. clear).

General Discussion

These data represent a first attempt to examine listeners' speech recognition performance when not only the intended, or target, speech is spoken clearly but also when other competing signals in the listening environment are spoken clearly. Native listeners of English with normal hearing have always performed significantly better on a sentence recognition task when the target speech was spoken clearly compared with conversationally (see, e.g., Picheny et al., 1985, for a similar result). Also, listeners performed significantly better when the masker speech was spoken in Croatian compared with English (see Garcia Lecumberri & Cooke, 2006; Rhebergen, Versfeld, & Dreschler, 2005; and Van Engen & Bradlow, 2007, for similar results with different languages). No significant differences in performance were observed when the masker varied between a conversational and a clear style. This was true for two-talker English and Croatian, one-talker English, and modulated-noise maskers.

Conversational Versus Clear-Speech Targets

The improvement in intelligibility gained from clear speech has proven to be a very robust phenomenon. In fact, this clear-speech benefit has survived competition from wide-band noise (Bradlow & Bent, 2002), speech-shaped noise (Krause & Braida, 2002; Payton et al., 1994), and multitalker babble (Ferguson & Kewley-Port, 2002) and now single-talker, foreign speech, and modulated-noise maskers. This consistent benefit is crucial because listening in noise proves to be most difficult for listeners with hearing impairment (Bentler, Wu, Kettel, & Hurtig, 2008; Kochkin, 2005), and clear-speech signal-processing strategies may hold promise for listeners with hearing loss. However, if the acoustical changes that occur in a clear-speech signal can be implemented into a signal processing strategy, it would be very difficult, if not impossible, to constrain those changes only to the intended speech signal. This would be true even for those assistive listening devices that incorporate directional microphones and beam-forming technology because often both the target and competing sounds come from similar directions. Therefore, it is imperative that listeners still be able to obtain a clear-speech benefit with other clear signals competing in the environment. On the basis of the data reported in Experiments 1 and 2, listeners continue to achieve a clear-speech benefit regardless of the competing speech signal. We calculated clear-target speech benefit scores (quantified as the difference between clear-speech performance score and conversational speech performance score [measured as percentage correct], holding all other variables constant) for both SNRs for the three types of masker conditions (two talker, one talker, and

modulated noise). Regardless of the masker type or the SNR, listeners received, on average, a minimum of a 22-percentage-point increase (for the -8 SNR, clear modulated-noise masker condition) and a maximum of a 35-percentage-point increase (for the -16 SNR conversational one-talker masker condition). These degrees of improved speech recognition in noise are dramatically greater than what is typically reported for hearing-aid users after audibility has been accounted for (i.e., through the use of directional microphone technology, digital noise reduction algorithms, or frequency gain characteristics of the listening device; Bentler et al., 2008; Keidser et al., 2006; Klomp & Dhar, 2008). It is also encouraging that the clear-speech benefit was greater for more difficult listening conditions—that is, the more difficult SNRs and the masker conditions providing informational masking. The significant interaction observed between the language of the masker and the style of the target speech in Experiment 1 may suggest that listeners can take even greater advantage of clear speech either when the listening condition becomes more difficult or when informational masking causes confusion for the listener (when listening in competing English vs. Croatian speech). These results showcase the promise of clear speech as an avenue to better speech recognition in noise.

Effect of the SNR

As explained in the beginning of this article, the reason for using two different SNRs (with the easier SNR always presented first) throughout these experiments was twofold. First, we wanted to be able to account for the amount of variability observed in previous informational masking experiments (e.g., Freyman et al., 2007; Kidd et al., 1994; Van Engen & Bradlow, 2007); that is, having two different SNRs would allow us to have meaningful data for all of our listeners in the event that some of the listeners showed a ceiling effect in the easier listening conditions (i.e., easier SNR and/or easier masker condition; e.g., Croatian vs. English, conversational vs. clear) or a floor effect in one of the more difficult SNR conditions. Second, it is well known that listeners' experience with a task can affect their performance. Therefore, by including two SNRs, with the easier SNR presented first, we tried to ensure that any practice effects observed would be counterbalanced by the increased difficulty of the second SNR conditions in the second half of the experiment.

For the modulated white noise and two-talker masker conditions, listeners' performance significantly decreased in the more difficult SNR condition. However, for the one-talker masker condition, listeners' scores, on average, did not significantly decrease in the more difficult SNR condition. We acknowledge that the difference in SNRs between the "easier" and "harder" conditions used

in these experiments was only 2 dB. Using a greater range of SNRs could possibly have yielded significant differences between SNR test conditions for the one-talker maskers. However, this result is still noteworthy because it reinforces the observation that, when conducting speech-in-speech experiments with maskers that are heavily dominated by informational masking (e.g., one-talker maskers; see Simpson & Cooke, 2005), one must consider perceptual learning within the experimental task (Kidd, Mason, & Richards, 2003; Lutfi, 1990). Although it is normal and expected for listeners to improve their recognition over the course of a task, different degrees of learning/improvement or adaptation may occur as a result of varying amounts of informational masking across different speech masker conditions (e.g., one-talker vs. two-talker vs. eight-talker maskers). This being said, in our results we did not see a significant interaction between the SNR and the language of the masker in Experiment 1. The lack of this significant interaction may suggest that the spectral and temporal density of the two-talker maskers causes enough energetic masking that listeners either require more time to learn/adapt to the task or simply adapt less *because* of the energetic masking contributions when compared with the one-talker masker conditions.

Van Engen and Bradlow (2007) reported improvements in performance across trials for speech-in-speech recognition tasks. They examined English sentence recognition in the presence of English and Mandarin speech maskers for monolingual native English listeners. They tested two groups of listeners in two SNR conditions in which the easier of the two SNRs was presented first. The first group performed the recognition task at SNRs of 5 and 0 dB. The second group of listeners performed the recognition task at SNRs of 0 and -5 dB. Their results indicated a significant difference in performance between the two groups at the 0-dB SNR condition; that is, those listeners who performed the 0-dB SNR condition in the second half of the experiment performed better at 0 dB SNR than those listeners who began the experiment at 0 dB SNR. These results suggest that the improvements observed in their speech-in-speech data were probably due to the time it took for listeners to gain either an understanding of the listening task or familiarity with the target voice. It is reasonable to conclude that, of the listening conditions included in the present experiments, learning which voice was the target voice would be most difficult and would take the most time in the one-talker masker conditions. However, it should be noted that in Experiment 2, we always had the listeners perform the modulated-noise masker conditions first (before the one-talker masker conditions) to familiarize them with the target voice. Also, although on average listeners' performance significantly decreased in the more difficult SNR condition in the modulated-noise masker

conditions, in which there were no competing voices for the listener to identify, six of the 20 listeners improved their score in the more difficult SNR condition. This indicates that even when the listener does not need to identify the target voice from a competing background voice, listeners' performance still can improve over the course of an experiment. These data imply that the improvements in performance observed in our data are more likely due to the listeners gaining a better understanding of the listening task rather than becoming more familiar with the target voice. It should be noted, however, that in these experiments, the masker was chosen from a long stimulus file in which the starting point randomly varied on every trial. This method differed from that used by Van Engen and Bradlow, in whose study the masker noise was frozen (not changing from trial to trial). This methodological difference could impact how much, how quickly, and why listeners improve their performance across a number of trials (see Felty, Buchwald, & Pisoni, 2009).

Clear- Versus Conversational Speech Maskers

An important finding of these experiments is that the style of the masker speech did not significantly affect performance. We began these experiments with two alternative hypotheses. The first was that clear speech would be a more effective masker, meaning that it would be more difficult to recognize the target signal when a more salient clear-speech signal was competing in the background. The second, alternative hypothesis was that clear speech would be a less effective masker, meaning it would be easier to recognize the target signal when a clear-speech signal, with more exaggerated temporal modulations, was competing in the background. On the basis of the results of all four masker types (two-talker English, two-talker Croatian, one-talker English, and modulated white noise), we can reject both hypotheses. Furthermore, we can conclude that the hypotheses are not rejected simply because they cancel each other out; that is, on the basis of the results from some of the masker conditions (e.g., the white noise maskers that would not have any linguistic content), we can conclude that, for the stimuli used in these experiments, the difference in the dips within the clear- and conversational speech signals is not enough to improve or degrade listeners' speech recognition. Moreover, our results indicate that clear speech is neither a more nor less effective masker than conversational speech. It appears that the acoustic-phonetic differences that arise in a typical clear-speech signal, though robust when imposed on the target speech, are too small to change the clear-speech signal's effectiveness as a masker. This suggests two things: (a) listeners' perception of speech they are trying to recognize is different from their perception of speech they are

trying to ignore and (b) small acoustic-phonetic differences (e.g., like those commonly observed between clear and conversational speech) do not change the effectiveness of the masker speech. However, larger acoustic-phonetic differences (e.g., those observed across languages) do significantly change the effectiveness of the speech masker. This suggests that a range of acoustic-phonetic and/or linguistic variations in the masker speech can contribute to its effectiveness. Further research examining a broader spectrum of acoustic-phonetic variation in the masker speech is needed. For example, researchers have often observed that different talkers produce clear speech in different ways (e.g., Bradlow et al., 2003; Ferguson & Kewley-Port, 2007; Krause & Braida, 2004). It would therefore be advantageous in future experiments to include a variety of speakers as maskers. This is especially important when one considers the significant interaction observed between the language of the masker and the style of the masker speech in Experiment 1. Although post hoc analyses indicated no significant differences between clear and conversational maskers when all other factors were held constant (which leads us to reject our original hypotheses; see Figure 3), the data revealed a trend that listeners performed better when listening in the presence of a clear English speech masker compared with a conversational one. However, the opposite was observed for the Croatian clear-speech maskers. It is plausible to assume—on the basis of the interaction of the language of the masker and the style of the masker speech—that with different speakers competing, clear speech could help the listener perform the task. In other words, contrary to our original hypothesis, a clear signal that would “pop out” may not actually cause greater difficulty for the listener but may help the listener auditorily stream the information, making it easier to ignore the masker speech (or at least determine which signal the listener is trying to ignore). A greater number of speakers producing different degrees of clear-speech intelligibility would help probe this idea. For example, the English speakers used in the masking conditions in our experiments did produce a clear-speech advantage while producing target sentences in another experiment (see talkers F6 and F2 in Table 1 of Smiljanic & Bradlow, 2008). However, the clear-speech advantage for these speakers reported by Smiljanic and Bradlow (2008) was smaller compared with clear-speech advantages observed for other speakers that have been reported in the literature. Perhaps if the clear-speech advantage for the speakers used to create the maskers had been larger, we might have observed a significant effect of masking between clear and conversational speech maskers.

Effect of the Croatian Masker

Although not the focus of these experiments, we observed in Experiment 1 a release in masking in the

presence of competing speech maskers spoken in a language other than English. These results are in agreement with previous research that has explored maskers spoken in a language that differs from the language of the target speech (Garcia Lecumberri & Cooke, 2006; Rhebergen et al., 2005; Van Engen & Bradlow, 2007). Garcia Lecumberri and Cooke (2006) investigated the recognition of English consonants for native (English) and nonnative (Basque/Spanish) listeners in the presence of several competing masker conditions at an SNR of 0 dB. They reported a small but significant improvement ($\eta^2 = .063$) in performance by native English listeners when the competing speech masker was spoken in Spanish (a language unknown to the native group) compared with the English masker condition. Van Engen and Bradlow (2007) reported a similar release in masking when native (English) listeners performed a sentence recognition task in the presence of Mandarin two-talker babble compared with English two-talker babble.

There are several possible explanations for the observed release in masking that occurs when the masker speech is spoken in a language that is different than the target speech. First, the English masker could be providing greater energetic masking because the target and masker speech are more acoustically similar (both spoken in the same language) compared with the Croatian condition (in which the target speech was spoken in English). However, another possibility is that the listener is more distracted by the English masker, causing interference at a linguistic level (potentially because of syllables, words, or other linguistic units) and imposing greater informational masking compared with the Croatian speech. This explanation assumes that the energetic contributions of the English and Croatian maskers are similar and that the differences in masking occur not in the auditory periphery but at the central level. Garcia Lecumberri and Cooke (2006) and Van Engen and Bradlow (2007) both suggested that the differences observed in performance may be due to the differences in linguistic interference from the masker languages (a known language vs. an unknown language).

In the present study, the language effect (increased performance in the non-English masker) was robust against variation in the style of both the target and masker speech; that is, performance in the presence of the Croatian maskers was always better than performance in the presence of the English maskers regardless of target or masker speaking style. Although these data cannot help identify the source of the language effect as either energetic or informational, they establish that it is indeed robust. Additional data are needed to probe the role of energetic and informational masking when conducting speech-in-speech experiments in which the masker speech is not linguistically matched to the target speech. One way to probe this question would be to test listeners'

performance on a speech-in-speech recognition task when the masker speech varies in intelligibility (e.g., from native speech to highly accented foreign speech, all spoken in the same language).

Modulated-Noise Maskers

The modulated-noise maskers used in Experiment 2 were developed from the clear and conversational speech of one talker. As discussed earlier, including a variety of speakers will be necessary when examining the effectiveness of clear speech as a masker in future experiments. The results depicted in Figure 5 indicate that the modulation differences between the clear- and conversational speech maskers used in Experiment 2 did not significantly affect performance. However, as discussed earlier, it is well documented that different talkers adjust their speech production differently when speaking clearly (Bradlow et al., 2003; Ferguson & Kewley-Port, 2007; Smiljanic & Bradlow, 2005). Furthermore, increases in temporal modulation are only one of the many acoustic-phonetic features that change between conversational and clear speech. Some talkers who demonstrate more dramatic changes in temporal modulations between their clear and conversational speech may prove to be more or less effective in terms of masking, depending on the other features that change in their clear speech relative to their conversational speech.

A promising next step would be to assess the clear-speech benefit of several speakers and then use the envelope patterns of these speakers' clear and conversational speech as maskers. Determining both the intelligibility advantage (in terms of target speech) and the masking effectiveness of clear versus conversational speech of the same speakers would potentially provide a more complete view of how speaking style modifications affect both target speech intelligibility and masker effectiveness.

Testing listeners with hearing loss would also be a critical next step. The differences in temporal modulations between some speakers' clear and conversational speech may change the effectiveness of the masking for some listeners with normal hearing. It has been reported, however, that listeners with hearing impairment are less affected by amplitude modulations in competing noise than are listeners with normal hearing (Eisenberg, Dirks, & Bell, 1995). Therefore, it is reasonable to predict that listeners with hearing impairment may not be able to distinguish between the temporal differences in clear-versus conversational speech maskers.

Potential Limitations

Previous research has shown that there is a large amount of variability in how different speakers change

their speech to speak more clearly (e.g., Ferguson & Kewley-Port, 2007; Krause & Braida, 2004; Picheny et al., 1986). One possible reason we did not observe significant differences between styles of maskers is that we used only four different talkers throughout these experiments (two Croatian and two English). If a clear-speech signal-processing strategy were to be implemented, it would be based on average clear-speech production (not on an individual talker). The rate of the clear competing speech spoken by the four talkers used throughout these experiments was faster compared with the target talker's rate of speech. If the masker talkers we had used had a more exaggerated style of clear speech, the clear-speech maskers may have had a different impact on these results.

Second, the clear-speech targets used in these studies were RMS normalized to the same pressure level as the conversational speech stimuli; however, there were significantly greater and longer pauses in the clear speech. Therefore, the intensity of the clear speech itself (not the RMS of the entire sentence that included the longer and more frequent pauses) may have given these listeners an advantage when listening to the clear-speech stimuli by inadvertently manipulating the SNR in their favor. If this were true, however, the clear-speech masker would have caused an equal disadvantage in terms of SNR. Liu and Zeng (2006) reported that when the pauses were removed from the clear- and conversational speech stimuli used in their study, the RMS was 0.2 dB greater for the clear speech in comparison to the conversational speech. Of course, the 0.2-dB difference reported by Liu and Zeng was specific to the stimuli used in their study; however, such a small difference is unlikely to account for performance differences observed between clear and conversational speech in these experiments. In addition, if the RMS normalization was slightly biased toward providing greater energy for clear speech, it actually is more promising that clear speech would not be detrimental to the listener if competing in the background.

Conclusion

This study was a first attempt to determine how listeners' speech recognition would be affected if a clear-speech strategy were implemented on not only the target speech but also competing signals, as would be the case if an electronic device indiscriminately transformed all incoming signals. The results are promising in that listeners still demonstrated a significant clear speech benefit even when the competing speech was also clear.

Acknowledgments

This project was supported, in part, by an ASH Foundation New Investigator Grant, the Hugh Knowles Foundation at

Northwestern University, and National Institute on Deafness and Other Communication Disorders Grants R01-DC005794 and 1F31DC009516-01A1. Portions of these data were presented at the 2008 American Speech-Language-Hearing Association National Convention, Chicago, IL, and at the 2008 Auditory Perception, Cognition, and Action Meeting, Chicago. We thank Chun-Liang Chan and Yan (Felicia) Gai for help with software development. Rajka Smiljanic provided the Croatian translations and recorded materials. We also thank Nah Eun (NahNah) Kim and Christina Yuen for help with data collection and management and Rebekah Abel for helpful comments throughout these experiments.

References

- American National Standards Institute.** (2004). *Specifications for audiometers* (ANSI S3.6-2004). New York, NY: Author.
- Bench, J., Kowal, A., & Bamford, J.** (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, *13*, 108–112.
- Bentler, R., Wu, Y. H., Kettel, J., & Hurtig, R.** (2008). Digital noise reduction: Outcomes from laboratory and field studies. *International Journal of Audiology*, *47*, 447–460.
- Boersma, P., & Weenink, D.** (2009). Praat: Doing phonetics by computer (Version 5.1.07) [Computer software]. Retrieved from <http://www.praat.org/>.
- Bradlow, A. R., & Alexander, J. A.** (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, *121*, 2339–2349.
- Bradlow, A. R., & Bent, T.** (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America*, *112*, 272–284.
- Bradlow, A. R., Kraus, N., & Hayes, E.** (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research*, *46*, 80–97.
- Davidson, S. A., Gilkey, R. H., Colburn, H. S., & Carney, L. H.** (2006). Binaural detection with narrowband and wideband reproducible noise maskers: III. Monaural and diotic detection and model results. *The Journal of the Acoustical Society of America*, *119*, 2258–2275.
- Eisenberg, L. S., Dirks, D. D., & Bell, T. S.** (1995). Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. *Journal of Speech and Hearing Research*, *38*, 222–233.
- Felty, R. A., Buchwald, A., & Pisoni, D. B.** (2009). Adaptation to frozen babble in spoken word recognition. *The Journal of the Acoustical Society of America*, *125*, EL93–EL97.
- Ferguson, S. H., & Kewley-Port, D.** (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *112*, 259–271.
- Ferguson, S. H., & Kewley-Port, D.** (2007). Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of Speech, Language, and Hearing Research*, *50*, 1241–1255.
- Festen, J. M., & Plomp, R.** (1990). Effects of fluctuating noise and interfering speech on the speech-reception

- threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America*, 88, 1725–1736.
- Freyman, R. L., Balakrishnan, U., & Helfer, K. S.** (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America*, 109, 2112–2122.
- Freyman, R. L., Helfer, K. S., & Balakrishnan, U.** (2007). Variability and uncertainty in masking by competing speech. *The Journal of the Acoustical Society of America*, 121, 1040–1046.
- Garcia Lecumberri, M. L., & Cooke, M.** (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America*, 119, 2445–2454.
- Helfer, K. S.** (1997). Auditory and auditory–visual perception of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 40, 432–443.
- Keidser, G., Rohrseitz, K., Dillon, H., Hamacher, V., Carter, L., Rass, U., & Convery, E.** (2006). The effect of multi-channel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers. *International Journal of Audiology*, 45, 563–579.
- Kidd, G., Jr., Mason, C. R., Deliwalla, P. S., Woods, W. S., & Colburn, H. S.** (1994). Reducing informational masking by sound segregation. *The Journal of the Acoustical Society of America*, 95, 3475–3480.
- Kidd, G., Jr., Mason, C. R., & Richards, V. M.** (2003). Multiple bursts, multiple looks, and stream coherence in the release from informational masking. *The Journal of the Acoustical Society of America*, 114, 2835–2845.
- Klemp, E. J., & Dhar, S.** (2008). Speech perception in noise using directional microphones in open-canal hearing aids. *Journal of the American Academy of Audiology*, 19, 571–578.
- Kochkin, S.** (2005). MarkeTrak VII: Hearing loss population tops 31 million people. *Hearing Review*, 127, 16–29.
- Krause, J. C., & Braida, L. D.** (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *The Journal of the Acoustical Society of America*, 112(5, Pt. 1), 2165–2172.
- Krause, J. C., & Braida, L. D.** (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, 115, 362–378.
- Krause, J. C., & Braida, L. D.** (2009). Evaluating the role of spectral and envelope characteristics in the intelligibility advantage of clear speech. *The Journal of the Acoustical Society of America*, 125, 3346–3357.
- Liu, S., & Zeng, F. G.** (2006). Temporal properties in clear speech perception. *The Journal of the Acoustical Society of America*, 120, 424–432.
- Lutfi, R. A.** (1990). How much masking is informational masking? *The Journal of the Acoustical Society of America*, 88, 2607–2610.
- Nazzi, T., Bertoni, J., & Mehler, J.** (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 756–766.
- Payton, K. L., Uchanski, R. M., & Braida, L. D.** (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *The Journal of the Acoustical Society of America*, 95, 1581–1592.
- Picheny, M. A., Durlach, N. I., & Braida, L. D.** (1985). Speaking clearly for the hard of hearing: I. Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research*, 28, 96–103.
- Picheny, M. A., Durlach, N. I., & Braida, L. D.** (1986). Speaking clearly for the hard of hearing: II. Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, 29, 434–446.
- Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. A.** (2005). Release from informational masking by time reversal of native and non-native interfering speech. *The Journal of the Acoustical Society of America*, 118(3, Pt. 1), 1274–1277.
- Schum, D. J.** (1996). Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology*, 7, 212–218.
- Simpson, S. A., & Cooke, M.** (2005). Consonant identification in *N*-talker babble is a nonmonotonic function of *N*. *The Journal of the Acoustical Society of America*, 118, 2775–2778.
- Smiljanic, R., & Bradlow, A. R.** (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, 118(3, Pt. 1), 1677–1688.
- Smiljanic, R., & Bradlow, A. R.** (2008). Temporal organization of English clear and conversational speech. *The Journal of the Acoustical Society of America*, 124, 3171–3182.
- Smiljanic, R., & Bradlow, A. R.** (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compass*, 3, 236–264.
- Studebaker, G. A.** (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 455–462.
- Tun, P. A., O’Kane, G., & Wingfield, A.** (2002). Distraction by competing speech in young and older adult listeners. *Psychology and Aging*, 17, 453–467.
- Uchanski, R. M., Choi, S. S., Braida, L. D., Reed, C. M., & Durlach, N. I.** (1996). Speaking clearly for the hard of hearing: IV. Further studies of the role of speaking rate. *Journal of Speech and Hearing Research*, 39, 494–509.
- Van Engen, K. J., & Bradlow, A. R.** (2007). Sentence recognition in native- and foreign-language multi-talker background noise. *The Journal of the Acoustical Society of America*, 121, 519–526.

Received September 25, 2009

Revision received February 4, 2010

Accepted April 29, 2010

DOI: 10.1044/1092-4388(2010/09-0210)

Contact author: Lauren Calandruccio, who is now with the Department of Linguistics and Communication Disorders, Queens College of the City University of New York, Flushing, NY 11367. E-mail: lauren.calandruccio@qc.cuny.edu.

The Effectiveness of Clear Speech as a Masker

Lauren Calandruccio, Kristin Van Engen, Sumitrajit Dhar, and Ann R. Bradlow
J Speech Lang Hear Res 2010;53;1458-1471; originally published online Aug 5,
2010;
DOI: 10.1044/1092-4388(2010/09-0210)

The references for this article include 3 HighWire-hosted articles which you can
access for free at: <http://jslhr.asha.org/cgi/content/full/53/6/1458#BIBL>

This information is current as of December 6, 2010

This article, along with updated information and services, is
located on the World Wide Web at:
<http://jslhr.asha.org/cgi/content/full/53/6/1458>



AMERICAN
SPEECH-LANGUAGE-
HEARING
ASSOCIATION