Recognition memory in noise for speech of varying intelligibility[a]

Rachael C. Gilbert[b]
Department of Linguistics, University of Texas at Austin, Austin, Texas 78712

Bharath Chandrasekaran
Department of Communication Sciences and Disorders, University of Texas at Austin, Austin, Texas 78712

Rajka Smiljanic
Department of Linguistics, University of Texas at Austin, Austin, Texas 78712

(Received 17 October 2012; revised 31 October 2013; accepted 20 November 2013)

This study investigated the extent to which noise impacts normal-hearing young adults' speech processing of sentences that vary in intelligibility. Intelligibility and recognition memory in noise were examined for conversational and clear speech sentences recorded in quiet (quiet speech, QS) and in response to the environmental noise (noise-adapted speech, NAS). Results showed that (1) increased intelligibility through conversational-to-clear speech modifications led to improved recognition memory and (2) NAS presented a more naturalistic speech adaptation to noise compared to QS, leading to more accurate word recognition and enhanced sentence recognition memory. These results demonstrate that acoustic-phonetic modifications implemented in listener-oriented speech enhance speech-in-noise processing beyond word recognition. Effortful speech processing in challenging listening environments can thus be improved by speaking style adaptations on the part of the talker. In addition to enhanced intelligibility, a substantial improvement in recognition memory can be achieved through speaker adaptations to the environment and to the listener when in adverse conditions. © *2014 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4838975]

PACS number(s): 43.70.Mn, 43.71.Sy, 43.72.Dv [BRM] Pages: 389–399

I. INTRODUCTION

Most communicative environments involve some degree of background noise. In everyday listening situations, accurate speech perception relies on the capacity of the auditory system to process degraded speech signals. This requires stable sensory representations and considerable cognitive effort to extract the signal from noise. Such a task is challenging even for listener groups with normal hearing and normal cognitive abilities (Assmann and Summerfield, 2004; Rogers *et al.*, 2006). Speech recognition in adverse conditions impacts perceptual processes, mental representations, and linguistic and memory functions (Mattys *et al.*, 2012). The current paper examines two loci of the impact of adverse listening conditions: word recognition and recognition memory for spoken sentences. Specifically, we examine whether difficulty of speech processing in noise can be overcome by talker adaptations to challenging listening conditions: in response to noise and to listener perceptual difficulty.

Extensive research has shown that noise impacts speech processing adversely (Mattys *et al.*, 2009; Summers *et al.*, 1988; Junqua, 1996; Assman and Summerfield, 2004). In terms of memory, Murphy *et al.* (2000) found that digitized word pairs were significantly more difficult to recall in 12-talker babble than in quiet at both low [-5 dB signal-to-noise ratio (SNR)] and moderate (-10 dB SNR) noise levels. Using sentences masked by 8-talker babble, Pichora-Fuller *et al.* (1995) found that an increase in noise level from +5 dB SNR to 0 dB SNR resulted in a significant drop in word recall for both younger and older adults. In another study, young adults exhibited poorer recall for spoken digits masked by narrowband noise compared to digits in quiet, even when these same digits were accurately identified (Rabbitt, 1968). Similarly, hearing-impaired older adults were found to recall fewer words compared to listeners with typical hearing, even though they identified words equally well (Rabbitt, 1991). These studies suggest that listening to speech that is difficult to process (due to hearing impairment or noise) adversely affects memory by detracting from encoding and rehearsal abilities.

These findings are in accord with a host of hypotheses which propose that stimuli that are easier to process will also be remembered better [cf. effortfulness hypothesis (EH): McCoy *et al.*, 2005, ease of language understanding (ELU): Rönnberg, 2003; Rönnberg *et al.*, 2008]. The shared assumption is that the processing resources are limited. When processing stimuli in challenging listening environments or when there is a mismatch between the signal and long-term memory representations (due to degraded, distorted or unintelligible signal), working-memory resources are allocated to inference making and disambiguation of the message. Fewer resources are then available for storing of speech content in memory. The corollary of these models is that reductions in the perceptual efforts required to successfully recognize a degraded signal should free up more processing resources

---

[a] Portions of this work were presented at the 164th meeting of the Acoustical Society of America in Kansas City, MO (October 2012).

[b] Author to whom correspondence should be addressed. Electronic mail: rachaelgilbert@utexas.edu

for memory-related tasks. However, work in perceptual fluency provides contrary evidence (Rhodes and Castel, 2008; Besken and Mulligan, 2013). Processing fluency reliably influences people's metacognitive judgments across a variety of different domains, such as judgments of distance or learnability (Alter and Oppenheimer, 2008; Besken and Mulligan, 2013; Rhodes and Castel, 2009). For instance, disfluently processed stimuli in both visual and auditory domains (e.g., words printed in a more difficult-to-read font and softer words) were interpreted to be farther from the subject's current position or to require more studying compared to easier to process stimuli. This perceptual interference reduced perceptual fluency (as measured by identification rates and reaction times) and impacted metacognitive judgments. However, it also led to equal or enhanced memory compared to intact control stimuli (Mulligan and Lozito, 2004; Nairne, 1988; Besken and Mulligan, 2013). In contrast to the above stated hypotheses, the actual memory performance is predicted to be equal or better for stimuli that require more cognitive effort to process.

It is well established that listener-oriented styles of speech production can enhance intelligibility, providing a release in perceptual effort (for a review of methodologies and results see Smiljanic and Bradlow, 2009). Talkers naturally adopt a "clear" speaking style when they are aware of a speech perception difficulty on the part of the listener (e.g., hearing impairment or low proficiency in the language) (Picheny *et al.*, 1986; Schum, 1996; Krause and Braida, 2002; Ferguson and Kewley-Port, 2002; Ferguson, 2004; Smiljanic and Bradlow, 2005, 2009). Clear speech modifications typically involve a decrease in speaking rate, a wider dynamic F0 range, greater sound-pressure levels, more salient stop releases, an expanded vowel space, greater obstruent RMS energy, and increased energy in 1000–3000 Hz range of long-term spectra (Smiljanic and Bradlow, 2005; Picheny *et al.*, 1986; Krause and Braida, 2004; Bradlow *et al.*, 2003; Liu *et al.*, 2004; Ferguson and Kewley-Port, 2002). The clear speech benefit has been documented for listeners in different languages, for young and old adults, for native and nonnative listeners, and for listeners with hearing impairments (Picheny *et al.*, 1985; Smiljanic and Bradlow, 2005, 2011; Smiljanic, 2013; Smiljanic and Sladen, 2013; Bradlow and Bent, 2002; Bradlow *et al.*, 2003; Liu *et al.*, 2004).

The majority of the studies investigating the clear speech intelligibility benefit have used speech signals that were recorded under quiet conditions and then tested after mixing with noise. This may be an issue given that "the speech signal presented to perceivers has no correlation with the noise accompanying it and may, therefore, introduce yet one more distortion in the already strained realism of laboratory test conditions" (Chung *et al.*, 2005, p. 55). Speech produced in the presence of actual noise, or noise-adapted speech (NAS, i.e., Lombard speech), thus presents a more realistic adaptation to such adverse listening conditions. Similar to clear speech, NAS is characterized by a decrease in speaking rate, an increase in intensity, higher peak F0, a flattening of spectral tilt, and an increase in average F0 (Lombard, 1911; Summers *et al.*, 1988; Junqua, 1993; Lane and Tranel, 1971).[1] Previous research has shown that NAS is more intelligible in noise than quiet speech (QS) mixed with noise, both for native and nonnative listeners (Dreher and O'Neill, 1957; Pittman and Wiley, 2001; Summers *et al.*, 1988; Cooke and Lecumberri, 2012).

While the beneficial effects of these two intelligibility-enhancing speaking style modifications are well established in terms of word recognition, very little is known about whether variation in intelligibility impacts the encoding of speech in memory. In a recent study, Van Engen *et al.* (2012) examined whether semantic and phonetic enhancements aided sentence recognition memory. Sentences that were easier to process in noise (meaningful and clear vs anomalous and conversational) were also easier to recognize as previously heard (old vs new). The recognition memory benefit may be due to the availability of either more salient acoustic-phonetic and semantic cues which made encoding more robust, or more working-memory resources available for memory recognition. Importantly, all the sentences in the memory experiment were presented in quiet and were thus completely intelligible to the listeners, minimizing the likelihood that the recognition results are merely a consequence of the poor encoding of the less intelligible stimuli. The extent to which the sentence recognition memory benefit for easier-to-process sentences applies to speech in noise has not been examined thus far.

The present study investigated the extent to which intra-talker variation in speech clarity impacted recognition memory for sentences in noise. We also compared the processing of speech recorded in quiet and mixed with noise with that of speech recorded in response to the environmental noise (NAS). One prediction stemming from the EH and ELU is that sentences in which linguistic input is better specified (through acoustic-phonetic enhancements associated with clear speech and NAS) and perceptual effort is reduced, more resources will be available for encoding speech in memory. Both clear and noise-adapted speech modifications are thus expected to enhance sentence recognition memory for listeners. Furthermore, clear speech produced in response to noise will show a cumulative processing benefit. That is, noise-adapted clear speech will be the most intelligible and (per these models) provide the largest recognition memory boost for sentences processed in noise. Alternatively, if processing difficulty impacts only word identification rate but not sentence recognition memory (cf. processing fluency), we would expect conversational sentences and sentences produced in quiet to be recognized as equally well as or more accurately than NAS and clear sentences.

Experiment 1 examined the intelligibility of conversational and clear speech produced in quiet and in the presence of noise for normal-hearing, native adult speakers of English. In Experiment 2, we tested the extent to which listener-oriented conversational and clear speech produced in quiet and in response to noise affect recognition memory for sentences presented in noise. Acoustic analyses were also performed on the sentences in order to examine the acoustic-phonetic changes characteristic of clear speech and noise-adapted speech. The results provide a direct comparison between listener- and environment-oriented acoustic-phonetic enhancements, as well as their effects on speech intelligibility. Furthermore, this is one of the first studies to investigate the extent to which clear speech and noise-adapted speech benefits extend to speech processing tasks in noise beyond word recognition.

II. EXPERIMENT 1: INTELLIGIBILITY

A. Materials

The stimuli consisted of 80 meaningful sentences modified from the Basic English Lexicon (BEL) sentence materials (Calandruccio and Smiljanic, 2012; Van Engen *et al.*, 2012). Sentences each contained four keywords for intelligibility scoring (e.g., *The small animal scared the baby*). One native female speaker of American English (aged 26 yr, with no speech or hearing impairment) was recorded producing the full set of 80 meaningful sentences over two sessions.

The two sessions differed in the type of talker response to the environment. In the first session, the speaker read the stimuli *in quiet* (quiet speech; QS). In the second session, the speaker read the same sentences *in the presence of 6-talker babble* presented via Sennheiser HD5 headphones (noise-adapted speech; NAS). The 6-talker babble was composed of six monolingual speakers of American English (three females and three males between the ages of 28 and 48 yr) producing semantically anomalous sentences in English (Van Engen and Bradlow, 2007). Anomalous sentences were used to minimize chance that listeners might extract a meaningful sentence other than the target. Both recording sessions took place in a sound-attenuated booth.

In each session, all sentences were read once in a conversational speaking style and once in a clear speaking style. Conversational speech was elicited by instructing the speaker to speak in a casual, conversational style, as if she was talking to someone familiar with her voice and speech patterns. The clear speech was elicited by instructing her to speak as though the listener was having a hard time understanding her, whether due to hearing difficulty or because the listener was a non-native speaker of English. For the speech produced in noise, the same instructions for eliciting the two listener-oriented speaking styles were given as in quiet. These instructions have been shown to be sufficient for elicitation of different speaking styles (for a review of clear speech and Lombard speech elicitation instructions, see Pichora-Fuller *et al.*, 2010 or Smiljanic and Bradlow, 2009).

Sentences were presented to the speaker one at a time on a computer monitor. Recordings were made using a Shure SM10A head-mounted microphone and a MOTU UltraLite-MK3 Hybrid recorder. The recorded sentences were segmented into individual files which were equalized for RMS amplitude across the entire sentence duration. The total set of recorded sentences was 320 (80 conversational QS; 80 clear QS; 80 conversational NAS; 80 clear NAS).

B. Listeners

Sixteen adults between the ages of 18 and 34 served as listeners. All participants were native, monolingual speakers of American English. All were University of Texas at Austin undergraduate students. They all passed a hearing-screening test [thresholds <25 dB hearing level (HL) at 1, 2, and 4 kHz]. Participants provided written informed consent and were either paid for their participation or received course credit.

C. Procedure

All sentences were presented to listeners for assessment of intelligibility. Speech-shaped noise (SSN) was created for each sentence type (conversational QS, clear QS, conversational NAS, and clear NAS) by filtering white noise to the long-term average spectrum of the full set of sentences. This approach was used to ensure that masking was constant across the recording types (following Van Engen *et al.*, 2012). Each file was digitally mixed with noise at a SNR of -5 dB sound pressure level (SPL). The SNR of -5 dB was determined from pilot testing to ensure that listeners would not perform at the ceiling in the easiest listening condition. Each of the final stimulus files consisted of a 400 ms silent lead, followed by 500 ms of noise, followed by the speech-plus-noise files, and ending with 500 ms of only noise. The noise preceding and following the speech stimulus was at the same level as the noise mixed with the speech.

For the intelligibility test, each participant was seated in front of a computer monitor in a sound-attenuated booth in the Phonetics Laboratory at the University of Texas at Austin. The stimuli were played over headphones (Sennheiser HD570 or Sony MDR-CD780) at a comfortable listening level set by the experimenter. Instructions and stimuli were presented using EPrime (Schneider *et al.*, 2002). The participant's task was to listen to each sentence and write down as much as they could onto a prepared answer sheet. After each trial, the participant pressed a button on the keyboard to move onto the next trial. Each trial was presented only once, but participants could take as much time as they wished to write down the sentences.

In order to familiarize participants with the task, the session began with four practice items not included in the subsequent test. Each participant then transcribed a total of 80 pseudorandomized sentences from either the QS recordings or the NAS recordings. Forty of these sentences were produced in conversational speech, and 40 were produced in clear speech; this was counterbalanced so that half of the participants heard sentences produced in the opposite speaking styles as the other half of the participants. Participants never heard the same sentence twice. Each sentence was scored by the number of keywords correctly identified (four per sentence) for a total of 160 keywords per sentence type per listener. In order to be considered correct, no morphemes could be added to or deleted from the keywords, but homophones were acceptable.

D. Results

Intelligibility results showed that, for QS, listeners identified 38.71% of the keywords in conversational speech [standard deviation (SD): 9.81%] and 77.19% of the keywords in clear speech (SD: 7.43%). For NAS, listeners correctly identified 52.25% of the keywords in conversational speech (SD: 9.03%) and 86.81% of the keywords in clear

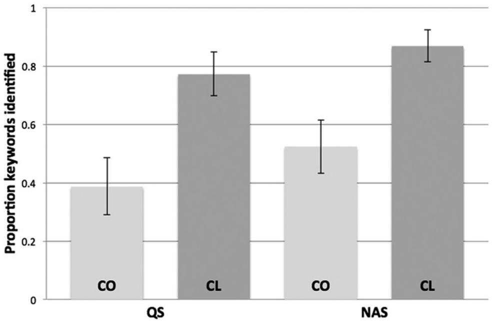

FIG. 1. Average proportion of keywords identified from conversational and clear sentences produced in quiet (QS) and in noise (NAS). Error bars represent standard error.

speech (SD: 5.46%). Proportion correct scores for conversational and clear sentences produced in QS and NAS are shown in Fig. 1.

Intelligibility data were analyzed with a linear mixed effects logistic regression using the lme4 package in R, consistent with the intelligibility analysis used in Van Engen *et al.* (2012). Running a regression on the results showed that variance due to Sentence (our second random effect) was 0.79—a good confirmation of including a second random effect. Since we wanted to include more than one random effect in our Experiment 1 analyses, it no longer met the assumptions of General Linear Models (independence assumption), and thus we chose to use mixed effects regression over analysis of variance (ANOVA). This statistical model allowed us to account for variance due to both Listener and Sentence. Keyword identification (i.e., correct or incorrect) was the dichotomous dependent variable. Listener and Sentence were included in the model as random factors and Listener-Oriented Speaking Style (conversational or clear), Environment-Oriented Speaking Style (QS or NAS), and their interactions as fixed effects. Random slopes were not included in the model, since we did not believe the effects of the speaking styles to vary by Sentence or Subject (similar to Van Engen *et al.*, 2012). Listener-Oriented Speaking Style was contrast coded (−0.5, 0.5) such that negative beta values are associated with clear speech and positive beta values are associated with conversational speech. Environment-Oriented Speaking Style was also contrast coded (−0.5, 0.5) such that negative beta values are associated with speech produced in response to noise and positive beta values are associated with speech produced in quiet. The results of the regression are presented in Table I.

The results revealed that the overall probability of correct keyword identification was significantly higher for NAS vs QS speech ($p < 0.001$) and for clear vs conversational speech ($p < 0.001$). Results also revealed a significant interaction between Listener-Oriented Speaking Style and Environment-Oriented Speaking Style ($p = 0.026$). The nature of this interaction was examined by performing a second round of mixed-effects logistic regressions on the QS and NAS sets individually. The results of these regressions are shown in Table II. Although the effect of listener-oriented speaking style was a highly significant predictor of correct keyword identification for both QS and NAS, the effect of style was greater for NAS ($\beta_{QS} = -2.033$ whereas $\beta_{NAS} = -2.291$).

Experiment 2 investigates the extent to which these differences in intelligibility impact recognition memory.

III. EXPERIMENT 2: RECOGNITION MEMORY

A. Materials

Forty sentences each in conversational and clear speech from the QS and NAS conditions (160 total) were used in the recognition memory experiment. This subset of the total sentences from Experiment 1 was selected to constrain the length of the experiment. The QS sentences were the same meaningful sentences used in the Van Engen *et al.* (2012) study. Results from Van Engen *et al.* showed that the QS sentences used as old and new for recognition memory did not differ in intelligibility. In order to confirm this pattern for NAS sentences used in the current study, an unpaired, 2-tailed *t*-test was conducted. The test showed no significant difference between the intelligibility of NAS old and new sentences.[2]

B. Listeners

Sixty native speakers of American English between the ages of 18 and 30 participated in the experiment. They were drawn from the same participant pool as in Experiment 1 but were different individuals. The same inclusion criteria applied as in Experiment 1. Participants provided written informed consent and were either paid for their participation or received course credit.

C. Procedure

The same 6-talker babble file played through headphones to the speaker during the production of NAS served as the noise masker in the exposure phase. To avoid masker familiarization, each meaningful sentence was digitally mixed with one of four different portions of the babble file. Sentences were mixed at signal-to-noise ratios of 0 and

TABLE I. Results of the linear mixed effects logistic regression on intelligibility data for all sentences.

						95% Confidence interval	
Fixed effects:	Estimate	Odds ratio	Std. error	z value	p	Lower bound	Upper bound
(Intercept)	0.948	2.580	0.126	7.520	<0.001***	2.012	3.316
Environment-oriented speaking style	−2.120	0.120	0.056	−38.180	<0.001***	0.269	0.503
Listener-oriented speaking style	−0.998	0.369	0.155	−6.460	<0.001***	0.107	0.134
Environment-oriented speaking style: Listener oriented speaking style	0.237	1.269	0.107	2.230	0.026*	1.029	1.565

TABLE II. Results of the linear mixed effects logistic regression on intelligibility data for sentences produced in quiet (QS) and in response to noise (NAS).

	Fixed effects:	Estimate	Odds ratio	Std. error	z value	p	95% Confidence interval Lower bound	95% Confidence interval Upper bound
QS	(Intercept)	0.454	1.575	0.131	3.455	<0.001***	1.213	2.048
	Listener-oriented speaking style	−2.033	0.131	0.071	−28.790	<0.001***	0.114	0.151
NAS	(Intercept)	1.503	4.494	0.178	8.421	<0.001***	3.151	6.463
	Listener-oriented speaking style	−2.291	0.101	0.088	−25.948	<0.001***	0.085	0.120

+3 dB SPL. Although presenting sentences at an SNR of −5 dB (that of Experiment 1) would provide a more direct comparison, pilot testing revealed that Experiment 2 required an easier SNR range to avoid results at floor. The two SNRs were chosen to investigate whether intelligibility-enhancing speaking styles prove more critical to recognition memory processes as the signal became further degraded.

Each of the final stimulus files consisted of a 400 ms silent lead, followed by 500 ms of noise, followed by the speech-plus-noise files, and ending with 500 ms of only noise. The noise preceding and following the speech stimulus was at the same level as the noise mixed with the speech.

Testing setup was the same as in Experiment 1. Participants took part in one of four experimental conditions: recognition memory for conversational and clear QS mixed with noise at 0 dB SNR ($n = 15$), for conversational and clear QS mixed with noise at +3 dB SNR ($n = 15$), for conversational and clear NAS mixed with noise at 0 dB SNR ($n = 15$), and for conversational and clear NAS mixed with noise at +3 dB SNR ($n = 15$). In each experimental condition, listeners were first exposed to 40 unique sentences embedded in the 6-talker babble and instructed to try to commit them to memory (the exposure phase). Twenty of the sentences were presented in conversational speech and 20 in clear speech. Sentences were presented only once with a 500 ms break between sentences. In the test phase, participants were instructed to listen to a second set of sentences and indicate by pressing one of two buttons whether each sentence was old (from the exposure phase) or new. All 40 of the exposure sentences were included along with 40 new sentences. Half of the sentences were in conversational speech and half were in clear speech. Sentences in the test phase were presented in quiet since the goal of this study was to examine the effects of noise on the initial speech processing (encoding) during which listeners were trying to commit sentences to memory.

The participants were instructed to press a third button on the response box to move from trial to trial. Each trial was presented only once. In both phases, sentence order was randomized for each participant.

D. Results

Recognition accuracy was analyzed using the d' statistic in order to provide a measure of accuracy independent of response bias (Lamont *et al.*, 2005). The d' statistic accounts for discrimination sensitivity and bias in each participant by subtracting the normalized probability of false alarms (i.e., identifying a new sentence as old) from the normalized probability of hits (i.e., identifying an old sentence as old), and then correcting the formula to account for values of 0 and 1. Table III lists all normalized hit rates, false alarm rates, d′ statistics, and C scores (a measure of changes in response bias). Table IV lists all confidence intervals. The average C scores across all conditions are positive, meaning that participants were generally biased to respond “new” more often than “old.” This bias was stronger for clear speech than for conversational speech. The overall results of Experiment 2 are shown in Fig. 2.

In line with the recognition memory analysis used in Van Engen *et al.* (2012), we ran an omnibus ANOVA on the d′ scores. D′ aggregates results across sentences, thus disallowing Sentence to be included as a random effect, unlike in Experiment 1. Since we only need to consider Subject as a random effect, an ANOVA is a sufficient model. D′ scores were submitted to a mixed ANOVA with Listener-Oriented Speaking Style (conversational or clear) as the within-subject factor and Environment-Oriented Speaking Style (QS or NAS) and SNR (0 dB or +3 dB) as between-subject factors. There was a main effect of Listener-Oriented Speaking Style [$F(1,56) = 22.310$, $p < 0.001$, $\eta_p^2 = 0.285$] and of Environment-Oriented Speaking Style [$F(1,1) = 10.223$, $p = 0.002$, $\eta_p^2 = 0.154$] on d′ scores. The effect of SNR (0 dB vs + 3 dB) was not significant [$F(1,1) = 0.105$, $p = 0.747$, $\eta_p^2 = 0.002$]. No significant interactions between

TABLE III. Normalized hit rates, false alarm rates, d′, and C values for the recognition memory test for sentences produced in quiet (QS) and in noise (NAS).

	SNR	Conversational speech Hit rate	Conversational speech False alarm rate	Conversational speech d′	Conversational speech C	Clear speech Hit rate	Clear speech False alarm rate	Clear speech d′	Clear speech C
QS	0	0.51	0.31	0.55	0.24	0.58	0.20	1.14	0.36
	3	0.60	0.31	0.76	0.13	0.64	0.26	1.12	0.19
NAS	0	0.60	0.20	1.17	0.34	0.66	0.14	1.56	0.33
	3	0.64	0.26	1.12	0.18	0.58	0.17	1.26	0.40

TABLE IV. Pairwise comparisons of the recognition memory data for speech produced in noise (NAS) vs quiet (QS); for clear speech (CL) vs conversational speech (CO); for sentences heard at an easier SNR (+3 dB) vs a harder SNR (0 dB).

					95% Confidence interval	
Factor	Levels	Mean difference	Std. error	p	Lower bound	Upper bound
Environment-oriented speaking style	NAS-QS	.387*	0.121	0.002	0.144	0.629
Listener-oriented speaking style	CL-CO	.371*	0.079	<0.001	0.214	0.528
SNR	3 dB-0dB	0.039	0.121	0.747	−0.203	0.282

Listener-Oriented Speaking Style, Environment-Oriented Speaking Style, and SNR were found [Listener-Oriented Speaking Style by Environment-Oriented Speaking Style: $F(1,1) = 1.718$, $p = 0.195$, $\eta_p^2 = 0.030$, Listener-Oriented Speaking Style by SNR: $F(1,1) = 2.261$, $p = 0.138$, $\eta_p^2 = 0.039$, Environment-Oriented Speaking Style by SNR: $F(1,1) = 1.277$, $p = 0.263$, $\eta_p^2 = 0.022$, Listener-Oriented Speaking Style by Environment-Oriented Speaking Style by SNR: $F(1,1) = 0.010$, $p = 0.920$, $\eta_p^2 = 0.000$].

The results showed that speech clarity significantly contributed to listeners' enhanced recognition memory for sentences. Listeners were better able to recognize previously heard sentences when they were produced in clear speech relative to conversational speech and in NAS relative to QS. This effect was significant even though listeners were committing these sentences to memory in noise. Finally, the results showed that recognition memory scores did not significantly differ across QS and NAS clear speech, unlike in Experiment 1.

IV. ACOUSTIC ANALYSES

A. Procedure

A series of acoustic analyses was performed on all sentences in order to assess whether the two listener-oriented speaking styles (conversational vs clear) and the two environment-oriented speaking styles (QS vs NAS) differed in their acoustic-articulatory characteristics. Four specific acoustic-articulatory features were measured: speech rate (syllables per second), F0 range (Hz), mean F0 (Hz), and energy in the 1–3 kHz range (dB). All acoustic features were measured per sentence and then averaged to obtain listener-oriented (conversational, clear) and environment-oriented (QS, NAS) speaking style values. Speech rate was calculated as the number of syllables divided by the sentence duration, excluding pauses greater than 100 ms. F0 range was calculated as the difference between the sentence's minimum F0 value and peak F0 value. Mean F0 was an average of F0 values over the entire sentence. Energy in the 1–3 kHz range was measured by averaging the long-term average spectrum energy between 1 and 3 kHz across each unleveled sentence. We focused on these temporal-, pitch-, and spectral-related features, as they are typical metrics for contrasting conversational speech against clear speech as well as quiet speech against Lombard speech (Pichora-Fuller *et al.*, 2010; Smiljanic and Bradlow, 2009).

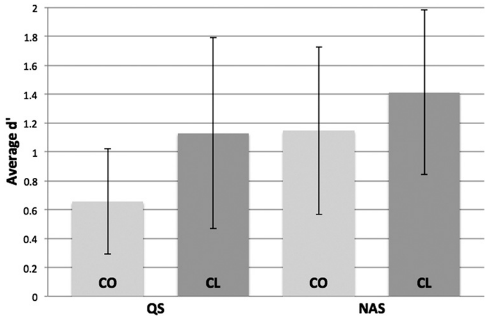

FIG. 2. Average d′ scores for conversational and clear sentences produced in quiet (QS) and in noise (NAS). Error bars represent standard error.

B. Results

The results of four acoustic measurements are given in Table V. For each of the four measurements, results were submitted to repeated measures ANOVA with Listener-Oriented Speaking Style (conversational or clear) and Environment-Oriented Speaking Style (QS or NAS) as the within-sentence factors. This is consistent with the acoustic analyses from Van Engen *et al.* (2012)—since there is no random effect of Listener, we did not need to use mixed effects regression, unlike in Experiment 1. In order to ensure that the same acoustic trends characterized materials used in both experiments, these same analyses were repeated on the sentence subset used in Experiment II (see Table VIII for a list of statistics). Confidence intervals for the entire set and the Experiment II subset can be found in Table IX.

For speech rate, there were main effects of Listener-Oriented Speaking Style [$F(1,159) = 3050.756$, $p < 0.001$, $\eta_p^2 = 0.950$] and Environment-Oriented Speaking Style [$F(1,159) = 279.977$, $p < 0.001$, $\eta_p^2 = 0.638$], with clearly produced sentences showing significantly slower speech rates than sentences in conversational speech, and NAS sentences showing significantly slower speech rates than QS sentences. There was no significant interaction between Listener-Oriented Speaking Style and Environment-Oriented Speaking Style [$F(1,159) = 0.025$, $p = 0.875$, $\eta_p^2 = 0.000$].

Significant main effects of Listener-Oriented Speaking Style and of Environment-Oriented Speaking Style were found for F0 range: [$F(1,159) = 100.200$, $p < 0.001$, $\eta_p^2 = 0.387$], [$F(1,159) = 14.233$, $p < 0.001$, $\eta_p^2 = 0.082$]. Sentences in clear speech showed significantly larger F0 ranges than sentences in conversational speech, whereas NAS showed significantly smaller F0 ranges than QS sentences. There was no significant interaction [$F(1,159) = 1.719$, $p = 0.192$, $\eta_p^2 = 0.011$].

There were significant main effects of Listener-Oriented Speaking Style and Environment-Oriented Speaking Style

TABLE V. Acoustic measures of sentence materials as produced in quiet (QS), in noise (NAS), in conversational speech (CO), and in clear speech (CL). Within each cell, the number on the left represents the average for all sentences (Experiment I) and the number on the right represents the average for the subset of sentences used in Experiment II.

Mean (SD)	QS CO		QS CL		NAS CO		NAS CL	
Speech rate (syllables/s)	5.30 (0.98)	5.14 (0.92)	2.81 (0.52)	2.79 (0.53)	4.75 (0.64)	4.66 (0.61)	2.27 (0.50)	2.29 (0.57)
Average F0 (Hz)	161.24 (8.85)	161.19 (9.22)	167.42 (7.90)	166.33 (7.95)	179.49 (5.78)	179.66 (5.21)	198.22 (7.32)	198.04 (7.38)
F0 range (Hz)	136.79 (108.25)	137.96 (108.71)	215.63 (122.69)	210.69 (119.18)	117.94 (38.44)	123.05 (43.46)	179.63 (60.52)	179.59 (59.01)
Energy: 1–3 kHz	22.61 (2.52)	22.64 (2.35)	22.17 (2.63)	22.83 (2.73)	28.60 (1.76)	28.76 (1.80)	28.66 (1.72)	28.98 (1.77)

on average F0: [$F(1,159) = 527.344$, $p < 0.001$, $\eta_p^2 = 0.768$], [$F(1,159) = 2339.655$, $p < 0.001$, $\eta_p^2 = 0.936$]. Sentences in clear speech and NAS exhibited significantly higher mean F0 than sentences in conversational speech and QS speech, respectively. A significant interaction between Listener-Oriented Speaking Style and Environment-Oriented Speaking Style was also found [$F(1,159) = 140.765$, $p < 0.001$, $\eta_p^2 = 0.470$], with NAS clear speech exhibiting a significantly higher mean F0 than could be attributed to the cumulative effects of NAS speech and clear speech alone. This has been broken down into simple effects for both the entire set and the subset (cf. Table VI).

Finally, there was a significant main effect of Environment-Oriented Speaking Style on energy in 1–3 kHz range [$F(1,159) = 2736.700$, $p < 0.001$, $\eta_p^2 = 0.945$]; NAS sentences had significantly greater energy in this range than QS sentences. This trend was present but not statistically significant for Listener-Oriented Speaking Style [$F(1,159) = 3.676$, $p = 0.057$, $\eta_p^2 = 0.023$]. A significant interaction between Listener-Oriented Speaking Style and Environment-Oriented Speaking Style was found as well [$F(1,159) = 10.819$, $p = 0.001$, $\eta_p^2 = 0.064$]; conversational QS sentences exhibited slightly higher 1–3 kHz energy compared to clear QS sentences, whereas the inverse held true for NAS sentences. This is unusual, given that clear speech in quiet typically exhibits higher energy in the 1–3 kHz range than conversational speech (Van Engen *et al.*, 2012). The simple effects of this interaction are shown below in Table VII. The interaction did not extend to the subset of materials used in Experiment 2.

The analyses of sentences used in Experiment I showed that sentences produced in a clear speaking style were overall slower, had higher F0 means, and exhibited wider F0 ranges compared to sentences produced in conversational speech. Speaking in response to the environmental noise (NAS) lead to slower speaking rates, higher F0 means, and greater energy in the 1–3 kHz range compared to speech produced in quiet (QS). NAS sentences showed significantly smaller F0 ranges compared to QS sentences which differs from previous findings in which F0 range was increased in Lombard speech (cf. Jessen *et al.*, 2005). However, it is likely that the increased F0 mean found for NAS relative to QS sentences in this study resulted in smaller F0 variation. Although non-significant, conversational QS exhibited slightly more 1–3 kHz energy than clear QS, whereas the inverse pattern held true for conversational and clear NAS. Finally, the cumulative effect of listener- and environment-oriented speaking styles was manifested in clear NAS exhibiting a significantly higher mean F0 than clear QS.

TABLE VI. Simple effects of the mean F0 interaction for all sentences and the subset of sentences used in Experiment II, as produced in quiet (QS), in noise (NAS), in conversational speech (CO), and in clear speech (CL).

	Environment-oriented speaking style	Mean difference (CL-CO)	Std. error	p	95% Confidence interval Lower bound	95% Confidence interval Upper bound
Total	QS	6.185*	0.874	<0.001	4.458	7.912
	NAS	18.736*	0.62	<0.001	17.512	19.96
Exp. II subset	QS	5.137*	1.284	<0.001	2.58	7.693
	NAS	18.382*	0.916	<0.001	16.559	20.204

Acoustic analyses on the subset of sentences used in Experiment II revealed similar trends (as illustrated in Table VIII and Table IX) as those found for the entire set. The only exception was that the marginal effect of Listener-Oriented Speaking Style on energy in the 1–3 kHz range was less significant in the Experiment II stimuli. Additionally, the unexpected interaction between Listener- and Environment-Oriented Speaking Styles on energy in the 1–3 kHz range (conversational > clear) was not present in the subset. The subset of sentences used in Experiment II thus very closely resembled the acoustic patterns present in the overall set of sentences.

V. DISCUSSION

This study examined the extent to which speaking style adaptations to the environment and to the listener facilitate word recognition and recognition memory for spoken sentences in noise. The results of Experiment 1 showed that listener- and environment-oriented acoustic-phonetic enhancements to the speech signal (clear speech and NAS) resulted in increased intelligibility, as evidenced by improved word recognition in noise (see also Uchanski, 2005; Picheny *et al.*, 1985; Smiljanic and Bradlow, 2009; Pichora-Fuller *et al.*,

TABLE VII. Simple effects of the 1–3 kHz interaction for all sentences, as produced in quiet (QS), in noise (NAS), in conversational speech (CO), and in clear speech (CL).

Environment-oriented speaking style	Mean difference (CL-CO)	Std. error	p	95% Confidence interval Lower bound	95% Confidence interval Upper bound
QS	−0.432*	0.147	0.004	−0.722	−0.141
NAS	0.062	0.091	0.499	−0.118	0.241

TABLE VIII. A list of statistics for the repeated-measures ANOVA performed on the subset of sentences used in Exp. II.

	Listener-oriented speaking style	Environment-oriented speaking style	Environment-oriented speaking style: Listener oriented speaking style
Speech Rate	$F(1,79) = 1366.233, p < 0.001, \eta_p^2 = 0.945^*$	$F(1,79) = 93.143, p < 0.001, \eta_p^2 = 0.541^*$	$F(1,79) = 0.032, p = 0.858, \eta_p^2 = 0.000$
F0 Range	$F(1,79) = 44.043, p < 0.001, \eta_p^2 = 0.358^*$	$F(1,79) = 5.682, p = 0.020, \eta_p^2 = 0.067^*$	$F(1,79) = 0.668, p = 0.416, \eta_p^2 = 0.008$
F0 Mean	$F(1,79) = 215.032, p < 0.001, \eta_p^2 = 0.731^*$	$F(1,79) = 1237.327, p < 0.001, \eta_p^2 = 0.940^*$	$F(1,79) = 72.987, p < 0.001, \eta_p^2 = 0.480^*$
LTAS	$F(1,79) = 2.237, p = 0.139, \eta_p^2 = 0.028$	$F(1,79) = 1353.628, p < 0.001, \eta_p^2 = 0.945^*$	$F(1,79) = 0.022, p = 0.883, \eta_p^2 = 0.000$

2010). The results further demonstrated that the intelligibility benefit from NAS clear speech was significantly greater compared to QS clear speech. This suggests that the talker response to environmental noise (i.e., producing deliberate clear speech in response to the actual background babble) enhances speech in ways that differ from clear speech produced in quiet. Clear speech in response to noise is more resistant to the masking effect of noise. Finally, the clear speech intelligibility benefit may be diminished when the listening conditions that do not correlate with the conditions in which the speech was produced (Dreher and O'Neill, 1957; Pittman and Wiley, 2001; Summers *et al.*, 1988).

The results of Experiment 2 provided new evidence that clear speech and NAS benefits extend to better sentence recognition memory compared to conversational speech and QS. Speaking style adaptations that enhanced intelligibility in Experiment 1 also enhanced recognition memory in noise. This finding expands upon recent work examining the effects of clear speech and semantic context on recognition memory in quiet (Van Engen *et al.*, 2012). Here, we show that more intelligible speech—even when presented in noise—allows for better speech recognition memory. In all conditions, recognition memory for both clear speech and NAS was characterized by lower rates of false alarm responses (i.e., identifying new sentences as old) than their conversational speech and QS counterparts (Table III). The correlation between exaggerated acoustic-phonetic cues and lower false alarm rates may be indicative of enhanced memory traces for more distinctive speaking styles (Van Engen *et al.*, 2012).

With respect to our predictions, recognition memory results provided support for the Effortfulness Hypothesis (McCoy *et al.*, 2005) and the Ease of Language Understanding model (Rönnberg *et al.*, 2008). Sentences that were easier to process were also easier to recognize as old or new. This is in contrast to the prediction arising from the perceptual fluency hypothesis which states that perceptual interference may lead to enhanced sentence recall (Besken and Mulligan, 2013). The recognition memory benefit found in the current study could arise in one of the two ways. The resources for speech processing are pooled with those for recognition memory, so that any increased demands for one task diminish the mental resources available for the other. The other possibility is that speech encoded (memory representations) under challenging processing conditions is of poorer quality than representations produced under more favorable conditions (less well integrated with other memories or lower levels of activation). Van Engen *et al.* (2012) found a recognition memory benefit for easier-to-process meaningful and clear sentences, even

TABLE IX. Pairwise comparisons of acoustic data for speech produced in noise (NAS) vs quiet (QS), and for clear speech (CL) vs conversational speech (CO).

		Factor	Levels	Mean diff.	Std. error	p	95% Confidence interval Lower bound	95% Confidence interval Upper bound
Speech rate	Total	Environment-oriented speaking style	NAS-QS	-0.545^*	0.033	<0.001	−0.609	−0.48
		Listener-oriented speaking style	CL-CO	-2.487^*	0.045	<0.001	−2.576	−2.398
	Exp. II subset	Environment-oriented speaking style	NAS-QS	-0.494^*	0.051	<0.001	−0.596	−0.392
		Listener-oriented speaking style	CL-CO	-2.362^*	0.064	<0.001	−2.49	−2.235
F0 Range	Total	Environment-oriented speaking style	NAS-QS	-27.426^*	7.269	<0.001	−41.783	−13.068
		Listener-oriented speaking style	CL-CO	70.267*	7.02	<0.001	56.403	84.131
	Exp. II subset	Environment-oriented speaking style	NAS-QS	-23.007^*	9.652	0.02	−42.218	−3.796
		Listener-oriented speaking style	CL-CO	64.635^*	9.739	<0.001	45.25	84.021
F0 Mean	Total	Environment-oriented speaking style	NAS-QS	25.095^*	0.713	<0.001	23.675	26.515
		Listener-oriented speaking style	CL-CO	11.759^*	0.802	<0.001	10.163	13.355
	Exp. II subset	Environment-oriented speaking style	NAS-QS	25.095^*	0.713	<0.001	23.675	26.515
		Listener-oriented speaking style	CL-CO	11.759^*	0.802	<0.001	10.163	13.355
LTAS	Total	Environment-oriented speaking style	NAS-QS	6.244^*	0.119	<0.001	6.008	6.479
		Listener-oriented speaking style	CL-CO	−0.185	0.096	0.057	−0.375	0.006
	Exp. II subset	Environment-oriented speaking style	NAS-QS	6.133^*	0.167	<0.001	5.801	6.464
		Listener-oriented speaking style	CL-CO	0.201	0.134	0.139	−0.067	0.469

though all sentences were presented in quiet and thus entirely intelligible. Given these findings, we believe that the current sentence recognition memory results cannot be attributed solely to the intelligibility data. With regard to our initial question, the results do show that speaking style modifications enhance memory recognition as well as intelligibility in noise regardless of the exact underlying mechanism. Although listening in noise is an effortful process, listener- and environment-oriented speaking styles can mitigate the processing load. It remains to be examined further which of the above interpretations (or both) can account for clear speech and NAS facilitatory effects in speech processing in adverse conditions.

In contrast to Experiment 1, Experiment 2 did not yield an interaction between listener- and environment-oriented speaking styles. While the intelligibility gain for NAS clear speech was larger compared to that for QS clear speech, the same effect was not observed in sentence recognition memory. Given that both of the SNR conditions in Experiment 2 were more beneficial than in Experiment 1, it is possible that the additional benefit of NAS clear speech would also emerge for recognition memory in more challenging listening conditions, i.e., at lower SNRs.

The result of the acoustic analysis confirmed that both clear speech and NAS sentences exhibited typical intelligibility-enhancing features (Pichora-Fuller *et al.*, 2010; Smiljanic and Bradlow, 2009). Compared to conversational speech, clear speech exhibited slower speech rates, higher pitch, and larger F0 ranges. Compared to QS, NAS was characterized by slower speech rates, higher pitch, and more energy in the 1–3 kHz range. These analyses showed that clear speech and NAS shared some acoustic-articulatory features. However, the larger NAS intelligibility and recognition memory benefit suggests that NAS is characterized by different enhancements as well. This could be realized through a larger magnitude of modifications (e.g., highest F0 mean for NAS clear sentences compared to all other productions) or through additional changes not investigated here (e.g., vocal effort, voice quality, phonological contrast enhancement). Despite the wealth of production and perception studies in clear speech and NAS research, a direct relationship between the acoustic-phonetic features and intelligibility is still rather tenuous (Ferguson, 2004; Picheny *et al.*, 1989; Stollman *et al.*, 1994; Uchanski *et al.*, 1996). It remains to be seen whether these are the same features that contribute to the observed improvements in recognition memory. Establishing the exact mapping between acoustic-articulatory modifications and perceptual benefits thus remains a challenging question to researchers.

It is important to note that the clear speech in this study was produced following specific instructions and not spontaneously in response to an actual listener. Recent research has shown that different instructions as well as speech produced in the presence of actual listeners can yield different acoustic-phonetic modifications (Hazan and Baker, 2011; Lam *et al.*, 2012), whereas others have found no significant acoustic-phonetic differences between deliberately or inadvertently produced clear speech (Hazan and Markham, 2004; Bond and Moore, 1994). Replicating the experiment using recordings of different speakers with and without communicative intent will be necessary to disambiguate the extent of the influence of these methodological choices. Further research is also necessary to examine the influence of the type and level of noise in noise-adapted speech recordings, given the evidence that the characteristics of intelligibility-enhancing speech vary according to the quality of the noise masker and the intent of the speaker (Hazan and Baker, 2011; Hazan *et al.*, 2012; Uther *et al.*, 2007; Burnham *et al.*, 2002).

VI. CONCLUSIONS

While it is well established that clear speech improves intelligibility for various listener groups and under different listening conditions (Picheny *et al.*, 1985; Smiljanic and Bradlow, 2005, 2011; Bradlow and Bent, 2002; Bradlow *et al.*, 2003; Liu *et al.*, 2004), the current paper provides new evidence that the listener- and environment-oriented speaking style changes enhance recognition memory in noise. In addition to a significant effect of talker voice on word recognition and recognition memory (Goldinger, 1996; Palmeri *et al.*, 1993), within-talker variability also impacts recognition memory. It remains to be determined precisely which mechanism underlies the word recognition and recognition memory benefit for the variation related to intelligibility. Finally, the results reported here that noise-adapted speech is both more intelligible and better remembered than quiet speech mixed with noise contributes to the growing area of work illustrating the need for more naturalistic speech-in-noise perception research (i.e., incorporating noise-adapted stimuli).

The results of this study have several practical and clinical implications. The finding that recognition memory is influenced by variability in speech intelligibility suggests that simple speaking style adaptations on the part of the talker can improve listener comprehension and recognition memory in real-world, noisy situations (e.g., in the classroom). Conversely, recognition memory may be adversely affected by other sources of variability in speech intelligibility such as speech production impairments or foreign-accented speech (i.e., any speech which requires additional cognitive effort to process). The results of this study highlight that memory difficulties for older adults or cochlear implant users, for instance, may in part arise from the extra effort to perceptually process speech. Finally, we found that speech adapted to noise was more intelligible and better remembered than speech recorded in quiet and mixed with noise. This suggests that the majority of speech-in-noise results utilizing the latter may be overestimating the effects of noise on speech perception. Ultimately, it is important to conduct speech perception research that simulates real-world communicative conditions as best as possible.

ACKNOWLEDGMENTS

We would like to thank Kristin Van Engen for her assistance with experimental design. We would also like to thank Lauren Franklin and Johanna Boardman for their assistance with data collection, and Erika Hale for her assistance with statistical analyses.

[1] An increase in mean F0 has been found in clear speech as well, albeit less robustly (Pichora-Fuller *et al.*, 2010).

[2] p = 0.0511 on the set of all sentences, with new sentences marginally more intelligible than old sentences. However, we believe this to be confounded by a strong learning effect, as listeners were always first presented with 10 old sentences followed by 10 new. A test on the latter 60 sentences (excluding the first 20) yields a p value of 0.227.

Alter, A. L., and Oppenheimer, D. M. (**2008**). "Effects of fluency on psychological distance and mental construal (or why New York is a large city, but New York is a civilized jungle)," Psychol. Sci. **19**, 161–167.

Assmann, P. F., and Summerfield, Q. (**2004**). "The perception of speech under adverse acoustic conditions," in *Speech Processing in the Auditory System*, edited by S. Greenberg, W. A. Ainsworth, A. N. Popper, and R. R. Fay (Springer Verlag, New York), pp. 231–308.

Besken, M., and Mulligan, N. W. (**2013**). "Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance," Mem. Cognit. **41**, 897–903.

Bond, Z., and Moore, T. (**1994**). "A note on the acoustic-phonetic characteristics of inadvertently clear speech," Speech Commun. **14**, 325–337.

Bradlow, A. R., and Bent, T. (**2002**). "The clear speech effect for non-native listeners," J. Acoust. Soc. Am. **112**, 272–284.

Bradlow, A. R., Kraus, N., and Hayes, E. (**2003**). "Speaking clearly for children with learning disabilities: Sentence perception in noise," J. Speech Lang. Hear. Res. **46**, 80–97.

Burnham, D., Kitamura, C., and Vollmer-Conna, U. (**2002**). "What's new pussycat? On talking to babies and animals," Science **296**, 1435.

Calandruccio, L., and Smiljanic, R. (**2012**). "The development of sentence recognition materials using a basic non-native English lexicon," J. Speech Lang. Hear. Res. **55**(5), 1342–1355.

Chung, V., Mirante, N., Otten, J., and Vatikiotis-Bateson, E. (**2005**). "Audiovisual processing of Lombard speech," Proc. Auditory-Visual Speech Process. Int. Conf. **2005**, 55–56.

Cooke, M., and Lecumberri, M. L. G. (**2012**). "The intelligibility of Lombard speech for non-native listeners," J. Acoust. Soc. Am. **132**(2), 1120–1129.

Dreher, J. J., and O'Neill, J. (**1957**). "Effects of ambient noise on speaker intelligibility for words and phrases," J. Acoust. Soc. Am. **29**, 1320–1323.

Ferguson, S. H. (**2004**). "Talker differences in clear and conversational speech: Vowel intelligibility for normal- hearing listeners," J. Acoust. Soc. Am. **116**, 2365–2373.

Ferguson, S. H., and Kewley-Port, D. (**2002**). "Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **112**, 259–271.

Goldinger, S. D. (**1996**). "Words and voices: Episodic traces in spoken word identification and recognition memory," J. Exp. Psychol. Learn. **22**, 1166–1183.

Hazan, V., and Baker, R. (**2011**). "Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions," J. Acoust. Soc. Am. **130**(4), 2139–2152.

Hazan, V., Grynpas, J., and Baker, R. (**2012**). "Is clear speech tailored to counter the effect of specific adverse listening conditions?" J. Acoust. Soc. Am. **132**(5), EL371–EL377.

Hazan, V., and Markham, D. (**2004**). "Acoustic-phonetic correlates of talker intelligibility for adults and children," J. Acoust. Soc. Am. **116**, 3108–3118.

Jessen, M., Koster, O., and Gfroerer, S. (**2005**). "Influence of vocal effort on average and variability of fundamental frequency," Int. J. Speech Lang. Law **12**(2), 174–213.

Junqua, J. C. (**1993**). "The lombard reflex and it role on human listener and automatic speech recognizers," J. Acoust. Soc. Am. **93**(1), 510–524.

Junqua, J. C. (**1996**). "The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex," Speech Commun. **20**, 13–22.

Krause, J. C., and Braida, L. D. (**2002**). "Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility," J. Acoust. Soc. Am. **112**, 2165–2172.

Krause, J. C., and Braida, L. D. (**2004**). "Acoustic properties of naturally produced clear speech at normal speaking rates," J. Acoust. Soc. Am. **115**, 362–378.

Lam, J., Tjaden, K., and Wilding, G. (**2012**). "Acoustics of Clear Speech: Effect of Instruction," J. Speech Hear. Res. **55**, 1807–1821.

Lamont, A. C., Stewart-Williams, S., and Podd, J. (**2005**). "Face recognition and aging: Effects of target age and memory load," Mem. Cognit. **33**, 1017–1024.

Lane, H., and Tranel, B. (**1971**). "The Lombard sign and the role of hearing in speech," J. Speech Hear. Res. **14**, 677–709.

Liu, S., Del Rio, E., Bradlow, A. R., and Zeng, F. G. (**2004**). "Clear speech perception in acoustic and electrical hearing," J. Acoust. Soc. Am **116**, 2374–2383.

Lombard, E. (**1911**). "Le signe de l'elevation de la voix" ("The sign of the rise in the voice"), Ann. maladies oreille, larynx, nez, pharynx **37**, 101–119.

Mattys, S. L., Brooks, J., and Cooke, M. (**2009**). "Recognizing speech under a processing load: Dissociating energetic from informational factors," Cogn. Psychol. **59**(3), 203–243.

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (**2012**). "Speech recognition in adverse conditions: A review," Lang. Cognit. Processes **27**(7–8), 953–978.

McCoy, S. L., Tun, P. A., Cod, L. C., Colangelo, M., Stewart, R. A., and Wingfield, A. (**2005**). "Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech," Q. J. Exp. Psychol. **58A**, 22–33.

Mulligan, N. W., and Lozito, J. P. (**2004**). "Self-generation and memory," Psychol. Learn. Motivation **45**, 175–214.

Murphy, D. R., Craik, F. I. M., Li, K. Z. H., and Schneider, B. A. (**2000**). "Comparing the effects of aging and background noise on short-term memory performance," Psychol. Aging **15**, 323–334.

Nairne, J. S. (**1988**). "A framework for interpreting recency effects in immediate serial recall," Mem. Cognit. **16**(4), 343–352.

Palmeri, T. J., Goldinger, S. D., and Pisoni, D. B. (**1993**). "Episodic encoding of voice attributes and recognition memory for spoken words," J. Exp. Psychol. Learn. **19**, 309–328.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1985**). "Speaking clearly for the hard of hearing I. Intelligibility differences between clear and conversational speech," J. Speech Hear. Res. **28**, 96–103.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1986**). "Speaking clearly for the hard of hearing II. Acoustic characteristics of clear and conversational speech," J. Speech Hear. Res. **29**, 434–446.

Picheny, M. A., Durlach, N. I., and Braida, L. D. (**1989**). "Speaking clearly for the hard of hearing III: An attempt to determine the contribution of speaking rate to difference in intelligibility between clear and conversational speech," J. Speech Hear. Res. **32**, 600–603.

Pichora-Fuller, M. K., Goy, H., and van Lieshout, P. (**2010**). "Effect on speech intelligibility of changes in speech production influenced by instructions and communication environments," Semin. Hear. **31**, 77–94.

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (**1995**). "How young and old adults listen to and remember speech in noise," J. Acoust. Soc. Am. **97**, 593–608.

Pittman, A. L., and Wiley, T. L. (**2001**). "Recognition of speech produced in noise," J. Speech Lang. Hear. Res. **44**, 487–496.

Rabbitt, P. M. A. (**1968**). "Channel capacity, intelligibility and immediate memory," Q. J. Exp. Psychol. **20**, 241–248.

Rabbitt, P. M. A. (**1991**). "Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ," Acta Otolaryngol., Suppl. **111**, 167–176.

Rhodes, M. G., and Castel, A. D. (**2008**). "Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions," J. Exp. Psychol.: Gen. **137**, 615–625.

Rhodes, M. G., and Castel, A. D. (**2009**). "Metacognitive illusions for auditory information: Effects on monitoring and control," Psychon. Bull. Rev. **16**(3), 550–554.

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., and Abrams, H. B. (**2006**). "Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing," Appl. Psycholinguist. **27**, 465–485.

Rönnberg, J. (**2003**). "Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model," Int. J. Audiol. **42**(S1), 68–76.

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (**2008**). "Cognition counts: A working memory system for ease of language understanding (ELU)," Int. J. Audiol. **47**(S2), S99–S105.

Schneider, W., Eschman, A., and Zuccolotto, A. (**2002**). *E-Prime User's Guide* (Psychology Software Tools, Inc, Pittsburgh, PA), pp. 1–160.

Schum, D. J. (**1996**). "Intelligibility of clear and conversational speech of young and elderly talkers," J. Am. Acad. Audiol **7**, 212–218.

Smiljanic, R. (**2013**). "Can older adults enhance the intelligibility of their speech?" J. Acoust. Soc. Am. **133**(2), EL129–EL135.

Smiljanic, R., and Bradlow, A. R. (**2005**). "Production and perception of clear speech in Croatian and English," J. Acoust. Soc. Am. **118**, 1677–1688.

Smiljanic, R., and Bradlow, A. R. (**2009**). "Speaking and hearing clearly: Talker and listener factors in speaking style changes," Linguist. Lang. Compass **3**, 236–264.

Smiljanić, R., and Bradlow, A. R. (**2011**). "Bidirectional clear speech perception benefit for native and high-proficiency non-native talkers and listeners: Intelligibility and accentedness," J. Acoust. Soc. Am. **130**(6), 4020–4031.

Smiljanic, R., and Sladen, D. (**2013**). "Acoustic and semantic enhancements for children with cochlear implants," J. Speech Hear. Res **56**, 1085–1096.

Stollman, M. H. P., Kapteyn, T. S., and Sleeswijk, B. W. (**1994**). "Effect of time-scaled modification of speech on the speech recognition threshold in noise for hearing-impaired and language-impaired children," Scand. Audiol. **23**, 39–46.

Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., and Stokes, M. A. (**1988**). "Effects of noise on speech production: acoustic and perceptual analyses," J. Acoust. Soc. Am. **84**(3), 917–928.

Uchanski, R. M. (**2005**). "Clear speech," in *Handbook of Speech Perception*, edited by D. B. Pisoni and R. E. Remez (Blackwell, Cambridge, MA), pp. 207–235.

Uchanski, R. M., Choi, S., Braida, L. D., Reed, C. M., and Durlach, N. I. (**1996**). "Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate," J. Speech Hear. Res. **39**, 494–509.

Uther, M., Knoll, M. A., and Burnham, D. (**2007**). "Do you speak E-N-G-L-I-SH? Similarities and differences in speech to foreigners and infants," Speech Commun. **49**, 2–7.

Van Engen, K., and Bradlow, A. R. (**2007**). "Sentence recognition in native- and foreign-language multi-talker background noise," J. Acoust. Soc. Am. **121**, 519–526.

Van Engen, K., Chandrasekaran, B., and Smiljanic, R. (**2012**). "Effects of speech clarity on recognition memory for spoken sentences," PLoS ONE **7**(9), e43753. doi:10.1371/journal.pone.0043753.