Research Article

The Effect of Talker and Listener Depressive Symptoms on Speech Intelligibility

Hoyoung Yi,^a Rajka Smiljanic,^b and Bharath Chandrasekaran^c

Purpose: This study examined the effect of depressive symptoms on production and perception of conversational and clear speech (CS) sentences.

Method: Five talkers each with high-depressive (HD) and low-depressive (LD) symptoms read sentences in conversational and clear speaking style. Acoustic measures of speaking rate, mean fundamental frequency (F0; Hz), F0 range (Hz), and energy in the 1–3 kHz range (dB) were obtained. Thirty-two young adult participants (15 HD, 16 LD) heard these conversational and clear sentences mixed with energetic masking (speech-shaped noise) at –5 dB SPL signal-to-noise ratio. Another group of 39 young adult participants (18 HD, 19 LD) heard the same sentences mixed with informational masking (one-talker competing speech) at –12 dB SPL signal-to-noise ratio. The key word correct score was obtained.

Results: CS was characterized by a decreased speaking rate, increased F0 mean and range, and increased energy in the 1–3 kHz range. Talkers with HD symptoms produced these modifications significantly less compared to talkers with LD symptoms. When listening to speech in energetic masking (speech-shaped noise), listeners with both HD and LD symptoms benefited less from the CS produced by HD talkers. Listeners with HD symptoms performed significantly worse than listeners with LD symptoms when listening to speech in informational masking (one-talker competing speech).

Conclusions: Results provide evidence that depressive symptoms impact intelligibility and have the potential to aid in clinical decision making for individuals with depression.

Depression is a common mental condition that affects a wide variety of chronic physical and social disabilities, such as addiction, unemployment, and suicide attempts (Kessler & Bromet, 2013). The World Health Organization has estimated as many as 300 million people suffer from depression worldwide and ranked depression as the single largest contributor to global disability with high societal costs all over the world (World Health Organization, 2017). It is widely acknowledged that individuals with depression have deficits in communication (Segrin, 1998). The American Psychiatric Association's (2013) *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* describes verbal and nonverbal indicators of depressive symptoms, including diminished ability

to think and concentrate, indecisiveness, reduced vocal intensity, slowed speech, and monotone pitch. Here, we focus on verbal communication in individuals with high levels of depressive symptoms as indicated by the Center for Epidemiological Studies–Depression (CES-D) scale (Radloff, 1977). The CES-D scale is a short self-report scale designed to measure depressive symptoms for use with general and clinical populations in order to identify elevated depressive symptoms with high internal consistency (Radloff, 1977). While these individuals are not medically diagnosed as clinically depressed, they have a higher likelihood of having major depressive disorder. We are interested in assessing intelligibility variation in talkers and listeners with high-depressive (HD) symptoms with an eye on aiding depression screening for clinicians. To that end, we first examine the extent to which “talkers” with HD symptoms can produce listener-oriented, intelligibility-enhancing speaking style adaptations (clear speech [CS]). Next, we examine whether “listeners” with HD symptoms can benefit from CS enhancements when recognizing speech in challenging listening situations, namely, when speech is masked by environmental noise (speech-shaped noise [SSN]) and by competing speech (one-talker competing speech [1-T]). Identifying deficits in speech

^aDepartment of Speech-Language-Hearing Sciences, Texas Tech University Health Sciences Center, Lubbock

^bDepartment of Linguistics, The University of Texas at Austin

^cDepartment of Communication Science and Disorders and Center for the Neural Basis of Cognition, University of Pittsburgh, PA

Correspondence to Rajka Smiljanic: rajka@austin.utexas.edu

Editor-in-Chief: Nancy Solomon

Received March 11, 2019

Revision received July 9, 2019

Accepted August 12, 2019

https://doi.org/10.1044/2019_JSLHR-S-19-0112

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

production and perception mechanisms provides a better understanding of the nature of communicative deficits in individuals with depressive symptoms and has a potential of aiding identification of depressive symptoms in clinical populations.

Production

Documented speech-related symptoms in major depressive disorder include indistinct, quiet, less variable, and slow speech output. Acoustically, speech produced by individuals with depression has reduced prosodic variability evidenced by reduced pitch range (Cannizzaro, Harel, Reilly, Chappell, & Snyder, 2004; France, Shiavi, Silverman, Silverman, & Wilkes, 2000; Nilsonne, 1987), slower speech rate, long silent pauses (Balsters, Krahmer, Swerts, & Vingerhoets, 2012; Cannizzaro et al., 2004; Nilsonne, 1987), reduced speech intensity (France et al., 2000; Kuny & Stassen, 1993), and reduced distinction between vowel categories (Scherer, Morency, Gratch, & Pestian, 2015). Cannizzaro et al. (2004) showed, for individuals with major depressive symptoms, that as the Hamilton Depression Rating Scale (Hamilton, 1960) scores increased, indicating HD symptoms, speaking rate and pitch variation were significantly reduced.

Similar speech patterns are also common in dysarthric speech disorders arising from cognitive impairments with associated effects on muscle tension and control (Kent, 2000; Kent & Kim, 2003). Acoustic similarities between patients with Parkinson's disease (PD), whose speech is typically characterized by dysarthria, and individuals with major depressive disorder have been reported (Flint, Black, Campbell-Taylor, Gailey, & Levinton, 1993). Patients with PD additionally demonstrate a high occurrence of depression (Cummings, 1992; Reijnders, Ehrt, Weber, Aarsland, & Leentjens, 2008), which has a negative impact on the progress of PD, including motor symptoms and cognitive deficits. A majority of people with dysarthria after stroke also experience depressive symptoms, further affecting poststroke recovery and functional outcomes in activities of daily living (Dickson, Barbour, Brady, Clark, & Paton, 2008; Whyte & Mulsant, 2002). Individuals diagnosed with communication disorders and co-occurring depression may thus be differentially disadvantaged.

Most of the literature focuses on the populations with clinical diagnoses of depression and lacks studies that involve individuals who were not clinically diagnosed with depression but exhibit elevated depressive symptoms and are at higher risk for developing major depressive disorder. Here, we are interested in examining whether individuals with HD symptoms also show diminished acoustic variation in their speech production. Speech patterns with these characteristics (both in healthy and clinical populations) may lead to decreased intelligibility and more effortful processing on the part of the listener (Cummins et al., 2015; Hazan, Romeo, & Pettinato, 2013; Liss, Timmel, Baxley, & Killingsworth, 2005; McAuliffe, Wilding, Rickard, & O'Beirne, 2012; Newman, Clouse, & Burnham, 2001; Torre & Barlow,

2009). Reduced acoustic variability and correspondingly monotonous sounding speech in individuals with HD symptoms could thus reduce their effectiveness in daily communication.

Although previous work has revealed consistent speech characteristics in individuals with depression, to our knowledge, no studies to date have examined the extent to which speech output and consequent intelligibility can be enhanced via listener-oriented speaking style modifications. When they are aware that their communication partner is having difficulty understanding them, talkers naturally adopt a listener-oriented clear speaking style (Gilbert, Chandrasekaran, & Smiljanic, 2014; Smiljanic & Bradlow, 2005, 2009; Van Engen, Chandrasekaran, & Smiljanic, 2012). Compared with conversational speech, CS is typically characterized by slower speech rate, louder vocal level, wider fundamental frequency (F0) range, greater obstruent root-mean-square energy, an expanded vowel space, and increased energy at higher frequencies (Ferguson & Kewley-Port, 2007; Gilbert et al., 2014; Smiljanic & Bradlow, 2005; Smiljanic & Gilbert, 2017a). Perceptually, CS improves word recognition for a variety of listener groups (Pichora-Fuller, Goy, & van Lieshout, 2010; Smiljanic & Bradlow, 2009). Within the hypo- and hyperarticulation (H&H) theory (Lindblom, 1990; Perkell, Zandipour, Matthies, & Lane, 2002), conversational speech is at the hypoarticulation end of the continuum, and CS is at the hyperarticulation end.

Young healthy adults can successfully modify their output from hypo- to hyperarticulated speech in response to different communication challenges. However, non-native talkers, children, and older adults have all been shown to be less efficient in making CS adjustments both in how they modify their speech and in the resulting intelligibility benefit (Hazan, Tuomainen, & Pettinato, 2016; Pettinato, Tuomainen, Granlund, & Hazan, 2016; Rogers, DeMasi, & Krause, 2010; Smiljanic & Bradlow, 2011; Smiljanic & Gilbert, 2017a, 2017b). In clinical populations, individuals with PD, amyotrophic lateral sclerosis, or multiple sclerosis, whose production of voice pitch, loudness, speaking rate, vowel formants, and distribution of energy are disrupted, were less able to modify their spoken output when instructed to speak clearly compared to the age-matched controls (Goberman & Elmer, 2005; Tjaden, Kain, & Lam, 2014; Tjaden, Lam, & Wilding, 2013; Tjaden, Sussman, & Wilding, 2014). The current study examines whether talkers with HD symptoms produce similar hyperarticulatory modifications when instructed to speak clearly and whether depressive symptoms affect their hypo- to hyperarticulation abilities. We expect that individuals with HD symptoms will be less efficient in producing listener-oriented intelligibility-enhancing speech modifications in comparison to individuals with low-depressive (LD) symptoms.

Perception

The H&H theory (Lindblom, 1990) proposes perception–production links as being a dynamic process

leading to higher order cognitive function. To our knowledge, no studies have examined CS enhancement in speech production and intelligibility in talkers and listeners with HD symptoms. Few studies have examined speech perception in individuals with elevated depressive symptoms (e.g., Chandrasekaran, Van Engen, Xie, Beevers, & Maddox, 2015), relative to the literature on speech production. From the listener perspective, speech communication often occurs under challenging listening conditions, including noisy restaurants, hospitals, or classrooms. A listener needs to understand the target speech signal, portions of which can be masked by overlapping noise while simultaneously disregarding irrelevant competing speech signals. These requirements for intelligible message transmission make speech comprehension a challenging task even for healthy, normal-hearing listeners (Mattys, Davis, Bradlow, & Scott, 2012).

The two types of interference with the speech signal commonly encountered are referred to as energetic masking (EM) and informational masking (IM; Brungart, Simpson, Ericson, & Scott, 2001; Chandrasekaran et al., 2015; Van Engen, Phelps, Smiljanic, & Chandrasekaran, 2014; Xie, Maddox, Knopik, McGeary, & Chandrasekaran, 2015). EM occurs in the auditory periphery and interferes with the perception of a speech stimulus due to the spectrotemporal overlap of the target signal and masker. IM occurs at higher levels of auditory and cognitive processing when target signal and a masker, other competing speech, are difficult to segregate from one another even though the target signal and interfering speech (masker) are relatively well represented in the auditory system. IM thus arises after the effect of EM has been accounted for (Cooke, Garcia Lecumberri, & Barker, 2008). The possible sources of IM are postperiphery consequence of masking, including misallocation of components between the noise and the target, competing attentional capture by the masker, linguistic interference, and associated cognitive load (Cooke et al., 2008). Speech recognition in IM incurs greater attentional and higher order cognitive cost to selectively pay attention to the target words while ignoring competing maskers.

The current study examines if individuals with HD symptoms experience larger listening deficit in challenging listening conditions, under both IM and EM maskers, compared to individuals with LD symptoms. Depressive symptoms have been shown to affect higher order cognitive functions, such as executive control (Austin et al., 1992; McDermott & Ebmeier, 2009), cognitive flexibility (Butters et al., 2004), working memory (Christopher & MacDonald, 2005; Clark, Chamberlain, & Sahakian, 2009), and cognitive inhibition (Joormann, 2010), all of which are recruited more in challenging listening environments. Recent research has suggested that the inhibitory deficit from negative emotional materials is a key mechanism in increasing risk for depression (Joormann, 2010). People with depression have difficulty preventing irrelevant negative material from entering and remaining in working memory. Similar inhibitory dysfunction may affect speech perception in noise, especially under IM, which requires listeners to inhibit irrelevant competing speech to detect and attend target

speech. Chandrasekaran et al. (2015) investigated speech perception in individuals with elevated depressive symptoms (i.e., nonclinical population) under various maskers. They found that listeners with HD symptoms were affected more by IM than EM compared to individuals with LD symptoms. The IM deficit may be driven by increased distractibility related to elevated depressive symptoms. In the current study, we examine the extent to which listeners with HD symptoms can utilize conversational-to-CS enhancements to improve their word recognition in noise and the extent to which these enhancements can ameliorate a selective deficit during IM.

Experiment Overview

The current article presents the results from production and perception experiments. The aim of the production experiment was to examine the extent to which individuals with elevated depressive symptoms can produce conversational-to-CS modifications such that their intelligibility is enhanced. We examined the acoustic-articulatory modifications that characterize conversational speech and CS produced by five talkers with HD symptoms and five talkers with LD symptoms based on the CES-D scale (Radloff, 1977). Given the cognitive impairments, muscle control and tension deficits, and reductions in intensity and subband energy variance in depression (Caligiuri & Ellwanger, 2000; Cummins, Epps, & Ambikairajah, 2013; Cummins, Epps, Sethu, Breakspear, & Goecke, 2013; Quatieri & Malyska, 2012; Sobin & Sackeim, 1997), we expected that talkers with HD symptoms will exhibit smaller and/or different acoustic-articulatory CS modifications. Furthermore, it is expected that these modifications will lead to smaller intelligibility gains compared to talkers with LD symptoms. Perception experiments tested word recognition in two types of noise (SSN and 1-T) for the listener-oriented, intelligibility-enhancing CS in listeners with HD and LD symptoms. Participants listened to sentences in background noise that represented the extreme ends of the continuum of maskers from purely energetic (SSN) to primarily informational (1-T; Chandrasekaran et al., 2015; Van Engen et al., 2014). We predicted that all listeners will benefit from CS when recognizing words with both energetic and informational maskers. However, we expected that the listeners with HD symptoms will receive less benefit from conversational-to-CS enhancements, especially in IM, which may incur greater attentional and higher order cognitive cost during inhibition of the irrelevant competing speech compared to listeners with LD symptoms.

Examining how individuals with HD symptoms adapt to different communicative challenges, whether as talkers or listeners, may provide important clues about the nature of their communicative deficits. Combined, the results will provide novel insights into the critical interactivities between the two in individuals with elevated depressive symptoms. Ultimately, speech production and perception data from individuals with HD symptoms could be used as observable behavioral signals to improve

diagnosis accuracy and treatment efficacy for individuals with depressive symptoms (Cummins, Epps, & Ambikairajah, 2013; Cummins, Epps, Sethu, et al., 2013; Cummins et al., 2015).

Method

Production

Participants

Twenty female undergraduate student participants were recruited from the University of Texas at Austin. To control for gender differences in intelligibility (Bradlow, Torretta, & Pisoni, 1996; Ferguson, 2004), only female participants were included in the current study. All participants completed the CES-D scale (Radloff, 1977). The CES-D scale is one of the most widespread brief scales for identifying depressive symptoms in both general and clinical populations. The questionnaire contains 20 items about the occurrence of the depressive symptoms in the week prior to the interview with response options ranging from 0 (*not at all or less than 1 day*) to 3 (*most or all of the time, 5–7 days*). The total score range is from 0 to 60, and the higher scores reflect increased depressive symptom severity. A standard cutoff point of 16 has been found to have appropriate sensitivity and specificity rates for identifying depressed individuals (e.g., Shean & Baldwin, 2008). Individuals with a score of 16 or more have had at least six of the 20 symptoms in the CES-D scale, and they are at risk for clinical depression (Radloff, 1977). The participants in the current study were classified as having elevated depressive symptoms based on a score of 16 or greater on the CES-D scale.

Following these criteria, five talkers were identified as having HD symptoms, and 15 talkers were identified as having LD symptoms. From the 15 talkers with LD symptoms, we selected five talkers with the lowest score on the CES-D scale. These LD talkers were age-matched with talkers with HD symptoms (see Table 1). Talkers with LD symptoms were between the ages of 19 and 24 years (average age = 21 years). Talkers with HD symptoms were between the ages of 18 and 20 years (average age = 19 years). All talkers passed a hearing screening test (threshold of < 25 dB HL at 1, 2, and 4 kHz) and completed the Language Experience and Proficiency Questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007) prior to testing to verify they were monolingual English speakers with no second-language exposure from main caregivers (i.e., parents) before the age of 12 years. Participants provided written informed consent

and received extra course credit. All study procedures were approved by the institutional review board at the University of Texas at Austin.

Stimuli and Procedure

Stimuli consisted of 80 meaningful sentences (e.g., *The gray mouse ate the cheese*) from the Basic English Lexicon (Calandruccio & Smiljanic, 2012). Each sentence included four key words for intelligibility scoring. Each talker was recorded producing the full set of 80 sentences first in conversational and then in clear speaking style. Conversational speech was elicited by instructing each participant to speak normally and casually as if talking to a friend or family member who is familiar with her voice and speech pattern. For CS, participants were asked to speak as if they were communicating with someone who has a hard time understanding them because of a low proficiency in English (i.e., nonnative English speaker) or who has hearing difficulty. These instructions have reliably shown to be sufficient for elicitation of diverse speaking styles in prior studies (Smiljanic & Bradlow, 2009; Smiljanic & Gilbert, 2017a, 2017b). Sentences were presented to participants one at a time on a computer monitor. Recordings were made using a stand microphone and a MOTU UltraLite-mk3 Hybrid recorder.

Analyses

In order to assess conversational-to-CS modifications, a series of acoustic analyses was performed on conversational and clear sentences produced by five HD and five LD talkers. Four specific acoustic-articulatory features were measured: speech rate (syllables per second), mean F0 (Hz), F0 range (Hz), and energy in the 1–3 kHz range (dB). Speech rate was calculated as the number of syllables divided by the sentence duration. F0 range was calculated as the difference between the sentence’s lowest and highest F0 values. 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 sentence. This measure was chosen since energy increases in this frequency band have been shown to typically accompany CS modifications (Hazan & Baker, 2011; Hazan et al., 2016; Krause & Braid, 2004; Smiljanic & Gilbert, 2017a). This energy increase is a sign of a reduction of spectral tilt, which characterizes speech produced with increased vocal effort (e.g., Glave & Rietveld, 1975; Sluijter & van Heuven, 1996). All measurements except speech rate were automatically derived using Praat scripts (Boersma & Weenink, 2014).

Perception

Participants

Listeners participated in one of the two listening experiments: (a) word recognition in SSN or (b) word recognition in 1-T. The inclusion/exclusion criteria were the same as for the talkers. All study procedures were approved

Table 1. Demographic information for the two talker groups.

Depressive Symptoms	Age (years)	F:M	CES-D
HD (5)	19.0 (18–20)	5:0	22 (18–31)
LD (5)	21.4 (19–24)	5:0	1.4 (1–3)

Note. F = female; M = male; CES-D = Center for Epidemiological Studies–Depression; HD = high depressive; LD = low depressive.

by the institutional review board at the University of Texas at Austin.

Word recognition in SSN. Thirty-two monolingual American English listeners were recruited from University of Texas undergraduates. Listener-participants self-reported no known hearing impairment and completed the CES-D scale to classify groups with HD and LD symptoms. The listener-participants were classified as having elevated depressive symptoms (i.e., HD) if they scored 16 or greater on the CES-D scale. They were classified as having LD symptoms if they scored 15 or lower on the CES-D scale. Among the 32 participants, a participant was excluded because of incomplete participation in the inclusionary protocol. On the basis of the inclusionary criterion, 15 listeners demonstrated elevated depressive symptoms (i.e., HD), and 16 listeners were classified in the LD symptom group (see Table 2). All listener-participants were between the ages of 18 and 32 years (average age = 22 years).

Word recognition in I-T. Thirty-nine undergraduates from the University of Texas at Austin participated in the experiment. All listeners were native, monolingual speakers based on the same criteria as in the intelligibility test with SSN. They all passed a hearing screening (threshold of < 25 dB HL at 0.5, 1, 2, and 4 kHz). Two participants among the 39 talkers were excluded from the analyses because they misunderstood the instructions. Listeners completed the CES-D scale. According to the same criterion in the production experiment, there were 19 listeners with LD symptoms and 18 listeners with HD symptoms. Participants provided written informed consent, and they either were paid for their participation or received course credit. All participants were between the ages of 18 and 35 years (average age = 21 years; see Table 3).

Procedure

Word recognition in SSN. A total of 1,600 sentences (80 sentences × 2 speaking styles × 10 talkers) were equated for average root-mean-square amplitude. SSN was generated by filtering white noise to the long-term average spectrum of the full set of sentences. Each file was digitally mixed with noise at a signal-to-noise ratio (SNR) of −5 dB SPL. The SNR level was determined through piloting to avoid a ceiling effect. Each of the 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

Table 2. Listeners for the intelligibility test with speech-shaped noise.

Depressive Symptoms	Age (years)	F:M	CES-D
HD (15)	20.4 (18–28)	11:4	31.1 (16–46)
LD (16)	21.6 (17–32)	10:6	6.5 (0–13)

Note. F = female; M = male; CES-D = Center for Epidemiological Studies–Depression; HD = high depressive; LD = low depressive.

Table 3. Listeners for the intelligibility test with one-talker competing speech.

Depressive Symptoms	Age (years)	F:M	CES-D
HD (18)	20.4 (18–35)	9:9	29.6 (17–50)
LD (19)	20.5 (18–29)	9:10	8.9 (3–14)

Note. F = female; M = male; CES-D = Center for Epidemiological Studies–Depression; HD = high depressive; LD = low depressive.

noise mixed with the speech. Each participant was seated in front of a computer monitor. The stimuli were played over headphones at a comfortable listening level set by the experimenter. Instructions and stimuli were presented using PsychoPy (Peirce, 2007). The participants were instructed to listen to each sentence and type what they heard onto a computer. Each trial was presented only once, but participants could take as much time as they wished to write down the sentences. Each listener heard a total of 85 trials, including five practice items. Eighty sentences were randomly selected out of a total pool of 1,600 sentences and presented without repetition for each listener. Each listener heard 20 clear and 20 conversational sentences produced by five talkers with LD symptoms and 20 clear and 20 conversational sentences produced by five talkers with HD symptoms. Thus, each participant listened to the sentences from all 10 talkers, but they each heard a different set of 80 sentences. Responses were scored by the number of key words correctly identified out of 320 words. Key words with added or omitted morphemes were scored as incorrect.

Word recognition in I-T. The same set of 1,600 target sentences (80 sentences × 2 speaking styles × 10 talkers) as above were mixed with one-talker masker (Chandrasekaran et al., 2015). A female talker of American English produced 30 simple English sentences used for one-talker masker (Chandrasekaran et al., 2015; Van Engen et al., 2010). To prevent masker familiarization, each target sentence was digitally mixed with one of the randomly selected portions of the masker at −12 dB SPL SNR. The SNR level was determined through piloting to avoid a ceiling effect. 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 and tasks were the same as in the word recognition in the SSN experiment.

Results

Production

Average speaking rate, average F0, F0 range, and 1–3 kHz energy for the two talker groups and speaking styles are provided in Table 4. The differences between CS and conversational speech styles based on these four acoustic measures are also reported in Table 4.

Table 4. Acoustic measures of sentences from talkers with low-depressive (LD) and high-depressive (HD) symptoms in clear and conversational (Conv) styles and the difference between conversational and clear speech styles.

	LD talkers			HD talkers		
	Conv, <i>M</i> (<i>SD</i>)	Clear, <i>M</i> (<i>SD</i>)	Conv–Clear	Conv, <i>M</i> (<i>SD</i>)	Clear, <i>M</i> (<i>SD</i>)	Conv–Clear
Speech rate (syllables/s)	4.88 (0.82)	2.81 (0.69)	2.07	4.69 (0.75)	3.08 (0.56)	1.61
Average F0 (Hz)	193.52 (15.43)	199.53 (18.17)	–6.01	193.28 (24.88)	192.94 (23.28)	0.35
F0 range (Hz)	156.89 (57.78)	210.86 (72.96)	–53.98	166.63 (62.90)	189.05 (70.85)	–22.41
Energy: 1–3 kHz (dB)	15.53 (6.42)	18.02 (6.70)	–2.49	16.01 (5.51)	16.60 (6.45)	–0.60

Note. Conv = conversational speech; Clear = clear speech; Conv–Clear = mean difference between conversational and clear speech styles; F0 = fundamental frequency.

For each of the four acoustic measurements, results were submitted to linear mixed-effects regression models with lme4 package (v1.1-14; Bates, Mächler, Bolker, & Walker, 2015) in R (v3.4.2). Fixed effects included talker group (HD group, LD group), speaking style (conversational, clear), and an interaction between talker group and speaking style for each acoustic measure. Talkers and sentences were included as random intercepts. The results of regression are presented in Table 5.

The Wald test was applied to examine the overall effect of talker group, speaking style, and their interaction on the acoustic measurements. Results revealed that, for all the acoustic measures, the main effect of listener-oriented speaking style was significant: speech rate: $\chi^2(1) = 4,462.671$, $p < .001$; mean F0: $\chi^2(1) = 32.403$, $p < .001$; range F0: $\chi^2(1) = 156.113$, $p < .001$; and energy in the 1–3 kHz range: $\chi^2(1) = 239.403$, $p < .001$. CS sentences were slower and had higher mean F0, greater F0 range, and more energy in the 1–3 kHz range than conversational sentences. There were no significant differences between talkers with HD and LD symptoms for all the acoustic characteristics: speech rate: $\chi^2(1) = 0.297$, $p = .586$; mean F0: $\chi^2(1) = 0.001$, $p = .981$;

range F0: $\chi^2(1) = 0.373$, $p = .541$; and energy in the 1–3 kHz range: $\chi^2(1) = 0.014$, $p = .906$. There were significant two-way interactions between talker group and speaking style for all the acoustic measures (all p values were smaller than .001). The interaction was further investigated using pairwise contrasts that were evaluated with Bonferroni-adjusted significance in R. For speech rate, there were significant effects on listener-oriented speaking style for both LD talkers, $B = 2.149$, $t(1509) = 66.803$, $p < .001$, and HD talkers, $B = 1.624$, $t(1509) = 50.407$, $p < .001$. Even though both talker groups slowed down significantly in CS sentences, the effect was greater for the talkers with LD symptoms than for those with HD symptoms. Conversational-to-CS speaking rate decrease was –2.07 syllables/s for talkers with LD symptoms and –1.61 syllables/s for talkers with HD symptoms. Only talkers with LD symptoms significantly increased mean F0 on CS sentences compared with conversational sentences: LD talkers: $B = -6.013$, $t(1509) = -5.692$, $p < .001$; HD talkers: $B = 0.347$, $t(1509) = 0.328$, $p = 1$. F0 range was significantly increased in CS compared to conversational speech for both talker groups; however, the effect of speaking style was greater for LD talkers, $B = -53.975$,

Table 5. Results of the linear mixed-effects regressions on acoustic measurements.

	Fixed effect	Estimate	SE	<i>t</i> (<i>df</i>)	<i>p</i>
Speech rate	(Intercept)	2.814	0.176	16.020 (9.8)	< .001
	Talker_HD	0.396	0.237	1.670 (8.1)	.133
	Style_Conversational	2.149	0.032	66.800 (1509)	< .001
	Talker_HD: Conversational	–0.525	0.045	–11.550 (1509)	< .001
Average F0	(Intercept)	199.531	7.052	28.296 (8.1)	< .001
	Talker_HD	–6.593	9.955	–0.662 (8.1)	.526
	Style_Conversational	–6.013	1.056	–5.692 (1509)	< .001
	Talker_HD: Conversational	6.360	1.494	4.257 (1509)	< .001
F0 range	(Intercept)	210.863	11.400	18.497 (9)	< .001
	Talker_HD	–21.817	15.952	–1.368 (8.6)	.206
	Style_Conversational	–53.975	4.320	–12.495 (1509.5)	< .001
	Talker_HD: Conversational	31.564	6.109	5.167 (1509.5)	< .001
Energy: 1–3 kHz	(Intercept)	18.018	2.855	6.311 (8.1)	< .001
	Talker_HD	–1.415	4.032	–0.351 (8)	.735
	Style_Conversational	–2.488	0.161	–15.473 (1509)	< .001
	Talker_HD: Conversational	1.891	0.227	8.318 (1509)	< .001

Note. Each intercept represents a reference condition: The talker group was the low-depressive group, and the sentence style was clear speech. HD = high depressive; F0 = fundamental frequency.

$t(1509) = -12.495, p < .001$, than for HD talkers, $B = -22.411, t(1509) = -3.709, p = .001$. Conversational-to-CS F0 range increase was 53.98 Hz for LD talkers and 22.41 Hz for HD talkers. Finally, the energy in the 1–3 kHz range was greater in CS for both LD and HD talkers, but the effect of speaking style was greater for LD talkers, $B = -2.488, t(1509) = -15.473, p < .001$, than for HD talkers, $B = -0.597, t(1509) = -3.709, p = .001$. Conversational-to-CS energy in the 1–3 kHz range change was –2.49 dB for LD talkers and –0.60 dB for HD talkers.

Perception

Word recognition data were analyzed with a linear mixed-effects logistic regression using the lme4 package (v1.1-14; Bates et al., 2015) in R (v3.4.2). Key word identification (i.e., correct or incorrect) was the dichotomous dependent variable. Fixed effects included listener group, talker condition, speech style, interaction of speech style with talker condition, and interaction of speech style with listener condition. Participants (listeners), sentences, and talkers were included in the model as random intercepts.

Word Recognition in SSN

The percentage of the key word correct score for HD and LD listeners for conversational and clear sentences produced by HD and LD talkers is shown in Figure 1. For all listeners, CS was more intelligible than conversational speech. Both LD and HD listeners identified sentences produced by talkers with LD symptoms more accurately than sentences produced by talkers with HD symptoms, especially for the CS sentences.

The results of regression are presented in Table 6. Overall, word recognition was not significantly affected by listener group ($p = .319$) or talker group ($p = .219$). The effect of speaking style was significant ($p < .001$), with the probability of correct key word identification significantly enhanced for CS compared to conversational speech. Results also revealed a significant interaction between speaking style and talker group ($p = .018$). The interaction effect was further analyzed using pairwise contrasts with Bonferroni-adjusted significance. Results revealed that, for the sentences produced by both LD and HD talkers, word recognition of the CS sentences was significantly more accurate than that of the conversational speech sentences. However, the effect of listener-oriented speaking style was greater for the sentences produced by talkers with LD symptoms ($B = 1.240, z = 18.966, p < .001$) than for the sentences produced by talkers with HD symptoms ($B = 1.022, z = 15.451, p < .001$). There was no significant interaction between speaking style and listener group ($p = .678$).

Word Recognition in 1-T

The percentage of the key word correct score for HD and LD listeners for conversational and clear sentences produced by HD and LD talkers is shown in Figure 2. For all listeners, word recognition in CS was more accurate than that in conversational speech when presented in 1-T. HD listeners were least accurate in recognizing words produced by HD talkers in conversational speech.

The results of regression are presented in Table 7. Results revealed that probability of correct key word identification was significantly higher for clear sentences than for conversational speech ($p < .001$). In addition, listener group differences were significant. Probability of correct identification was lower for HD listeners than for LD

Figure 1. Percentage of key words identified for listeners with low-depressive (LD) and high-depressive (HD) symptoms from conversational and clear sentences produced by talkers with LD and HD symptoms. The center line on each box plot indicates the median score, the edges of the box denote the 25th and 75th percentiles, and the whiskers extend to data points that lie within 1.5 times the interquartile range. Dots outside the range are outliers. SSN = speech-shaped noise.

Table 6. Results of the linear mixed-effects logistic regression on intelligibility data with speech-shaped noise.

Fixed effect	Estimate	SE	z	p
(Intercept)	0.830	0.313	6.349	< .008
Listener_HD	−0.151	0.152	−0.955	.319
Talker_HD	−0.497	0.404	−7.601	.219
Style_Conversational	−1.259	0.079	−18.646	< .001
Listener_HD: Conversational	0.038	0.092	0.415	.678
Talker_HD: Conversational	0.218	0.092	2.365	.018

Note. The intercept represents a reference condition: The listener group was the low-depressive group, the talker group was the low-depressive group, and the sentence style was clear speech. HD = high depressive.

listeners. The effect of talker group was not significant on word identification. The two-way interactions between speaking style and listener group ($p = .227$) and speaking style and talker group ($p = .895$) were not significant.

Discussion

The goal of this study was to examine speech production and perception in challenging listening environments in individuals with elevated depressive symptoms. In terms of the production, acoustic analyses revealed that all talkers decreased speaking rate, increased F0 mean and range, and increased energy in the 1–3 kHz range when instructed to speak clearly. The analyses confirmed that conversational and CS sentences differed in their acoustic–articulatory characteristics along dimensions that are typically found in listener-oriented speaking style adaptations (Smiljanic & Gilbert, 2017a; Van Engen et al., 2012). These enhancements may serve a goal of augmenting global salience and audibility of the speech signal as well as decrease cognitive

effort on the part of the listener via a reduction in rate at which information is transmitted or through better instantiation of phonetic categories (Cooke, King, Garnier, & Aubanel, 2014; Lansford, Liss, Caviness, & Utianski, 2011). Importantly though, individuals with HD symptoms produced all of these modifications significantly less frequently compared to LD talkers. These results give evidence of reductions in global spread of acoustic variance with increasing levels of talker depressive symptoms. These quantitative differences match the reports of less distinct phonetic events in speech of individuals diagnosed with depression (Cummins et al., 2015; Scherer et al., 2015; Trevino, Quatieri, & Malyska, 2011). Similar reductions in vocal intensity, vowel space area, and speaking rate were observed in dysarthric speech of individuals with PD, amyotrophic lateral sclerosis, or multiple sclerosis, who also have a high occurrence of depression, compared to healthy older adults when instructed to speak clearly or in response to noise (Adams, Winnell, & Jog, 2010; Goberman & Elmer, 2005; Lam & Tjaden, 2016; Tjaden et al., 2013).

Figure 2. Percentage of key words identified for listeners with low-depressive (LD) and high-depressive (HD) symptoms from conversational and clear sentences produced by talkers with LD and HD symptoms. The center line on each box plot indicates the median score, the edges of the box denote the 25th and 75th percentiles, and the whiskers extend to data points that lie within 1.5 times the interquartile range. Dots outside the range are outliers.

Table 7. Results of the linear mixed-effects logistic regression on intelligibility data with one-talker babble.

Fixed effect	Estimate	SE	z	p
(Intercept)	2.804	0.323	8.678	< .001
Listener_HD	−1.118	0.368	−3.039	.002
Talker_HD	−0.278	0.249	−1.119	.263
Style_Conversational	−0.802	0.107	−7.524	< .001
Listener_HD: Conversational	−0.139	0.115	−1.209	.227
Talker_HD: Conversational	0.015	0.111	0.132	.895

Note. The intercept represents a reference condition: The listener group was the low-depressive group, the talker group was the low-depressive group, and the sentence style was clear speech. HD = high depressive.

Intelligibility results confirmed that, at least in the SSN condition, listeners benefited less from the CS produced by HD talkers, indicating that these conversational-to-CS modifications were less efficient in increasing word recognition in noise. The lack of such talker-related effect on intelligibility in 1-T could be attributed to the overall high performance in this condition (i.e., the ceiling effect may obscure the talker-related differences). The differences in CS articulatory–acoustic adjustments in individuals with HD and LD symptoms may point to the cognitive changes and changes in the rate and precision of articulatory control and phonatory function (Caligiuri & Ellwanger, 2000; Quatieri & Malyska, 2012; Sobin & Sackeim, 1997). However, the firm link between the communicative skills and depressive symptoms remains to be elucidated further for these subclinical and clinical populations. The results further suggest that individuals with HD symptoms may not be able to change their spoken output along the hypo- to hyperarticulated speech (H&H theory; Lindblom, 1990) or fine-tune their response to different communication barriers to the same extent as individuals with LD symptoms (Hazan & Baker, 2011; Lam, Tjaden, & Wilding, 2012). Future work should investigate the role that depressive symptoms and depression play in the ability to produce intelligibility-enhancing adaptations to the listener and other communication barriers such as in response to environmental noise. Individuals with a wider range of depressive symptoms should be included in this examination to more precisely determine how these differences occur and how they determine talker intelligibility. From a clinical perspective, speech-oriented behavioral therapy techniques using rate reduction increased vocal intensity, and CS (Darling & Huber, 2011; Duffy, 2013; Lam & Tjaden, 2016) may be beneficial intervention techniques for talkers with depression whose speech intelligibility may be compromised (for similar speech output enhancement therapy with talkers with dysarthria, see Beukelman, Fager, Ullman, Hanson, & Logemann, 2002; Park, Theodoros, Finch, & Cardell, 2016; Tjaden, Sussman, et al., 2014). Due to frequent comorbidity of depression in individuals with PD and dysarthria, future studies are needed to identify depression in these clinical populations to examine CS benefits on speech intelligibility.

In terms of perception, the results showed no differences between listeners with LD and HD symptoms when

listening to speech in EM. In contrast, when listening to speech in IM, listeners with HD symptoms performed significantly worse than listeners with LD symptoms.

In the condition of IM (1-T condition), listeners with HD symptoms showed a high variability of accuracy in responses. On further analyses, this variability was driven by three listeners with HD symptoms who achieved the lowest word recognition accuracy. Among these three listeners, two with HD symptoms appeared to have difficulty paying attention to the target sentences exhibited by their reports of multiple words from the competing speech. This finding is consistent with that of Chandrasekaran et al. (2015), who also showed a selective perceptual deficit during information-masking conditions for HD listeners. Elevated depressive symptoms affect listeners' ability to selectively focus on the target speech in conditions when competing speech is present. This processing deficit in IM could arise from a number of sources, including the misattribution of noise components to the target (and vice versa), competing attention from the masker, increased cognitive load, and linguistic interference (Cooke et al., 2008; Shinn-Cunningham, 2008). Interfering speech thus requires greater use of executive functions such as inhibitory control, working memory, and cognitive flexibility, all of which have been shown to be affected by depression (Clark et al., 2009; Joormann, 2010; Joormann & Gotlib, 2008).

A novel finding of this study is that listeners with HD symptoms significantly benefited from CS adaptations. This is in line with previous work demonstrating improved intelligibility of CS over conversational speech for a variety of listener groups, including adults with normal or impaired hearing (Ferguson, 2004; Ferguson & Kewley-Port, 2002; Maniwa, Jongman, & Wade, 2008; Uchanski, Choi, Braid, Reed, & Durlach, 1996), elderly adults (Helfer, 1998; Schum, 1996), native and nonnative listeners (Bradlow & Alexander, 2007; Bradlow & Bent, 2002; Smiljanić & Bradlow, 2011), and children with and without learning impairments (Bradlow, Kraus, & Hayes, 2003), and degraded listening conditions (Van Engen et al., 2014). Although the magnitude of the CS intelligibility advantage varied across listener groups and listening conditions, these findings showed that the CS can be used to improve performance for various listener groups with different communicative challenges.

The results also revealed that, even though CS word recognition was enhanced in both EM and IM, CS benefit was smaller for HD listeners compared to LD listeners in IM. This is despite the fact that the overall performance for both listener groups was higher in 1-T than in SSN. In conditions when target speech is masked by competing speech, the enhanced CS acoustic–phonetic cues could aid listeners in focusing on the target speech and/or in taking advantage of “glimpses”—spectrotemporal regions in which the target signal is least affected by the background noise (Cooke, 2005). The current results showed that LD listeners were more efficient at stream segregation and in using such spectrotemporal “dips” in the masker to guide their word recognition compared to HD listeners. CS, while overall improving their performance, provided less of a masking release to the HD listeners. It remains to be determined precisely at what level of signal processing this listener–difference occurs.

Our findings have important practical implications since typical social settings usually involve communicating with listeners who have perceptual difficulties and in the presence of competing speech. Comprehension difficulties and increased perceptual effort in such listening situations could exacerbate social and communicative difficulties in individuals with elevated depressive symptoms. The findings further highlight difficulties beyond perceptual problems that individuals with depressive symptoms encounter when communicating in adverse conditions. Their own speech may not be understood well in environments such as noisy classrooms and offices, which may lead to breakdown in communication. The challenge remains to identify a precise mechanism by which CS enhances communication for individuals with depressive symptoms.

In summary, this study has demonstrated that individuals with HD symptoms produced reduced acoustic–articulatory modifications when instructed to speak clearly and exhibited more perceptual challenges when listening to speech masked by competing speech compared to individuals with LD symptoms. There are a couple of limitations of the current study. It remains to be determined whether talkers and listeners with clinically diagnosed depression would perform similar to the individuals tested here who have HD symptoms but may nonetheless not meet the criteria for a major depressive episode. We did not identify whether the participants who reported elevated depressive symptoms were taking antidepressant medications to treat their depressive symptoms. Most evidence examining the link between selective serotonin reuptake inhibitors, for instance, and cognitive functioning such as memory provides conflicting results (e.g., Sayyah, Eslami, AlaiShehni, & Kouti, 2016). The impact of antidepressants and depressed maternal mood has been examined on speech perception in infants (Weikum, Oberlander, Hensch, & Werker, 2012). Those results showed that prenatal depressed maternal mood and selective serotonin reuptake inhibitor exposure were found to shift developmental milestones bidirectionally in infant speech perception tasks. Future studies with individuals with clinically diagnosed depression should identify

and control antidepressant medication effects to examine the impact of elevated depressive symptoms on speech perception and production more precisely.

Furthermore, future studies are needed to identify potential mechanisms that lead to these selective production and perception deficits in individuals with depressive symptoms. Though our findings provide important insight into the effect of elevated depressive symptoms on speech production and perception, they do not allow us to pinpoint the locus of this difficulty or how precisely CS contributes to the enhanced speech perception in IM and EM. Despite these limitations, the current results from this subclinical population of HD individuals complement production and perception difficulties reported for individuals diagnosed with depression and could thus constitute important markers of future depression. While not meant for formal diagnosis, this work highlights observable behavioral signals that could augment existing clinical protocols.

Acknowledgments

Portions of this work were presented at the 171st Meeting of the Acoustical Society of America in Salt Lake City, UT, May 2016. This work was supported by National Institute on Deafness and Other Communication Disorders Grant R01DC015504 (awarded to Bharath Chandrasekaran).

References

- American Psychiatric Association.** (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Retrieved from <https://doi.org/10.1176/appi.books.9780890425596>
- Adams, S. G., Winnell, J., & Jog, M.** (2010). Effects of interlocutor distance, multi-talker background noise, and a concurrent manual task on speech intensity in Parkinson's disease. *Journal of Medical Speech-Language Pathology*, 18(4), 1–8.
- Austin, M. P., Ross, M., Murray, C., O'Carroll, R. E., Ebmeier, K. P., & Goodwin, G. M.** (1992). Cognitive function in major depression. *Journal of Affective Disorders*, 25(1), 21–29.
- Balsters, M. J. H., Krahmer, E. J., Swerts, M. G. J., & Vingerhoets, A. J. J. M.** (2012). Verbal and nonverbal correlates for depression: A review. *Current Psychiatry Reviews*, 8(3), 227–234.
- Bates, D., Mächler, M., Bolker, B., & Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Beukelman, D. R., Fager, S., Ullman, C., Hanson, E., & Logemann, J.** (2002). The impact of speech supplementation and clear speech on the intelligibility and speaking rate of people with traumatic brain injury. *Journal of Medical Speech-Language Pathology*, 10(4), 237–242.
- Boersma, P., & Weenink, D.** (2014). *Praat speech processing software*. The Netherlands: Institute of Phonetics Sciences, University of Amsterdam. Retrieved from <http://www.praat.org>
- Bradlow, A. R., & Alexander, J. A.** (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America*, 121(4), 2339–2349.
- Bradlow, A. R., & Bent, T.** (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America*, 112(1), 272–284.

- Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research*, 46(1), 80–97.
- Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, 20, 255–272.
- 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(5, Pt. 1), 2527–2538.
- Butters, M. A., Whyte, E. M., Nebes, R. D., Begley, A. E., Dew, M. A., Mulsant, B. H., ... Becker, J. T. (2004). The nature and determinants of neuropsychological functioning in late-life depression. *Archives of General Psychiatry*, 61, 587–595.
- Calandruccio, L., & Smiljanic, R. (2012). New sentence recognition materials developed using a basic non-native English lexicon. *Journal of Speech, Language, and Hearing Research*, 55(5), 1342–1355.
- Caligiuri, M. P., & Ellwanger, J. (2000). Motor and cognitive aspects of motor retardation in depression. *Journal of Affective Disorders*, 57(1–3), 83–93.
- Cannizzaro, M., Harel, B., Reilly, N., Chappell, P., & Snyder, P. J. (2004). Voice acoustical measurement of the severity of major depression. *Brain and Cognition*, 56(1), 30–35.
- Chandrasekaran, B., Van Engen, K., Xie, Z., Beevers, C. G., & Maddox, W. T. (2015). Influence of depressive symptoms on speech perception in adverse listening conditions. *Cognition and Emotion*, 29(5), 900–909.
- Christopher, G., & MacDonald, J. (2005). The impact of clinical depression on working memory. *Cognitive Neuropsychiatry*, 10(5), 379–399.
- Clark, L., Chamberlain, S. R., & Sahakian, B. J. (2009). Neurocognitive mechanisms in depression: Implications for treatment. *Annual Review of Neuroscience*, 32, 57–74.
- Cooke, M. (2005). Making sense of everyday speech: A glimpsing account. In P. Divenyi (Ed.), *Speech separation by humans and machines* (pp. 305–314). Boston, MA: Springer.
- Cooke, M., Garcia Lecumberri, M. L., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America*, 123(1), 414–427.
- Cooke, M., King, S., Garnier, M., & Aubanel, V. (2014). The listening talker: A review of human and algorithmic context-induced modifications of speech. *Computer Speech & Language*, 28(2), 543–571.
- Cummings, J. L. (1992). Depression and Parkinson's disease: A review. *The American Journal of Psychiatry*, 149(4), 443–454.
- Cummins, N., Epps, J., & Ambikairajah, E. (2013, May). Spectrotemporal analysis of speech affected by depression and psychomotor retardation. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (pp. 7542–7546). Piscataway, NJ: Institute of Electrical and Electronics Engineers.
- Cummins, N., Epps, J., Sethu, V., Breakspear, M., & Goecke, R. (2013, August). *Modeling spectral variability for the classification of depressed speech*. In *Interspeech*, pp. 857–861.
- Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, 71, 10–49.
- Darling, M., & Huber, J. E. (2011). Changes to articulatory kinematics in response to loudness cues in individuals with Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, 54(5), 1247–1259.
- Dickson, S., Barbour, R. S., Brady, M., Clark, A. M., & Paton, G. (2008). Patients' experiences of disruptions associated with post-stroke dysarthria. *International Journal of Language & Communication Disorders*, 43(2), 135–153.
- Duffy, J. R. (2013). *Motor speech disorders—E-book: Substrates, differential diagnosis, and management*. St. Louis, MO: Elsevier Mosby.
- Ferguson, S. H. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. *The Journal of the Acoustical Society of America*, 116(4, Pt. 1), 2365–2373.
- Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 112(1), 259–271.
- Ferguson, S. H., & Kewley-Port, D. (2007). Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of Speech, Language, and Hearing Research*, 50, 1241–1255.
- Flint, A. J., Black, S. E., Campbell-Taylor, I., Gailey, G. F., & Levinton, C. (1993). Abnormal speech articulation, psychomotor retardation, and subcortical dysfunction in major depression. *Journal of Psychiatric Research*, 27(3), 309–319.
- France, D. J., Shiavi, R. G., Silverman, S., Silverman, M., & Wilkes, M. (2000). Acoustical properties of speech as indicators of depression and suicidal risk. *IEEE Transactions on Biomedical Engineering*, 47(7), 829–837.
- Gilbert, R. C., Chandrasekaran, B., & Smiljanic, R. (2014). Recognition memory in noise for speech of varying intelligibility. *The Journal of the Acoustical Society of America*, 135(1), 389–399.
- Glave, R. D., & Rietveld, A. C. M. (1975). Is the effort dependence of speech loudness explicable on the basis of acoustical cues? *The Journal of the Acoustical Society of America*, 58(4), 875–879.
- Goberman, A. M., & Elmer, L. W. (2005). Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease. *Journal of Communication Disorders*, 38(3), 215–230.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, & Psychiatry*, 23(1), 56.
- Hazan, V., & Baker, R. (2011). Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions. *The Journal of the Acoustical Society of America*, 130(4), 2139–2152.
- Hazan, V., Romeo, R., & Pettinato, M. (2013, June). The impact of variation in phoneme category structure on consonant intelligibility. *Proceedings of Meetings on Acoustics ICA2013*, 19(1), 060103.
- Hazan, V., Tuomainen, O., & Pettinato, M. (2016). Suprasegmental characteristics of spontaneous speech produced in good and challenging communicative conditions by talkers aged 9–14 years. *Journal of Speech, Language, and Hearing Research*, 59(6), S1596–S1607.
- Helfer, K. S. (1998). Auditory and auditory-visual recognition of clear and conversational speech by older adults. *Journal of the American Academy of Audiology*, 9(3), 234–242.
- Joormann, J. (2010). Cognitive inhibition and emotion regulation in depression. *Current Directions in Psychological Science*, 19(3), 161–166.
- Joormann, J., & Gotlib, I. H. (2008). Updating the contents of working memory in depression: Interference from irrelevant negative material. *Journal of Abnormal Psychology*, 117(1), 182–192.

- Kent, R. D.** (2000). Research on speech motor control and its disorders: A review and prospective. *Journal of Communication Disorders*, 33(5), 391–428.
- Kent, R. D., & Kim, Y. J.** (2003). Toward an acoustic typology of motor speech disorders. *Clinical Linguistics & Phonetics*, 17(6), 427–445.
- Kessler, R. C., & Bromet, E. J.** (2013). The epidemiology of depression across cultures. *Annual Review of Public Health*, 34, 119–138.
- Krause, J. C., & Braid, L. D.** (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, 115(1), 362–378.
- Kun, S., & Stassen, H. H.** (1993). Speaking behavior and voice sound characteristics in depressive patients during recovery. *Journal of Psychiatric Research*, 27(3), 289–307.
- Lam, J., & Tjaden, K.** (2016). Clear speech variants: An acoustic study in Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, 59(4), 631–646.
- Lam, J., Tjaden, K., & Wilding, G.** (2012). Acoustics of clear speech: Effect of instruction. *Journal of Speech, Language, and Hearing Research*, 55(6), 1807–1821.
- Lansford, K. L., Liss, J. M., Caviness, J. N., & Utianski, R. L.** (2011). A cognitive-perceptual approach to conceptualizing speech intelligibility deficits and remediation practice in hypokinetic dysarthria. *Parkinson's Disease*, 2011, 150962.
- Lindblom, B.** (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht, the Netherlands: Springer.
- Liss, M., Timmel, L., Baxley, K., & Killingsworth, P.** (2005). Sensory processing sensitivity and its relation to parental bonding, anxiety, and depression. *Personality and Individual Differences*, 39(8), 1429–1439.
- Maniwa, K., Jongman, A., & Wade, T.** (2008). Perception of clear fricatives by normal-hearing and simulated hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 123(2), 1114–1125.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M.** (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940–967.
- Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K.** (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, 27(7–8), 953–978.
- McAuliffe, M. J., Wilding, P. J., Rickard, N. A., & O'Beirne, G. A.** (2012). Effect of speaker age on speech recognition and perceived listening effort in older adults with hearing loss. *Journal of Speech, Language, and Hearing Research*, 55(3), 838–847.
- McDermott, L. M., & Ebmeier, K. P.** (2009). A meta-analysis of depression severity and cognitive function. *Journal of Affective Disorders*, 119(1–3), 1–8.
- Newman, R. S., Clouse, S. A., & Burnham, J. L.** (2001). The perceptual consequences of within-talker variability in fricative production. *The Journal of the Acoustical Society of America*, 109(3), 1181–1196.
- Nilsson, Å.** (1987). Acoustic analysis of speech variables during depression and after improvement. *Acta Psychiatrica Scandinavica*, 76(3), 235–245.
- Park, S., Theodoros, D., Finch, E., & Cardell, E.** (2016). Be clear: A new intensive speech treatment for adults with nonprogressive dysarthria. *American Journal of Speech-Language Pathology*, 25(1), 97–110.
- Peirce, J. W.** (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13.
- Perkell, J. S., Zandipour, M., Matthies, M. L., & Lane, H.** (2002). Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues. *The Journal of the Acoustical Society of America*, 112(4), 1627–1641.
- Pettinato, M., Tuomainen, O., Granlund, S., & Hazan, V.** (2016). Vowel space area in later childhood and adolescence: Effects of age, sex and ease of communication. *Journal of Phonetics*, 54, 1–14.
- Pichora-Fuller, M. K., Goy, H., & van Lieshout, P.** (2010, May). Effect on speech intelligibility of changes in speech production influenced by instructions and communication environments. *Seminars in Hearing*, 31(2), 77–94.
- Quatieri, T. F., & Malyska, N.** (2012). Vocal-source biomarkers for depression: A link to psychomotor activity. In *Thirteenth Annual Conference of the International Speech Communication Association*.
- Radloff, L. S.** (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385–401.
- Reijnders, J. S., Ehrt, U., Weber, W. E., Aarsland, D., & Leentjens, A. F.** (2008). A systematic review of prevalence studies of depression in Parkinson's disease. *Movement Disorders*, 23(2), 183–189.
- Rogers, C. L., DeMasi, T. M., & Krause, J. C.** (2010). Conversational and clear speech intelligibility of /bVd/ syllables produced by native and non-native English speakers. *The Journal of the Acoustical Society of America*, 128(1), 410–423.
- Sayyah, M., Eslami, K., AlaiShehni, S., & Kouti, L.** (2016). Cognitive function before and during treatment with selective serotonin reuptake inhibitors in patients with depression or obsessive-compulsive disorder. *Psychiatry Journal*, 2016, 5480391.
- Scherer, S., Morency, L. P., Gratch, J., & Pesticelli, J.** (2015, April). Reduced vowel space is a robust indicator of psychological distress: A cross-corpus analysis. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. (pp. 4789–4793). Piscataway, NJ: Institute of Electrical and Electronics Engineers.
- Schum, D. J.** (1996). Intelligibility of clear and conversational speech of young and elderly talkers. *Journal of the American Academy of Audiology*, 7(3), 212–218.
- Segrin, C.** (1998). Interpersonal communication problems associated with depression and loneliness. In P. A. Anderson & L. A. Guerrero (Eds.), *Handbook of communication and emotion* (pp. 215–242). New York, NY: Academic Press.
- Shean, G., & Baldwin, G.** (2008). Sensitivity and specificity of depression questionnaires in a college-age sample. *The Journal of Genetic Psychology*, 169(3), 281–288.
- Shinn-Cunningham, B. G.** (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, 12(5), 182–186.
- Sluijter, A. M., & van Heuven, V. J.** (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*, 100(4, Pt. 1), 2471–2485.
- Smiljanić, R., & Bradlow, A. R.** (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, 118(3, Pt. 1), 1677–1688.
- Smiljanić, R., & Bradlow, A. R.** (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and Linguistics Compass*, 3(1), 236–264.
- Smiljanić, R., & Bradlow, A. R.** (2011). Bidirectional clear speech perception benefit for native and high-proficiency non-native

- talkers and listeners: Intelligibility and accentedness. *The Journal of the Acoustical Society of America*, 130(6), 4020–4031.
- Smiljanic, R., & Gilbert, R. C.** (2017a). Acoustics of clear and noise-adapted speech in children, young, and older adults. *Journal of Speech, Language, and Hearing Research*, 60(11), 3081–3096.
- Smiljanic, R., & Gilbert, R. C.** (2017b). Intelligibility of noise-adapted and clear speech in child, young adult, and older adult talkers. *Journal of Speech, Language, and Hearing Research*, 60(11), 3069–3080.
- Sobin, C., & Sackeim, H. A.** (1997). Psychomotor symptoms of depression. *The American Journal of Psychiatry*, 154(1), 4–17.
- Tjaden, K., Kain, A., & Lam, J.** (2014). Hybridizing conversational and clear speech to investigate the source of increased intelligibility in speakers with Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, 57(4), 1191–1205.
- Tjaden, K., Lam, J., & Wilding, G.** (2013). Vowel acoustics in Parkinson's disease and multiple sclerosis: Comparison of clear, loud, and slow speaking conditions. *Journal of Speech, Language, and Hearing Research*, 56(5), 1485–1502.
- Tjaden, K., Sussman, J. E., & Wilding, G. E.** (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in Parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research*, 57(3), 779–792.
- Torre, P., III., & Barlow, J. A.** (2009). Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders*, 42(5), 324–333.
- Trevino, A. C., Quatieri, T. F., & Malyska, N.** (2011). Phonologically-based biomarkers for major depressive disorder. *EURASIP Journal on Advances in Signal Processing*, 2011(1), 42.
- Uchanski, R. M., Choi, S. S., Braida, L. D., Reed, C. M., & Durlach, N. I.** (1996). Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate. *Journal of Speech, Language, and Hearing Research*, 39(3), 494–509.
- Van Engen, K. J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., & Bradlow, A. R.** (2010). The Wildcat Corpus of native- and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech*, 53(Pt. 4), 510–540.
- Van Engen, K. J., Chandrasekaran, B., & Smiljanic, R.** (2012). Effects of speech clarity on recognition memory for spoken sentences. *PLOS ONE*, 7(9), e43753.
- Van Engen, K. J., Phelps, J. E., Smiljanic, R., & Chandrasekaran, B.** (2014). Enhancing speech intelligibility: Interactions among context, modality, speech style, and masker. *Journal of Speech, Language, and Hearing Research*, 57(5), 1908–1918.
- Weikum, W. M., Oberlander, T. F., Hensch, T. K., & Werker, J. F.** (2012). Prenatal exposure to antidepressants and depressed maternal mood alter trajectory of infant speech perception. *Proceedings of the National Academy of Sciences of the United States of America*, 109(Suppl. 2), 17221–17227.
- Whyte, E. M., & Mulsant, B. H.** (2002). Post stroke depression: Epidemiology, pathophysiology, and biological treatment. *Biological Psychiatry*, 52(3), 253–264.
- World Health Organization.** (2017). *Depression and other common mental disorders: Global health estimates*. Geneva, Switzerland: Author.
- Xie, Z., Maddox, W. T., Knopik, V. S., McGeary, J. E., & Chandrasekaran, B.** (2015). Dopamine receptor D4 (DRD4) gene modulates the influence of informational masking on speech recognition. *Neuropsychologia*, 67, 121–131.