

## Research Article

## Exploring Use of the Coordinate Response Measure in a Multitalker Babble Paradigm

Larry E. Humes,<sup>a</sup> Gary R. Kidd,<sup>a</sup> and Daniel Fogerty<sup>b</sup>

**Purpose:** Three experiments examined the use of competing coordinate response measure (CRM) sentences as a multitalker babble.

**Method:** In Experiment I, young adults with normal hearing listened to a CRM target sentence in the presence of 2, 4, or 6 competing CRM sentences with synchronous or asynchronous onsets. In Experiment II, the condition with 6 competing sentences was explored further. Three stimulus conditions (6 talkers saying same sentence, 1 talker producing 6 different sentences, and 6 talkers each saying a different sentence) were evaluated with different methods of presentation. Experiment III examined the performance of older adults with hearing impairment in a subset of conditions from Experiment II.

**Results:** In Experiment I, performance declined with increasing numbers of talkers and improved with asynchronous sentence onsets. Experiment II identified conditions under which an increase in the number of talkers led to better performance. In Experiment III, the relative effects of the number of talkers, messages, and onset asynchrony were the same for young and older listeners.

**Conclusions:** Multitalker babble composed of CRM sentences has masking properties similar to other types of multitalker babble. However, when the number of different talkers and messages are varied independently, performance is best with more talkers and fewer messages.

The work described in the three experiments below had two primary purposes. First, we wanted to learn more about what makes a competing babble disruptive. The number of talkers has long been established as a key factor, but it is not clear how the characteristics of a *multitalker babble*, such as the number of different voices or the number of different messages, contribute to the amount of interference. Second, we wanted to better understand how a particular set of stimuli, the *coordinate response measure* (CRM; Bolia, Nelson, Ericson, & Simpson, 2000), could be used as multitalker babble. When studying the perception of speech in competing speech, use of the CRM has become one of the most popular tools for English speech materials. All of the CRM sentences have the same rigid syntactic structure making use of the following form: “Ready <call sign> go to <color> <number> now.” There are eight call signs, four colors, and eight numbers, resulting in a total of 256 unique combinations. Each of

the 256 sentences is spoken by each of eight talkers, four men and four women. The syntactic structure of the sentences comprising the CRM enable the use of a lexical item, the *call sign*, as the marker for the target sentence to which the listener should attend. Although not a requirement, the convention in most studies making use of the CRM has been to make sentences with the call sign “Baron” the target sentence and all others the competition. The listener’s task in this case is to select the color–number combination produced in the sentence with “Baron” as the call sign while ignoring competing color–number combinations spoken in sentences with other call signs. The use of a call-sign marker for the target sentence to which the listener should attend enables the examination of a myriad of target-competition factors, such as the use of competition comprising: (a) the same talker as the target sentence; (b) different talkers of the same sex as the target talker; (c) different talkers of the opposite sex of the target talker; and (d) various combinations of these competing speech conditions. Each of these factors can also be examined further by varying the number of competing talkers and the target-to-masker ratio in dB. Brungart and colleagues (e.g., Brungart, 2001; Brungart, Simpson, Ericson, & Scott, 2001; Iyer, Brungart, & Simpson, 2010), in a series of systematic studies, have explored many of these parameters. One general finding to emerge across all of these studies is that each of these factors can affect

<sup>a</sup>Department of Speech and Hearing Sciences, Indiana University Bloomington

<sup>b</sup>Department of Communication Sciences and Disorders, University of South Carolina, Columbia

Correspondence to Larry E. Humes: humes@indiana.edu

Editor: Nancy Tye-Murray

Associate Editor: Mitchell Sommers

Received February 3, 2016

Revision received August 18, 2016

Accepted September 1, 2016

[https://doi.org/10.1044/2016\\_JSLHR-H-16-0042](https://doi.org/10.1044/2016_JSLHR-H-16-0042)

**Disclosure:** The authors have declared that no competing interests existed at the time of publication.

the perception of a target sentence in a background of similar competing target sentences.

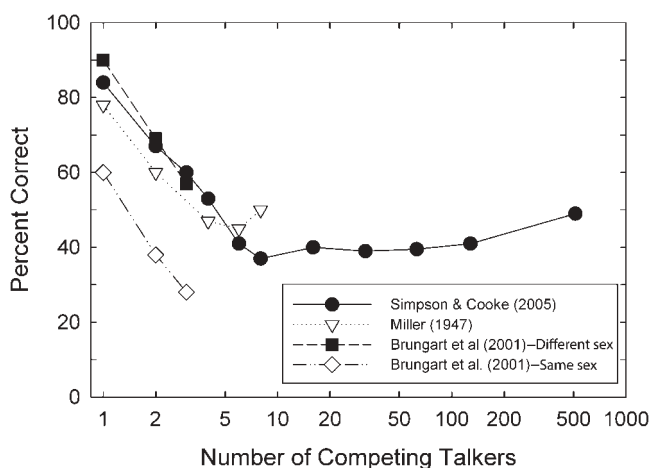
For the most part, the bulk of the data obtained with the CRM have been obtained with one, two, or three competing talkers. Yet, multitalker babble, a commonly used background signal when studying speech perception, typically comprises a blend of six to 12 talkers. A number of studies have investigated speech perception as a function of the number of competing talkers with a variety of speech materials (e.g., Bronkhorst & Plomp, 1992; Carhart, Tillman, & Greetis, 1969; Miller, 1947; Pollack & Pickett, 1958; Rosen, Souza, Ekelund, & Majeed, 2013; Simpson & Cooke, 2005). Figure 1 summarizes the data for the earliest (Miller, 1947) and one of the more recent (Simpson & Cooke, 2005) systematic studies of the number of competing talkers on speech perception. For comparison, the mean data from Brungart et al. (2001) are shown for competition comprising talkers of the same sex or different sex relative to the target talker for the CRM. The data for Brungart et al. (2001) and Miller (1947) are for a target-to-masker ratio of 0 dB, whereas the target-to-masker ratio used by Simpson and Cooke (2005) was -6 dB. Further, for each condition shown in Figure 1, Miller (1947) used an equal number of men and women as competition for a male target talker, and Simpson and Cooke (2005) made use of all-male competition and target stimuli. Last, across these three studies, the type of target material varied from sentences to words to vowel-consonant-vowel syllables, and the nature of the competition varied as well. Despite these differences across studies, the slope of the function from one to six or eight talkers is remarkably similar. The largest decline in performance is always from one to two competing talkers,

referred to as the *multitalker penalty* (Durlach, 2006; Iyer et al., 2010), with smaller decrements for each successive additional talker, reaching an asymptote at six or eight talkers. The data for babble competition from Rosen et al. (2013) are also consistent with these trends across number of talkers, although their data show little change in performance for more than two talkers. Although these findings are in general agreement, they also suggest that differences in speech materials can influence the point at which the effect of number of talkers reaches a maximum. Experiment I of this study seeks to confirm that the CRM follows the pattern established from other studies of multitalker competition by examining two-, four-, and six-talker combinations of CRM competing speech.

The general shape of the functions in Figure 1 conventionally is described as the result of a combination of so-called *energetic masking* and *informational masking*. As the number of talkers increases, the amount of spectral and temporal overlap with the target speech increases (thus increasing the potential for energetic masking) and the intelligibility of the competing speech decreases (decreasing the potential for a type of semantic interference or distraction generally associated with informational masking). However, these two types of masking are not as easily distinguished as is often implied. Although the definitions have been discussed at length (see especially Durlach et al., 2003, with elaboration provided more recently by Durlach, 2006), they remain problematic. Most recently, Stone and colleagues (Stone, Anton, & Moore, 2012; Stone, Füllgrabe, Mackinnon, & Moore, 2011; Stone, Füllgrabe, & Moore, 2012; Stone & Moore, 2014) have argued that even the masking of speech by “notionally steady-state noise” is not energetic masking in the sense of the masker’s internal effects swamping or suppressing that of the target signal. Instead, they suggest that various forms of modulation masking or modulation interference underlie even this assumed baseline energetic masking condition.

Rather than using the term *masking* to refer to all types of interference with speech understanding due to the presence of other co-occurring sounds, we prefer to use a term that does not seem to imply covering up or suppressing mechanisms. Thus, the detrimental effects of competing sounds on the understanding of the target speech will be referred to here as *interference*. Rather than attempting to identify the relative influence of energetic and informational mechanisms, we examine the amount of interference associated with different types of sounds that have been designed to include or exclude specific properties that have the potential to interfere with the understanding of co-occurring speech sounds. These types of sounds are as follows: (a) *natural speech*, consisting of meaningful spoken sentences with all of the natural spectral and temporal complexity; (b) *time-reversed speech*, consisting of spoken sentences with spectral and temporal complexity comparable to that of meaningful speech but without any semantic information; and (3) *speech-modulated noise*, with a long-term spectrum and wide-band amplitude envelope matched to the speech stimuli (plus the short-term random fluctuation of random noise), but without

**Figure 1.** Mean percent-correct performance from several studies for speech-identification or speech-recognition tasks for groups of young adults with normal hearing as a function of the number of competing talkers constituting the background. Data for the coordinate response measure (CRM) from Brungart et al. (2001) for one, two, or three competing talkers are shown when the sex of the target talker is the same as or different from the competing talker(s).



the semantic information or fine structure of speech. The examination of differences across these three types of interference are examined in Experiment II for various combinations of six talkers or six messages from the CRM.

As noted, there appear to be no published data for the CRM when using more than three competing talkers despite its widespread use for conditions involving one or two competing talkers. The addition of equal numbers of male and female competing talkers would enable up to six competing talkers for a given target talker with the CRM. In addition to those advantages of the CRM noted above for the study of speech-on-speech masking, the CRM has several other desirable features for research and potential clinical application. For example, although the CRM makes use of meaningful sentences, the rigid syntactical structure and limited vocabulary size reduce the predictability due to context and minimize learning effects common to many other sentence-based speech materials. Because the color-number responses of the listener are in the form of closed-set identification, response collection and scoring is easily automated, perhaps even allowing use of adaptive estimates of speech-perception performance (Eddins & Liu, 2012).

Another advantage associated with the use of the CRM as a tool to examine multitalker speech competition has to do with its rigid syntactic structure. This structure results in a fairly uniform wide-band envelope for all sentences and talkers in the CRM corpus. In typical multitalker babble, the competing speech is added with arbitrary or random temporal overlaps among the messages comprising the babble. Thus, the competing babble is inherently asynchronous with the target stimulus. The CRM enables one to explore the influence of the relative asynchrony of the target and competing speech stimuli. Although multitalker babble is an inherently asynchronous mix of competing speech signals, most often, the CRM has typically been administered with no onset asynchrony among any of the target and competing sentences, although this has not always been the case (Lee & Humes, 2012). The relative influence of asynchrony was explored in all three experiments in this project.

In the third experiment, older adults with impaired hearing were tested on a subset of CRM conditions from the second experiment that used natural time-forward speech competitors. This will assist in establishing the generality of the findings from Experiment II and also evaluate the potential clinical utility of a CRM-based multitalker babble. Older adults frequently complain of difficulty understanding speech in a background of competing speech. The CRM has been used previously to study this problem for one or two competing talkers (Humes & Coughlin, 2009; Humes, Kidd, & Lentz, 2013; Humes, Lee, & Coughlin, 2006; Lee & Humes, 2012), but not with six competing talkers or with variation in the number of different competing messages. To minimize the likely influence of inaudibility of the high-frequency portions of the speech stimuli, all speech stimuli were spectrally shaped to increase their amplitude in the regions of hearing loss (Humes, 2007; Humes et al., 2013).

## Experiment I: Two-, Four-, or Six-Talker CRM Competition

### Participants

Two groups of 10 young adults with normal hearing served as participants in Experiment I; one group for each of two different asynchrony values explored. Participants ranged in age from 19 to 23 years ( $M = 21.1$  years), and 16 of the 20 were women. These participants had hearing thresholds  $\leq 25$  dB HL (American National Standards Institute [ANSI], 2004) from 250 to 8000 Hz in both ears, had no evidence of middle-ear pathology, and had English as their native language. The test ear was always the right ear for these young participants. Participants were recruited primarily by flyers and university online postings and were paid for their participation.

### Stimuli and Apparatus

The CRM corpus (Bolia et al., 2000) was used for the target and competing speech in this study. As noted, the CRM measures the ability to identify two key words (color and number words) in a spoken target sentence (always identified here by the call sign “Baron”). The target sentence was always spoken by the same male talker (Talker 1; Bolia et al., 2000) and presented in a background of similar competing sentences from the same CRM corpus. For this experiment, the competing stimuli were either: (a) two CRM sentences, one spoken by a different man and the other by a woman; (b) four CRM sentences, two spoken by men other than the target talker and two spoken by women; or (c) six CRM sentences, three spoken by men other than the target talker and three spoken by female talkers.

For a given number of competing talkers, the competing sentences were either presented synchronously or asynchronously for a block of trials. For synchronous presentations, the onsets of all target and competing sentences were simultaneous. Despite the efforts to approximate equal durations across the entire stimulus set by Bolia et al. (2000), the offsets had an average difference of approximately 120 ms in the synchronous condition (less than half of the duration of the final word, “now”). For the asynchronous condition, two different approaches and values were explored, one for each group of 10 participants. For one group, a 50-ms onset asynchrony was applied so that 50 ms separated the onsets of all sentences. A second group had a 150-ms asynchrony between target and competing sentence onsets, and a 50-ms asynchrony between the other sentences. For the asynchronous conditions, half of the competing messages preceded the onset of the target sentence and half followed the onset of the target sentence. So, for example, for the six-talker competition, there were three competing male (M) voices and three competing female (F) voices producing CRM sentences with call signs other than “Baron” and color-number pairs differing from that produced in the target (T) sentence. For the asynchronous condition, the sequence of sentences on a given trial could have been M-F-F-T-M-F-M, with 50-ms onset asynchronies

between each competing sentence and either 50- or 150-ms separating the target onset from the onsets of adjacent competing sentences. This arrangement of target and competing talkers is illustrated in Figure 2 for the six-talkers competition in the asynchronous conditions. Although onset asynchrony also resulted in asynchronous endings, none of the background sentences ended until after the number in the target sentence had been presented.

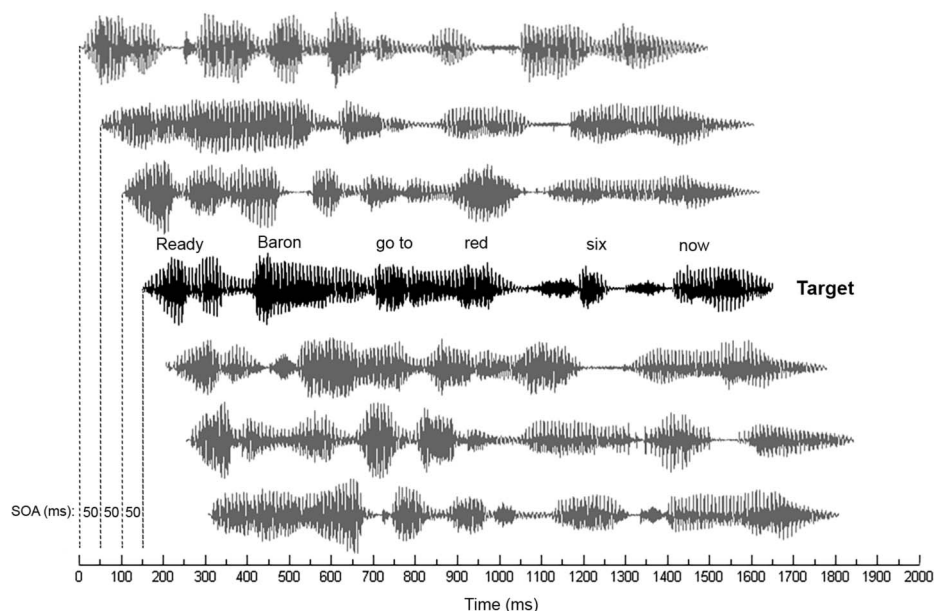
For the 150-ms condition, the competing CRM sentences immediately preceding and following the target sentence used a 150-ms onset asynchrony. The onset asynchrony among preceding or following competing sentences, however, was the same 50-ms value used with the other group. Onset asynchronies between the competing sentences were limited to 50 ms in this case to ensure overlap of all competing sentences with the color and number words in the target sentence with larger numbers of competing sentences. For just two sentences, one target and one competing, it was previously established that a 50-ms onset asynchrony improved performance by about 10 percentage points relative to the synchronous condition, and an increase to 150 ms further improved performance by another 10–20 percentage points (Lee & Humes, 2012). The 150-ms asynchrony value examined in the present study provides a test of a larger target–competitor asynchrony (150 ms) in the context of multiple competing sentences with a smaller asynchrony value (50 ms) among them.

All testing was done in a sound-treated booth that met or exceeded ANSI guidelines for permissible ambient noise for earphone testing (ANSI, 1999). Stimuli were presented monaurally, using an Etymotic Research ER-3A

insert earphone. A disconnected earphone was inserted in the nontest ear to block extraneous sounds. Stimuli were presented by computer using Tucker Davis Technologies System-III hardware (RP2 16-bit D/A converter, 48828-Hz sampling rate, HB6 headphone buffer). Each listener was seated in front of a touchscreen monitor, with a keyboard and mouse available.

The stimulus levels were 85 dB SPL for the target sentence and 78 dB SPL for *each* of the competing sentences for a target-to-masker ratio of +7 dB. (This relatively high presentation level was comparable to the levels used with the older participants with hearing impairment in Experiment III. Presentation levels for those participants were adjusted, on the basis of individual audiograms, to ensure audibility of the stimuli, as described below. Because this process results in relatively high presentation levels, similar levels were utilized in the present experiment to avoid between-groups differences due to the poorer intelligibility associated with presentation levels above 80 dB SPL [see Dubno et al., 2005a, 2005b; Studebaker et al., 1999]). The phrase *target-to-masker ratio* is used, as suggested by Brungart et al. (2001), to represent the level of the target relative to each competing sentence. As noted by Brungart et al. (2001), the overall *signal-to-noise ratio* (SNR) will then vary with the number of competing talkers. For this experiment, the SNRs are approximately +4, +1, and –1 dB for two, four, and six competing talkers, respectively. All levels were established via presentation of a noise, shaped to match the long-term spectrum of the CRM materials, through the ER-3A insert phone coupled to an HA-2 2-cm<sup>3</sup> coupler (ANSI, 2004; Lee & Humes, 2012).

**Figure 2.** Schematic illustration of the target and competing CRM stimuli in the six-talkers competition for Experiment I. The target sentence is shown in black, and each of the competing sentences is shown in gray. In this illustration, there is a 50-ms stimulus onset asynchrony (SOA) between all sentences. For the 150-ms asynchrony, only the competing sentences immediately preceding and following the target sentence had this SOA.





## Procedure

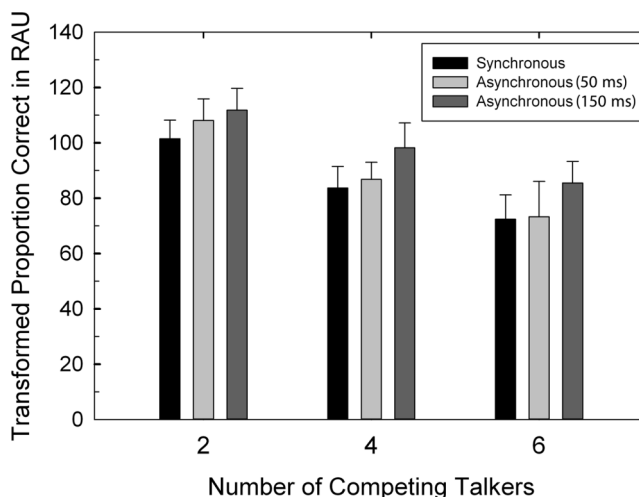
On each trial, target and competing sentences were presented simultaneously, and the task was to listen to the voice that said “Baron” as the call sign and report the color and number subsequently spoken in that voice. Each trial began with the word “listen” presented visually on the display, followed 500 ms later by presentation of the sentences. After each presentation, participants responded by touching (or clicking with a mouse) virtual buttons on a touchscreen display to indicate the color and number spoken by the talker who said “Baron.” All four colors and all eight numbers were included on the response display, which remained in view throughout each block of trials. The next trial was initiated by either clicking on (with the mouse) or touching a box on the monitor labeled “OK.” All trial blocks consisted of 32 trials, except for the practice trial block. Practice consisted of five trials of a no-distractor condition followed by five trials of each of the six conditions (two, four, or six background talkers, each synchronously and asynchronously combined), grouped in consecutive sets of five trials in the same condition and increasing the number of background talkers over trials. For the main experiment, the order of conditions was counterbalanced, with half of the participants tested with the following sequence of conditions in forward or reverse order. Each condition consisted of four trial blocks of the target sentence spoken by a male talker in the presence of: (a) two synchronous competing sentences (one man, one woman); (b) two asynchronous competing sentences (one man, one woman); (c) four synchronous competing sentences (two men, two women); (d) four asynchronous competing sentences (two men, two women); (e) six synchronous competing sentences (three men, three women); and (f) six asynchronous competing sentences (three men, three women).

## Results and Discussion

Figure 3 shows the means and standard deviations for the CRM in *rationalized arcsine units* (RAUs; Studebaker, 1985). RAUs, as the name implies, make use of an arcsine transformation to stabilize the error variance across the range of scores and then convert the resulting values to a range that matches the original percent-correct scores between about 20% and 80% correct. The mean scores for the synchronous condition for the 50-ms asynchrony group and the 150-ms asynchrony group differed by less than 2 RAUs and were pooled as a result. Because no asynchrony was involved for the synchronized comparison condition, this was essentially a replication of one of the conditions across the two groups. As one can see in Figure 3, as the number of competing talkers increased, performance declined. Further, the asynchronous conditions tended to yield higher scores, especially for the 150-ms asynchrony.

Two separate general linear model (GLM) analyses were performed on these data, one for each of the two groups. Repeated-measures factors were synchrony (synchronous or asynchronous) and number of competing

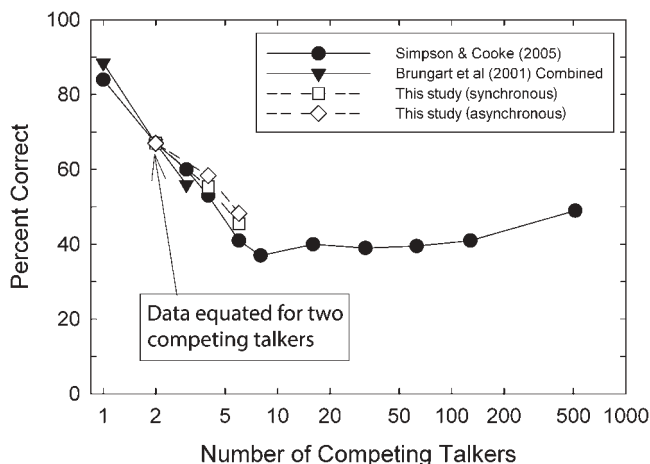
**Figure 3.** Means and standard deviations (+1) for rationalized arcsine units (RAU)-transformed percent-correct scores for two, four, and six competing talkers. The different colored bars represent different synchronies among the sentences constituting the stimulus ensemble. The data for the synchronous condition represent the pooled results for both groups ( $N = 20$ ) whereas the asynchronous conditions are for each group ( $n = 10$ ) separately.



talkers (2, 4, or 6). The number of competing talkers had a significant ( $p < .05$ ) influence on performance for both the 50-ms asynchrony group,  $F(1, 9) = 166.16$ ; and the 150-ms asynchrony group,  $F(1, 9) = 100.96$ ; with follow-up Bonferroni-adjusted  $t$  tests indicating that each value for number of talkers differed significantly ( $p < .05$ ) from all other values. For synchrony, the main effect was significant ( $p < .05$ ) for only the 150-ms asynchrony group,  $F(1, 9) = 123.59$ , although it approached significance ( $p = .057$ ) for the other group,  $F(1, 9) = 4.76$ . The interaction between the factors of synchrony and number of competing talkers was significant for the 50-ms asynchrony,  $F(1, 9) = 7.01$ ; but not for the 150-ms asynchrony,  $F(1, 9) = 0.89$ . The significant interaction for 50-ms is due to a significant effect of synchrony for two competing talkers, but not for four or six competing talkers—that is, the addition of a 50-ms asynchrony improved performance for two competing talkers, but not for four or six competing talkers, whereas the 150-ms asynchrony always resulted in better performance compared to synchronous presentations, regardless of the number of talkers.

Figure 4 compares the effects of number of competing talkers for the CRM measured in this study to that of Brungart et al. (2001), also for the CRM, but at a more difficult target-to-masker ratio (and pooling the same-sex and different-sex competition conditions). Also shown for comparison are the more extensive data of Simpson and Cooke (2005) for vowel–consonant–vowel syllables in babble (with SNR held constant, rather than the target-to-masker ratio). To facilitate comparison across the various test conditions and materials, the percent-correct scores were normalized to the same value for two competing talkers,

**Figure 4.** Comparison of the mean data from this study (open symbols) to a subset of the results from the literature as a function of number of competing talkers. The data from Simpson and Cooke (2005) are for consonant-identification targets in multitalker babble, whereas all the other data are for CRM target sentences in a background of CRM competing sentences. To reduce overall differences in performance across tasks and materials, the scores were all equated for two competing talkers.



the fewest number of talkers common to all studies compared. The data from this experiment are clearly in good agreement with both prior datasets regarding the relative change in performance with increasing number of competing talkers.

As noted above, the synchrony factor yielded somewhat mixed results, with a significant effect only observed for the 150-ms asynchrony group. For both groups, however, at least descriptively, the asynchronous condition leads to higher scores than the synchronous condition. The asynchronous conditions are expected to lead to better segregation of each of the speech streams (see, for example, Bregman, 1990; Carlyon, 2004), which, in turn, leads to better performance. The CRM task requires listeners to follow the speech stream containing “Baron” so that they can identify the correct color–number response. If it is more challenging to segregate the target from competing speech streams, then performance should be worse. In conditions with fewer competing talkers, stimulus onset asynchronies in the range used here have been shown to be effective cues for the segregation of target and competing speech streams (Lee & Humes, 2012). The 150-ms offset, even though it is present only between the target sentence and the immediately preceding and following competitors, appears to be more effective in helping to segregate the target, especially with larger numbers of talkers. It is important to note that such effects of onset asynchrony have been observed even when there are other cues, such as sex-related differences in voice fundamental frequency, present in the ensemble of target and competing talkers (Lee & Humes, 2012). Recall that half of the talkers constituting the competition in this study were of the same sex as the target talker, and half were of the opposite sex.

Another factor that may underlie the superior performance observed in the asynchronous conditions is the

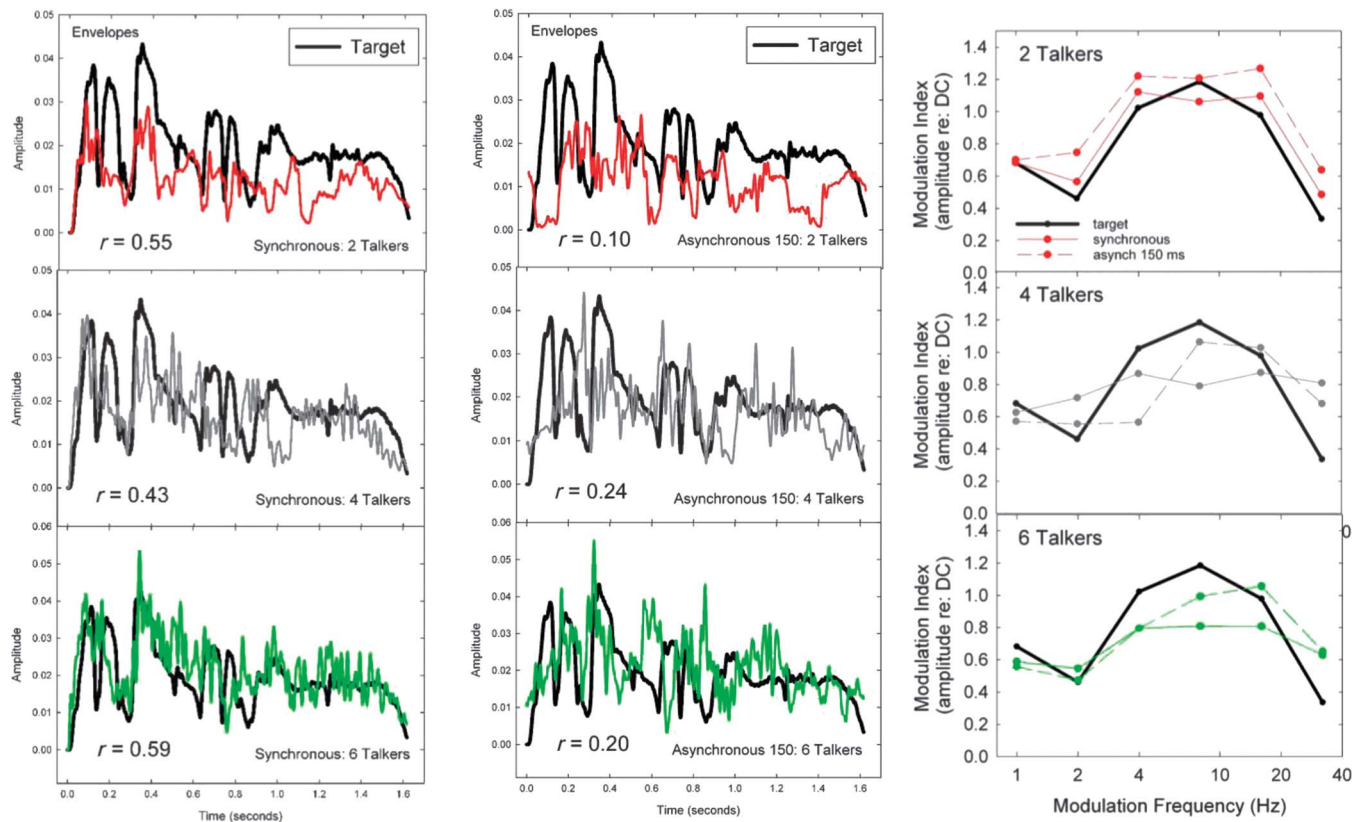
correlation of the envelopes. This is illustrated in Figure 5. The left panels illustrate the envelopes extracted from the synchronous conditions for sample trials in which the target CRM sentence (black line) was presented in backgrounds of two (top, red), four (middle, gray), and six (bottom, green) competing talkers, in each case speaking a unique CRM sentence. Envelopes were extracted by half-wave rectifying the signal followed by low-pass filtering via a sixth-order Butterworth filter at 50 Hz. In each panel, the correlation ( $r$ ) between these wide-band envelopes is also provided. For the synchronous conditions (left panels), the correlations between the target and competing envelopes are all moderate, ranging from .43 to .59. In contrast, the middle panels of Figure 5 show comparable analyses of the envelopes for the 150-ms asynchrony condition. Here the envelope correlations are all low, ranging from .10 to .24. Similarly low envelope correlations were observed for the 50-ms asynchrony condition ( $r$ s = .09–.20). Last, the envelopes were low-pass filtered at 1000 Hz, fast-Fourier transformed, with amplitudes normalized to the DC component, then summed into different modulation frequency bins with center frequencies of 1, 2, 4, 8, 16, and 32 Hz, similar to the method used by Gallun and Souza (2008). The resulting octave-band modulation spectra appear in the right panels. The modulation spectra reveal a relative shift toward higher modulation frequencies (16 Hz) for four- and six-talker competition, especially for the asynchronous condition. Only 150-ms asynchrony is shown here, but the modulation spectra for the 50-ms asynchrony were very similar to those shown in Figure 5 for 150-ms. This similarity is not surprising, given that the 150-ms condition still maintained 50-ms delays between the competitor stimuli for these same competing stimuli.

From these acoustic analyses, it appears that the asynchronous condition may be less difficult than the synchronous condition, at least in part because of the resulting dissimilarities of the envelopes (lower correlations) and of the modulation spectra between the target and the competition. The envelopes or modulations of the asynchronous competing stimuli, as a result, interfere less with the target stimulus leading to higher identification performance. The impact of the shift of modulation spectra would also be consistent with less modulation masking in the easier asynchronous conditions, as hypothesized by Stone and colleagues (Stone, Anton, & Moore, 2012; Stone et al., 2011; Stone, Füllgrabe, & Moore, 2012; Stone & Moore, 2014). The influence of target and masker onset asynchrony is further explored for several variations of the condition with six competing CRM stimuli in the next experiment.

## Experiment II: Independent Variation in Talkers and Messages for Six Competing CRM Stimuli

Having confirmed that the CRM shows a dependence on the number of talkers constituting the competing speech signal that is similar to that observed in several earlier studies

**Figure 5.** The left panels illustrate the envelopes extracted from the synchronous conditions for sample trials in which the target CRM sentence (black line) was presented in backgrounds of two (top, red), four (middle, gray), and six (bottom, green) competing talkers, in each case speaking a unique CRM sentence. In each panel, the correlation ( $r$ ) between these wide-band envelopes is also provided. Synchronous conditions are shown in the left panels, and asynchronous conditions are shown in the middle panels. The octave-band modulation spectra derived from the envelopes appear in the right panels. DC = direct current, or 0-Hz.



using other speech materials, we next examined the six-stimulus combinations in more detail. Because increases in the number of talkers have generally involved increasing the number of different messages as well as the number of different voices, the relative importance of these two variables in the number-of-talkers effect is unclear. To evaluate this, the numbers of different voices and messages in the babble were varied independently in Experiment II. Further, each of these three conditions was examined for synchronous and asynchronous (50-ms) combinations of competitor sentences. In addition, to establish the relative influence of the meaningfulness of the competition on target-identification performance, additional types of competition were used that shared some of the acoustic characteristics of the typical CRM-based babble but lacked linguistic meaning. The impact of the linguistic meaning of the competing stimulus on target identification performance has been studied many times by either time-reversing the competing stimulus or extracting the envelope from the competing speech and modulating a speech-shaped noise (e.g., Carhart et al., 1969; Dirks & Bower, 1969; Duquesnoy, 1983; Festen & Plomp, 1990; Humes & Coughlin, 2009; Rhebergen, Versfeld, & Dreschler, 2005; Van Engen & Bradlow, 2007) with variation

of the languages used for the target and competing stimuli being a more recent approach (e.g., Calandruccio, Brouwer, Van Engen, Dhar, & Bradlow, 2013; Calandruccio, Dhar, & Bradlow, 2010; Calandruccio & Zhou, 2014; Freyman, Balakrishnan, & Helfer, 2001; Van Engen & Bradlow, 2007). However, the competing speech in these earlier studies typically comprised one to three talkers, not six competing speech stimuli as in this study. To examine the influence of the meaningfulness of the competing speech for the six-stimulus CRM competition used here, both time-reversed CRM and envelope-modulated speech-spectrum noise were examined. Additional details follow.

### Participants

Twenty-nine young adults with normal hearing served as participants in Experiment II. The participants were assigned to one of three groups, one for each of three different types of background competition: (a) CRM time-forward ( $n = 10$ ); (b) CRM time-reversed ( $n = 9$ ); and (c) CRM envelope-modulated ( $n = 10$ ) noise. Participants ranged in age from 18 to 27 years ( $M = 20.9$  years), and 18 of the 29 were women. These participants had hearing thresholds



$\leq 25$  dB HL (ANSI, 2004) from 250 to 8000 Hz in both ears, had no evidence of middle-ear pathology, and had English as their native language. The test ear was always the right ear for these participants. Participants were recruited primarily by flyers and university online postings and were paid for their participation.

### ***Stimuli and Apparatus***

The background competition composed of six-sentence time-forward CRM sentences (as used in Experiment I) was the focus of this experiment. The time-reversed CRM sentences and the modulated noise were used as comparison conditions having envelope information similar to the time-forward CRM competition but with reduced linguistic or periodicity information (e.g., Brungart et al., 2001; Iyer et al., 2010). The target talker (T1; Bolia et al., 2000) and call-sign (“Baron”) was the same throughout this experiment as in the prior experiment.

For the time-forward CRM competition, the ensemble consisted of six CRM sentences with variation in the number of talkers and the number of messages. We examined three versions of six-stimulus competition: (a) one talker speaking six different CRM messages; (b) six talkers each speaking the same message and (c) six talkers each producing a unique sentence. Each of these three stimulus combinations was examined for synchronous and asynchronous (50 ms) combinations of stimuli. The single-talker conditions utilized six instances of a single male talker, and the six-talker conditions used three male (including the competing talker used for the single-talker conditions) and three female talkers. The target talker was never included among the background talkers. The call signs in the competition always differed from the target, and the target’s color–number combination was not used in any competitor, nor was any color–number combination used more than once in the competition. However, color or number duplication was allowed to occur (through random selection) within the set of target and competitor sentences on each trial.

The target-to-masker ratio (target relative to each competing sentence, as described above) was fixed at +7 dB for the time-forward CRM competition. For most of the six listening conditions, this would be expected to yield an overall SNR of about  $-0.8$  dB ( $10 \times \log_{10} [6] = 7.8$  dB). The measured SNR was slightly better, averaging  $0 \pm 0.5$  dB across five of six listening conditions. It was closer to  $-1$  dB for the one talker–six messages condition.

For the time-reversed CRM competition, each of the competing sentences was played backwards. Otherwise, the listening conditions were identical to those used for the time-forward CRM competition. For envelope-modulated noise, a fifth-order Butterworth low-pass filter with a cutoff frequency of 20 Hz was used to extract the envelope of the six-sentence competition, which was then used to modulate the amplitude of a noise having a long-term average spectrum identical to that of the CRM corpus. The root-mean-square amplitude of the modulated noise was equated to that of the time-forward and time-reversed CRM competition.

Pilot data were used to adjust the target-to-masker ratios for the time-reversed CRM and modulated-noise competition so as to place the overall performance in the same range. It is well known that the reduction or absence of linguistic content in competing sounds yields considerably better performance than with the conventional time-forward CRM sentences (Brungart, 2001; Brungart et al., 2001; Iyer et al., 2010). On the basis of these prior data and our own pilot data, target-to-masker ratios of +5 and +1 dB were selected for the time-reversed CRM and modulated-noise competition, respectively. This corresponds to overall SNRs of about  $-2$  and  $-6$  dB, respectively. Note that for each competition type, these SNRs (and underlying target-to-masker ratios) were fixed for all six conditions using that competition type—that is, the different SNRs used for each competition type were just used to place the overall performance for each competition type in the same range to facilitate comparisons of relative performance across the six stimulus conditions with different types of competition.

For onset asynchrony, a 50-ms onset asynchrony throughout the seven CRM-sentence ensemble was used. Unlike Experiment I, however, for which the target sentence was always preceded and followed by three competing talkers of the six-talker competition, for this experiment, the target sentence was randomly selected to be either the third, fourth, or fifth sentence of the seven-sentence ensemble in the asynchronous condition. The position of the target sentence varied randomly within a block of trials with an equal proportion of each of the three sequence positions in a trial block.

The equipment used in this experiment was identical to that used in Experiment I. The same test environment and response collection methods described in Experiment I were also used here.

### ***Procedure***

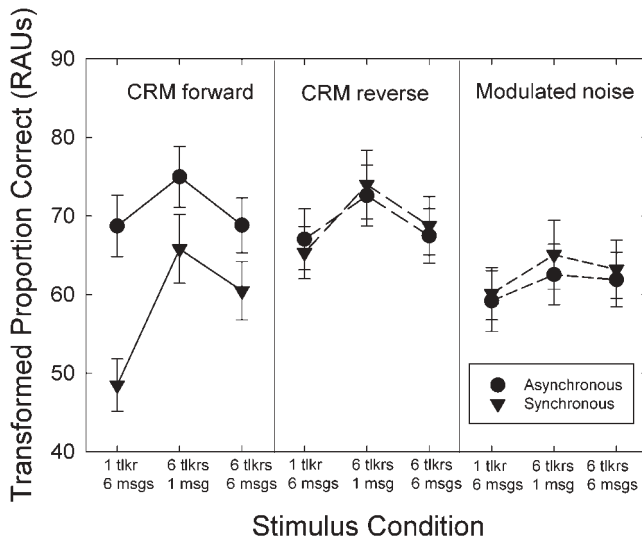
The same procedures used in Experiment I were used here. As noted, each participant group received one of the three types of competition. All trial blocks consisted of 32 trials, except for the practice block, which consisted of 30 trials; 10 with no competition and 20 with six talkers and six sentences, 10 synchronous and 10 asynchronous. After practice, participants were presented with four consecutive trial blocks of each condition, using a different random order of conditions for each participant. The same set of random orders was used for the group of participants in each of the three masker types.

### ***Results and Discussion***

The mean performance for the six listening conditions for the typical time-forward CRM stimuli, transformed from proportion correct to RAUs, are shown in the left panel of Figure 6. The error bars represent  $\pm 1$  SEM. The circles show the asynchronous CRM competing stimuli, and the inverted triangles show the data for the synchronous stimuli. Two trends are apparent in these data for these typical



**Figure 6.** The mean performance for the six listening conditions, transformed from proportion correct to rationalized arcsine units (RAUs). The error bars represent  $\pm 1$  SEM. The left, middle, and right panels represent the typical time-forward coordinate response measure (CRM) competition, the time-reversed CRM competition, and the envelope-modulated noise competition types, respectively. Circles represent the synchronous competition; inverted triangles represent the asynchronous competition in each panel. tlkr = talker; msg = message.



CRM time-forward competing stimuli: (a) the relative synchrony of the competing messages with the target onset matters, with performance being 10–20 RAUs higher for the asynchronous competing CRM messages; and (b) there is a trend for the six talkers–one message competition (middle data points) to be the easiest among the competing stimuli. The middle and right panels of Figure 6 show means for the two control comparison conditions, time-reversed CRM competition (middle) and envelope-modulated noise (right). As noted above, the target-to-masker ratios for each type of competition were manipulated to put overall performance in the same general range. Differences in overall performance between competition comprising meaningful time-forward speech versus either meaningless time-reversed speech or envelope-modulated noise had been well established previously and were not of interest here (e.g., Bronkhorst & Plomp, 1992; Brungart et al., 2001; Carhart et al., 1969; Dirks & Bower, 1969; Festen & Plomp, 1990; Humes & Coughlin, 2009; Iyer et al., 2010; Simpson & Cooke, 2005). Rather, the relative differences across talker/message/synchrony conditions for each type of masker were of greater interest in this study.

The data illustrated in Figure 6 indicate that when the target-to-masker ratios were adjusted to put overall performance levels in the same general range, the relative effects of number of talkers, number of CRM messages, and asynchrony differed across the three types of competition. Most obvious from visual inspection of the data in Figure 6 is the negligible influence of stimulus synchrony on performance for the time-reversed CRM (middle) and

the modulated noise (right). There is a trend visually, however, across all three types of interference for the competition on the basis of the one talker–six messages configuration being the most difficult and the six talkers–one message combination being the easiest among the three configurations evaluated here.

The trends apparent from visual inspection of the data in Figure 6 were largely supported by a mixed-model GLM analysis examining the between-participants factor of type of competition (CRM, time-reversed CRM, or modulated noise) and the within-participant factors of asynchrony (0 or 50 ms) and competing stimulus configuration (one talker–six messages, six talkers–one message, six talkers–six messages). The resulting  $F$  values are presented in Table 1. First, the top three rows show the results for the main effects of type of competing stimulus, competing stimulus configuration, and asynchrony, respectively. Note that, due to our control of performance levels, the effect of competition type was not significant ( $p > .05$ ) but that competition configuration and asynchrony were significant ( $p < .05$ ). Consistent with the visual pattern of the data in Figure 6, the effect of asynchrony interacted with the type of competition, being much larger for typical time-forward CRM presentation than for the other two interferers. Asynchrony also interacted with the competition configuration, due mainly to the larger effect of asynchrony in the one talker–six messages condition than in the other two configurations, observed primarily in the typical time-forward CRM condition.

The pattern of results for the time-forward CRM competition in Figure 6 is generally consistent with the explanation of the effects of synchrony offered for the results from Experiment 1—that is, the onset asynchrony of the sentences facilitates segregation of the target and competing CRM sentences and reduces the interference between target and competition. The temporal onset asynchrony of the competing and target messages allowed for considerable improvement in performance, presumably because the listener could now segregate the target from the six competing speech streams. When multiple voices are introduced in the competition, as in the 2 six-talker competition configurations, the benefit of onset asynchrony diminishes considerably, although it is still apparent across all of the listening conditions.

**Table 1.**  $F$  values from mixed-model general linear model (GLM) analysis of the data from Experiment II

GLM term	df	$F$	$p$
Competition type (T)	2, 27	1.22	.30
Competition configuration (C)	2, 54	16.00	<.001*
Asynchrony (A)	1, 27	18.00	<.001*
T $\times$ C	4, 54	1.53	.21
T $\times$ A	2, 27	29.18	<.001*
C $\times$ A	2, 54	3.79	.03*
T $\times$ M $\times$ A	4, 54	1.50	.21

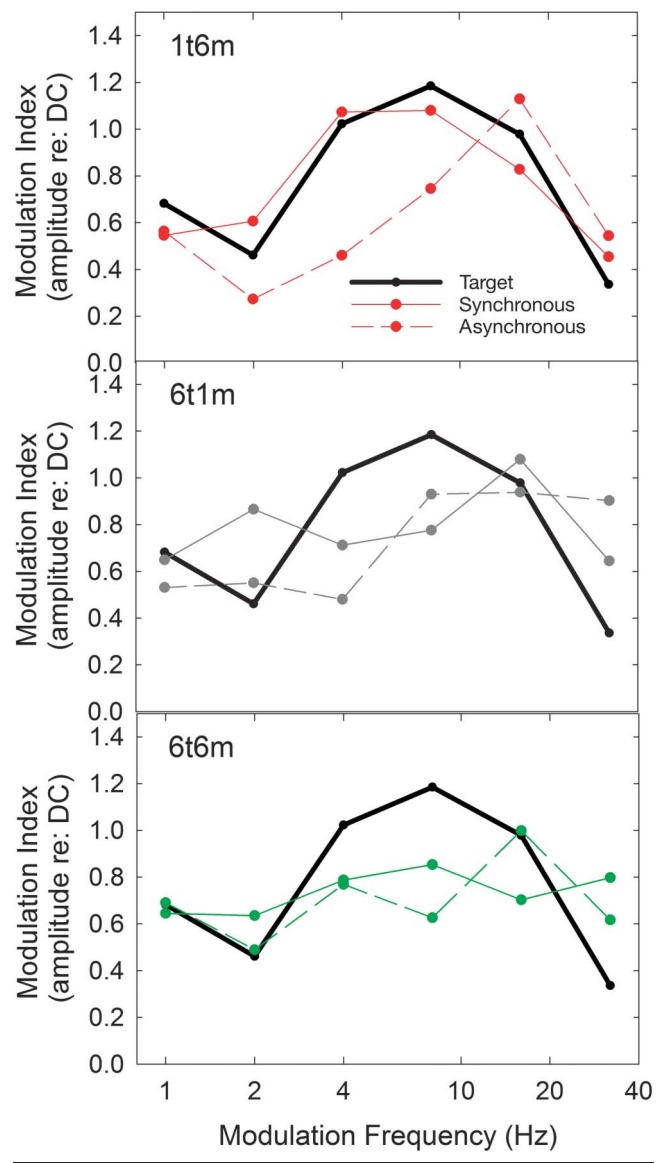
Note. The degrees of freedom (df) and significance levels ( $p$ ) are also in this table. Significant ( $p < .05$ )  $F$  values are marked with asterisks.

Analyses of the wide-band envelope and the modulation spectra for the time-forward CRM competitor, the competitor exhibiting the greatest effect of asynchrony, were also performed to determine whether insights could be gained with regard to factors contributing to this effect. Correlations between the envelopes of the target and competitor were weak for the one talker–six messages and six talkers–one message conditions, both synchronous and asynchronous, with  $.17 \leq r \leq .26$ , but moderate ( $r_s = .44$  and  $.50$ ) for the six talkers–six messages conditions. Unlike similar analyses for the conditions in Experiment I (see Figure 5), there were no systematic differences in correlation between each synchronous/asynchronous competitor pair.

With regard to modulation spectra, Figure 7 shows the modulation spectra for a representative target sentence and synchronous/asynchronous competitor pairs for conditions with multiple talkers or messages: one talker–six messages (top), six talkers–one message (middle), and six talkers–six messages (bottom). As in Experiment I, the modulation spectra for the easier asynchronous conditions show an upward shift of the spectrum relative to the target. As a result, less modulation masking or interference would be expected to occur for the asynchronous conditions, as was observed. Moreover, the difference between the synchronous and asynchronous conditions would be expected to be greatest for the one talker–six messages condition, as this condition shows the biggest shift in modulation spectra. This was also observed. Further, in terms of absolute identification performance, the synchronous one talker–six messages competitor shows the greatest overlap with the target's modulation spectrum, suggesting that this would be the most difficult of these six listening conditions. This is also the case. Thus, at least qualitatively, the pattern of results across these six time-forward CRM competing backgrounds appears to be consistent with the modulation masking or interference predictions of Stone and colleagues (Stone, Anton, & Moore, 2012; Stone et al, 2011; Stone, Füllgrabe, & Moore, 2012; Stone & Moore, 2014).

These reductions in modulation masking would also be expected to occur for the nonlinguistic competitors, particularly the noise competitor that was modulated by the wide-band temporal envelope derived from the time-forward competitors analyzed here. However, the large effect of asynchrony and its strong dependence on the babble configuration occurred only for the time-forward competition. Thus, asynchrony may facilitate selective attention to the target by desynchronizing it temporally from competing keywords in the meaningful competition. Without the segregation cue of onset asynchrony, an increase of the number of different talkers (i.e., distinct voices) in the background actually helps segregate the target talker, regardless of the number of different messages in the competition. Increasing the number of talkers may have a similar effect as asynchrony (i.e., shifting the modulation spectrum of the babble), as can be observed from Figure 7. In contrast to temporal onset asynchrony, the asynchrony introduced by using multiple talkers did result in a trend toward better performance for the nonlinguistic competitors. Overall, these

**Figure 7.** Octave-band modulation spectra for a representative target sentence and synchronous/asynchronous competitor pairs for the following competing stimulus configurations: one talker–six messages (top, 1t6m), six talkers–one message (middle, 6t1m), and six talkers–six messages (bottom, 6t6m). DC = direct current, or 0-Hz.



results suggest that the functional significance of the upward-shifted modulation spectrum for asynchronous conditions is that it serves as a segregation cue to reduce linguistic interference from a meaningful competitor.

### Experiment III: Extension to Older Adults With Hearing Loss

To test the generalization of the findings from young listeners with normal hearing in Experiment II to another group of listeners, older adults were recruited for

Experiment III and presented with the time-forward CRM stimuli from Experiment II. Many older adults have significant high-frequency hearing loss (Cruickshanks, Zhan, & Zhong, 2010), and they also represent about two-thirds of the hearing aid purchasers in the United States (Kochkin, 2009). As a result, older adults with hearing loss represent a commonly encountered group in research laboratories and in the clinic. In this experiment, we evaluated whether the relative effects of CRM competing-stimulus configuration and asynchrony observed in young adults with normal hearing also applied to older adults with impaired hearing. We were not interested in the performance of older adults on the CRM when they could not hear the stimuli entirely. Rather, we minimized the influence of inaudibility of both the target and competing CRM messages by amplifying or spectrally shaping the stimuli on the basis of the hearing loss of the individual participants. In addition, we were most interested in the effect of the most ecologically valid competition, time-forward CRM, on the performance of older adults with impaired hearing.

### Participants

Eleven older adults with hearing loss served as the participants for this experiment. Participants ranged in age from 59 to 77 years ( $M = 67.5$  years), and seven of the 11 were men. The mean air-conduction hearing thresholds (ER-3A transducer; re: ANSI, 2004) are provided in Table 2 for 250–8000 Hz in the right ear. The test ear was always the right ear for these participants. Participants also had no evidence of middle-ear pathology, no signs of dementia (Mini Mental State Examination  $> 25$ ; Folstein, Folstein, & McHugh, 1975), and English as their native language. Participants were recruited primarily by flyers and university online postings. Participants were paid for their participation.

### Stimuli and Apparatus

The stimuli and apparatus used were identical to that of Experiment II for the time-forward CRM competition, with one primary exception: All CRM materials, target and competition alike, were spectrally shaped to compensate for the inaudibility of higher frequencies in these participants. Although there are various ways in which compensation for inaudibility could be implemented (Humes, 2007),

**Table 2.** Means and standard deviations for the hearing thresholds (re: ANSI, 2004) of the 11 older adults with impaired hearing tested in Experiment III.

Test frequency (Hz)	<i>M</i> threshold in dB HL	<i>SD</i> in dB HL
250	17.3	10.3
500	18.2	7.2
1000	14.1	8.0
2000	28.2	13.5
3000	42.3	13.7
4000	52.7	12.3
6000	62.3	15.6
8000	65.5	16.0

this experiment made use of the same approach described recently by Lee and Humes (2012) and Humes et al. (2013). For the older listeners, the long-term spectrum of the full set of stimuli was measured, and a filter was applied to shape the spectrum according to each listener's audiogram (Humes et al., 2013). The shaping was applied with a 68 dB SPL overall speech level for the unshaped long-term spectrum as the starting point, and gain was applied as necessary at each one-third-octave band to produce speech presentation levels at least 13 dB above threshold from 125 to 4000 Hz.

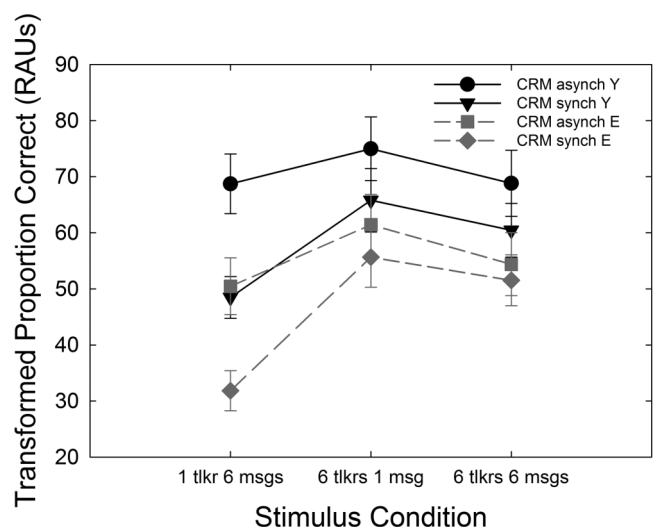
### Procedure

The procedures were identical to those used in Experiment II. Here, however, only the time-forward CRM masker was used with the older participants. As noted, this condition is most representative of everyday multitalker interference. In addition, as observed in Experiment II, various factors, such as competing-stimulus asynchrony, appear to uniquely influence this type of competition.

### Results and Discussion

Figure 8 shows the means and standard errors of the RAU-transformed CRM scores for the older adults in the present experiment (gray symbols and lines) as well as for the young adults from Experiment II (black symbols and lines). The data for the young participants with normal hearing are the same as shown previously for the time-forward condition in Figure 6. Two general trends are clearly apparent in these data. First, the results from the older adults

**Figure 8.** Means and standard errors ( $\pm 1$ ) of the rationalized arcsine unit (RAU)-transformed percent-correct scores for the six listening conditions from the 11 older adults with hearing impairment (E; gray symbols and lines) in Experiment III. Comparison data from young adults with normal hearing (Y; black symbols and lines) from Experiment II (Figure 6) are also plotted. All data are for the time-forward coordinate response measure (CRM) competition. *asynch* = asynchronized; *synch* = synchronized; *tlkr* = talker; *msg* = message.



with hearing impairment are roughly 15–20 RAUs worse than those of the young adults with normal hearing across the entire set of eight listening conditions. Second, the relative pattern of performance across the six conditions is quite similar between these two groups.

These general trends visible in Figure 8 were confirmed via a mixed-model GLM analysis of the data in this figure. Table 3 presents the *F* values from this analysis. First, the main effect of participant group was not significant but approached significance ( $p = .056$ ), and there were no significant interactions of the group variable with the other independent variables. Second, the pattern of relative effects across the six within-participant listening conditions is identical to that described previously (see Table 1) for the young adults with normal hearing only.

Assuming a slope of about 10 RAUs/dB for the linear portion of the psychometric function relating performance to target-to-masker ratio (Brungart et al., 2001; Eddins & Liu, 2012), the overall difference in performance of 15–20 RAUs observed in Figure 8 suggests that the older adults were operating at about a 2-dB deficit in target-to-masker ratio relative to the young adults, despite spectral shaping to restore audibility. Although the age difference was only marginally significant (due to large individual differences and considerable overlap in performance between the two age groups), the trend is consistent with earlier research. As has been noted frequently in the past for older adults with normal hearing or older adults with hearing loss listening to spectrally shaped stimuli, it is often the case that they function like young adults at a worse SNR, especially for speech in backgrounds of competing speech (Humes & Dubno, 2010; Lee & Humes, 2012). For spectrally shaped speech, the speech-perception performance of older listeners with hearing loss listening in backgrounds of competing speech is influenced by a number of auditory-perceptual and cognitive-linguistic factors (e.g., Akeroyd, 2008; Humes & Dubno, 2010; Humes et al., 2013). Age-related differences in these factors may underlie the group differences across the six listening conditions observed in this experiment, and large individual differences in these factors may account for much of the variability within each group.

**Table 3.** *F* values from mixed-model general linear model (GLM) analysis of the data from Experiment III.

GLM term	<i>df</i>	<i>F</i>	<i>p</i>
Age group (G)	1, 19	4.12	.056
Competition configuration (C)	2, 38	37.90	<.001*
Asynchrony (A)	1, 19	99.10	<.001*
G × C	2, 38	1.89	.16
G × A	1, 19	2.55	.12
C × A	2, 38	17.90	<.001*
G × M × A	2, 38	0.29	.74

*Note.* The degrees of freedom (*df*) and significance levels (*p*) are also in this table. Significant ( $p < .05$ ) *F* values are marked with asterisks.

The overall deficit of about 15–20 RAUs in the older adults with impaired hearing, relative to the young adults with normal hearing, across all listening conditions (see Figure 8), is consistent with prior findings from similar groups reported in Lee and Humes (2012) for the CRM. In Lee and Humes (2012), only one competing CRM talker was evaluated. The competing talker's voice was digitally manipulated to change the fundamental frequency by zero, three, or six semitones. Lee and Humes (2012) examined the effect of onset asynchrony for asynchrony values of 0, 50, 150, 300, and 600 ms, using a target-to-masker ratio of 0 dB. These onset asynchronies span a range that is relevant to the competition used here, with a 50-ms asynchrony between each sentence in the seven-sentence CRM ensemble and a cumulative asynchrony of 300 ms from the onset of the first competing sentence to the onset of the last competing sentence in the CRM ensemble. Lee and Humes (2012) found that older adults with impaired hearing, similar to the older adults in this study, performed about 20 RAUs worse than young adults with normal hearing under the same spectrally shaped stimulus conditions and across all values of onset asynchrony. This was true, moreover, despite demonstrating that both participant groups were equally able to identify the “Baron” call sign cueing the correct color–number pair later in the sentence. Further, in quiet listening conditions without a competing talker, both participant groups demonstrated color–number pair identification accuracy of at least 98%. Thus, the spectral shaping restored the audibility of the CRM color–number pairs to enable accurate identification by the older adults with impaired hearing. Despite this, the older adults in Lee and Humes (2012), as those in this experiment, performed about 20 RAUs worse than the young adults with normal hearing across all listening conditions. It is interesting that the magnitude of the age-related deficit for CRM color–number identification in the presence of competing CRM speech is roughly the same, whether there is one competing talker (Lee & Humes, 2012) or six (this experiment).

Although the older adults with impaired hearing tended to perform worse than the young adults with normal hearing across the range of listening conditions in this experiment, the relative differences in performance were quite similar. Thus, the relative benefits of onset asynchrony were the same for both groups, as were the effects of the number of competing talkers or messages.

Because the two participant groups in Experiment III differed in age and hearing loss, it is not possible to pin down whether the overall performance decrement in the older group may be attributable to age, peripheral pathology, or the combination. The results from Lee and Humes (2012) were mixed in that regard. When the results of older adults with impaired hearing were compared to older adults with normal hearing, there were significant differences between these two groups for one experiment but not for the other. Further research is required to better elucidate the mechanisms underlying the overall performance decrement observed for CRM identification performance in the presence of competing CRM sentences. Humes et al. (2013), on the



basis of the analysis of individual differences in aided speech-recognition and speech-identification performance in older adults, including the CRM, suggested that the mechanism(s) may involve higher level cognitive-linguistic processing.

## Summary and Conclusions

This series of experiments explored the use of the CRM speech materials in a multitalker babble paradigm. The first experiment demonstrated that performance on the CRM depended on the number of competing talkers in a manner similar to that observed previously for other speech materials. The first experiment also confirmed that, given the CRM's rigid syntactic structure and constrained vocabulary, the synchronization of the competing talkers comprising the babble can affect the measured speech-perception performance. The advantage for asynchronous presentation was limited with the 50-ms asynchrony (significant only with two competing talkers) but more robust for all numbers of competing talkers with a 150-ms asynchrony on either side of the target sentence.

Experiment II focused on various competing sentence manipulations exclusively for the competing-stimulus ensembles with six CRM stimuli. Synchrony of the stimulus ensemble again influenced performance so that asynchronous conditions tended to yield higher scores than synchronous conditions for the typical CRM presentation mode (time-forward target and competing CRM sentences). However, asynchrony had relatively little effect on performance when the interfering stimuli lacked meaning (time-reversed CRM and wide-band envelope-modulated noise). The effects of the number of different talkers and messages in the six CRM stimuli were also more pronounced in the time-forward condition. With these meaningful stimuli, performance was best with six talkers and a single message and worst with synchronous presentation of a single talker producing six messages. This suggests that the effect of the number of talkers observed in Experiment I and in other studies may have more to do with the increase in the number of different messages than with the number of different voices. Although increasing the number of messages leads to poorer performance, increasing the number of voices can facilitate selective listening when there are multiple stimuli in the babble, especially when they are presented synchronously. Overall, the results of Experiment II were generally consistent with explanations in terms of the segregation of the target and competing stimuli on the basis of spectral and temporal information, as well as potential modulation-masking mechanisms across the conditions. However, the effects of asynchrony, number of talkers, and number of messages were much more pronounced with meaningful time-forward babble, despite the fact that the time-reversed and modulated-noise competition shared many of the acoustic properties of the meaningful competition.

Finally, Experiment III examined the effects of aging and hearing loss on CRM performance for the same eight conditions included in Experiment II (time-forward CRM

only). The CRM target and competition were spectrally shaped to restore audibility. The results indicated an overall performance deficit in the older participants of about 15–20 RAUs, although not quite statistically significant ( $p = .056$ ), but very similar relative performance for both groups across the eight listening conditions with different potential sources of interference. Similar performance decrements have been observed in older adults for asynchronous CRM sentences with only one competing talker (Lee & Humes, 2012).

## Acknowledgments

This work was supported, in part, by a research grant (R01 AG008293) from the National Institute on Aging. The authors thank Kristen Baisley and Kristin Quinones for help with data collection.

## References

- Akeroyd, M. (2008). Are individual differences in speech reception related to individual differences in cognitive ability: A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, 47(Suppl. 2), S53–S71.
- American National Standards Institute. (1999). *Maximum permissible ambient levels for audiometric test rooms* (ANSI S3.1-1999). New York: Author.
- American National Standards Institute. (2004). *Specification for audiometers* (ANSI S3.6-2004). New York: Author.
- Bolia, R. S., Nelson, W. T., Ericson, M. A., & Simpson, B. D. (2000). A speech corpus for multitalker communication research. *The Journal of the Acoustical Society of America*, 107, 1065–1066.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Bronkhorst, A., & Plomp, R. (1992). Effects of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing. *The Journal of the Acoustical Society of America*, 92, 3132–3139.
- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, 109, 1101–1109.
- Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, 110, 2527–2538.
- Calandruccio, L., Brouwer, S., Van Engen, K. J., Dhar, S., & Bradlow, A. R. (2013). Masking release due to linguistic and phonetic dissimilarity between the target and masker speech. *American Journal of Audiology*, 22, 157–164.
- Calandruccio, L., Dhar, S., & Bradlow, A. R. (2010). Speech-on-speech masking with variable access to the linguistic content of the masker speech. *The Journal of the Acoustical Society of America*, 128, 860–869.
- Calandruccio, L., & Zhou, H. (2014). Increase in speech recognition due to linguistic mismatch between target and masker speech: Monolingual and simultaneous bilingual performance. *Journal of Speech, Language, and Hearing Research*, 57, 1089–1097.
- Carhart, R., Tillman, T. W., & Greetis, E. S. (1969). Perceptual masking in multiple sound backgrounds. *The Journal of the Acoustical Society of America*, 45, 694–703.

- Carlyon, R. P.** (2004). How the brain separates sounds. *Trends in Cognitive Science*, 8, 465–471.
- Cruikshanks, K. J., Zhan, W., & Zhong, W.** (2010). Epidemiology of age-related hearing impairment. In S. Gordon-Salant, R. D. Frisina, R. R. Fay, & A. N. Popper (Eds.), *The aging auditory system* (pp. 259–274). New York, NY: Springer.
- Dirks, D. D., & Bower, D. R.** (1969). Masking effects of speech competing messages. *Journal of Speech and Hearing Disorders*, 12, 229–245.
- Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B.** (2005a). Word recognition in noise at higher-than-normal levels: Decreases in scores and increases in masking. *The Journal of the Acoustical Society of America*, 118, 914–922.
- Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B.** (2005b). Recognition of filtered words in noise at higher-than-normal levels: Decreases in scores with and without increases in masking. *The Journal of the Acoustical Society of America*, 118, 923–933.
- Duquesnoy, A. J.** (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *The Journal of the Acoustical Society of America*, 74, 739–743.
- Durlach, N.** (2006). Auditory masking: Need for improved conceptual structure. *The Journal of the Acoustical Society of America*, 120, 1787–1790.
- Durlach, N. I., Mason, C. R., Kidd, G., Jr., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G.** (2003). A note on informational masking. *The Journal of the Acoustical Society of America*, 113, 2984–2987.
- Eddins, D. A., & Liu, C.** (2012). Psychometric properties of the coordinate response measure corpus with various types of background interference. *The Journal of the Acoustical Society of America*, 131, EL177–EL183.
- Festen, J. M., & Plomp, R.** (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, 88, 1725–1736.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R.** (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.
- 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.
- Gallun, F., & Souza, P.** (2008). Exploring the role of the modulation spectrum in phoneme recognition. *Ear and Hearing*, 29, 800–813.
- Humes, L. E.** (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology*, 18, 590–603.
- Humes, L. E., & Coughlin, M. P.** (2009). Aided speech-identification performance in single-talker competition by older adults with impaired hearing. *Scandinavian Journal of Psychology*, 50, 485–494.
- Humes, L. E., & Dubno, J. R.** (2010). Factors affecting speech understanding in older adults. In S. Gordon-Salant, R. D. Frisina, R. R. Fay, & A. N. Popper (Eds.), *The aging auditory system* (pp. 211–258). New York, NY: Springer.
- Humes, L. E., Kidd, G. R., & Lentz, J. J.** (2013). Auditory and cognitive factors underlying individual differences in aided speech understanding among older adults. *Frontiers in Systems Neuroscience*, 7, 55. doi:10.3389/fnsys.2013.00055
- Humes, L. E., Lee, J. H., & Coughlin, M. P.** (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *The Journal of the Acoustical Society of America*, 120, 2926–2937.
- Iyer, N., Brungart, D. S., & Simpson, B. D.** (2010). Effects of target-masker contextual similarity on the multimasker penalty in a three-talker diotic listening task. *The Journal of the Acoustical Society of America*, 128, 2998–3010.
- Kochkin, S.** (2009). MarkeTrak VIII: 25-year trends in the hearing health market. *Hearing Review*, 16(11), 12–31.
- Lee, J. H., & Humes, L. E.** (2012). Effect of fundamental-frequency and sentence-onset differences on speech-identification performance of young and older adults in a competing-talker background. *The Journal of the Acoustical Society of America*, 132, 1700–1717.
- Miller, G. A.** (1947). The masking of speech. *Psychological Bulletin*, 44, 105–129.
- Pollack, I., & Pickett, J. M.** (1958). Stereophonic listening and speech intelligibility against voice babble. *The Journal of the Acoustical Society of America*, 30, 131–133.
- 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, 1274–1277.
- Rosen, S., Souza, P., Ekelund, C., & Majeed, A. A.** (2013). Listening to speech in a background of other talkers: Effects of talker number and noise vocoding. *The Journal of the Acoustical Society of America*, 133, 2431–2443.
- 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.
- Stone, M. A., Anton, K., & Moore, B. C. J.** (2012). Use of high-rate envelope speech cues and their perceptually relevant dynamic range for the hearing impaired. *The Journal of the Acoustical Society of America*, 132, 1141–1151.
- Stone, M. A., Füllgrabe, C., Mackinnon, R. C., & Moore, B. C. J.** (2011). The importance for speech intelligibility of random fluctuations in “steady” background noise. *The Journal of the Acoustical Society of America*, 130, 2874–2881.
- Stone, M. A., Füllgrabe, C., & Moore, B. C. J.** (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *The Journal of the Acoustical Society of America*, 132, 317–326.
- Stone, M. A., & Moore, B. C. J.** (2014). On the near non-existence of pure energetic masking release for speech. *The Journal of the Acoustical Society of America*, 135, 1967–1977.
- Studebaker, G. A.** (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Disorders*, 28, 455–462.
- Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M., & Gwaltney, C. A.** (1999). Monosyllabic word recognition at higher-than-normal speech and noise levels. *The Journal of the Acoustical Society of America*, 105, 2431–2444.
- Van Engen, K. J., & Bradlow, A.** (2007). Sentence recognition in native- and foreign-language multi-talker background noise. *The Journal of the Acoustical Society of America*, 121, 519–526.