Speech-in-Noise Perception in Older Adults:
Impact of Emotional Semantic Valence and Clinical Depression

by

Emily J. Alexander

A thesis submitted in conformity with the requirements
for the degree of Master of Arts (Psychology)
Graduate Department of Psychology
University of Toronto

© Copyright by Emily J. Alexander (2019)
EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION

Speech-in-Noise Perception in Older Adults:
Impact of Emotional Semantic Valence and Clinical Depression

Master of Arts (Psychology), 2019

Emily J. Alexander

Department of Psychology
University of Toronto

Abstract
Depression influences and can be influence by interpersonal interactions, which rely on speech-in-noise comprehension. The semantic emotional salience of words may influence listening abilities in noise. This study examined the effects of emotional valence and depression on speech-in-noise comprehension in older adults (OA). Participants were 28 older adults ($M_{\text{age}} = 72.75$, $SD = 5.93$) with current depression ($n = 9$), remitted depression ($n = 9$), or healthy controls ($n = 10$). In an experimental-word-in-noise task, participants heard spoken words that were neutral, positive, or negatively-valenced. It was hypothesized that older adults with depression would show a reduction in the expected positive bias observed in healthy older adults on speech comprehension. No group differences were found in speech-in-noise comprehension ability in older adults. Across groups, there were main effects of signal-to-noise ratio and valence on EWIN task accuracy. Findings partially support a positivity bias in older adults, regardless of depression status.
Acknowledgements

First and foremost, thank you to my supervisor, Dr. Claude Alain, who provided his energy, support, compassion, and mentorship throughout this research project and writing process. Thanks to my subsidiary supervisor, Dr. Jed Meltzer, for his contributions and suggestions to this work. Thanks to team members in the EEG and MEG labs at Baycrest, for their assistance and feedback on this project, particularly Vanessa Chan for all her help. Thank you to the research assistants, Parnian Tajbakhsh and Leah Monson, who helped collect and process data. My appreciation goes out to members of Baycrest Health Sciences, the Psychology Department at the University of Toronto, and to fellow graduate students who provided feedback and support this past year. Thank you to my partner, Joshua Tolkamp, who has been my rock and helped in every way he could. Finally, a special thank you to all of the research participants who took part in this study; your generosity, patience, and kindness is appreciated.
# Table of Contents

Abstract ................................................................................................................................. ii

Acknowledgements ............................................................................................................... iii

Table of Contents ................................................................................................................... iv

Introduction ............................................................................................................................. 1

Methods ................................................................................................................................... 7

Participants ............................................................................................................................. 7

Protocol .................................................................................................................................. 12

Experimental Word-in-Noise (EWIN) Task ......................................................................... 12

Stimuli .................................................................................................................................... 12

Procedure .............................................................................................................................. 13

Additional testing .................................................................................................................. 14

Data Analysis ......................................................................................................................... 17

Results .................................................................................................................................... 17

QuickSIN ............................................................................................................................... 17

BKB ....................................................................................................................................... 18

PANAS ................................................................................................................................... 18

EWIN Task ............................................................................................................................. 20

Discussion ............................................................................................................................. 24

Conclusion ............................................................................................................................. 30

References ............................................................................................................................. 32

Appendix .............................................................................................................................. 36
Depression is a mental illness that can have serious effects on individuals and negative implications for society. Approximately 12.2% of people report having at least one major depressive episode in their lifetime (Patten et al., 2006), and rates of depression are rising. Globally, an estimated 322 million people live with depression, and rates of depression have increased by 18.4% between 2005 and 2015 (Friedrich, 2017). Major depressive disorder (MDD, also referred to as clinical depression) is distinctly different from normal feelings of sadness or loss (APA, 2013). Symptoms of clinical depression can include chronic depressed mood, marked loss of interest or pleasure from formerly enjoyable activities, feelings of worthlessness, weight gain or weight loss, fatigue or loss of energy, loss of concentration or focus, and recurrent suicidal ideation (APA, 2013). Given that depression is one of the leading causes of disability worldwide, understanding the impacts and implications of a diagnosis of clinical depression are paramount to helping individuals living with depression, and to those who may be at risk.

An important facet of human activity is connection with others through communication (i.e., interpersonal interactions). Interpersonal interactions can both influence and be affected by depression (Cohen et al., 2005; Rottenberg et al., 2005; Sass et al., 2014). Much research has investigated causes, potential risk factors, and symptoms of depression. Various psychological models of depression propose that depression spurs on self-propagating behaviours based on interpersonal interaction, but how these behaviours interact with cognitive and neurological factors of depression has not been deeply explored (Hames, Hagan, & Joiner, 2013).

Human communication is at the intersection of intra- and interpersonal factors in humans. The ability, or inability, to communicate with others can strongly influence feelings of depression
EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION

and loneliness (Horowitz, 2003). For example, depressed individuals who are admitted to nursing homes have a notably increased risk of death in the first year, which among other factors may in part be related to social isolation (APA, 2013). While past research has focused primarily on the effect of depression on speech production (reviewed in Hames et al., 2013), recent research has emerged on the relationship between depression and speech perception (Chandrasekaran, Van Engen, Xie, Beevers, & Maddox, 2015; Lidzba, Schwilling, Grodd, Krägeloh-Mann, & Wilke, 2011; Xie, Zinszer, Riggs, Beevers, & Chandrasekaran, 2019). This research suggests that individuals with depressive symptoms show impairments in listening to other speakers in noisy environments, despite having no signs of sensorineural hearing loss (Chandrasekaran et al., 2015; Keidser, Seeto, Rudner, Hygge, & Rönnber, 2015, Xie et al., 2019).

These studies, however, have primarily limited their samples to individuals with depressive symptoms, but not those with a clinical diagnosis of depression (i.e., non-clinical samples). Due to the marked difference between having symptoms of depression and having a clinical diagnosis of MDD, sampling individuals with symptoms of depression may not be the most reliable or representative method in studying depressed individuals. It is difficult to assess the direction of causality with the results of such non-clinical studies, since it is possible that other factors, such as hearing loss, especially at younger ages, could be the cause of feelings of social isolation and depression (Keidser et al., 2015). One proposition is that impairment in executive functioning and the ability to pay attention to one auditory stream among competing streams may contribute to the challenges reported by depressed individuals listening in noisy environments (Chandresekaran et al., 2015). A recent study examined the impact of depression on speech perception in noise ability in a clinical population of adults with a current diagnosis of MDD (Xie et al., 2019), which may be the first of its kind to use a clinical population in this area of research.
In older adults, some age-related hearing loss is expected, but even when this is taken into account, older adults still face significantly greater challenges at understanding speech in acoustically challenging environments (Sheft, Shafiro, Wang, Barnes, & Shah, 2015; Spyridakou & Bamiou, 2015). Notably, even when older individuals have a normal audiogram (i.e. normal, healthy hearing), they still fare relatively worse in understanding speech-in-noise (Spyridakou & Bamiou, 2015). Pure-tone audiometric assessments indicate how loud in decibels a sound must be for an individual to detect it. This is a purely auditory task, relying on the first stage of central auditory processing. In comparison, speech-in-noise (SIN) comprehension is a very complex auditory cognitive task. Understanding SIN, similar to the pure-tone audiogram, requires low-level automatic processing of the sound that is heard. This endogenous process is the first of many hierarchical stages of the central auditory system and does not require attention. The second stage of the central auditory system is a high-level cognitive processing that occurs later in time. This endogenous process is dependent on attention and cognition, both of which are known to decline in aging (Sheft et al., 2015; Zekveld, Rudner, Johnsrude, Heslenfeld, & Rönnber, 2012). Notably, cognitive decline in older adults contributes to an increased difficulty that older adults face when trying to understand SIN (Carvalho, Gonzalez, & Iorio, 2017). Older adults therefore face challenges not only in sensory decline but in the cognitive abilities required to process speech presented in noise.

Being able to selectively attend to speech in noisy environments is a key component of human communication that is often referred colloquially as the ‘cocktail party’ problem (Cherry, 1953). The need for selective attention has been demonstrated at the neural level, with neuronal responses in speech-related regions showing preferential following for an attended stream of speech compared to an unattended stream (Zion, Golumbic, et al., 2013). When attention cannot be dedicated to speech, speech recognition and comprehension also suffers; divided attention has
been shown to impair recognition of clear speech (Mattys & Wiget, 2011). Interestingly, induced anxiety produced a similar effect in recognition (Mattys, Seymour, Attwood, Munafò, 2013). Given that anxiety and depression symptoms often coincide (Lamers et al., 2011), it is unsurprising that individuals with depressive symptoms would also show difficulty listening in ‘cocktail party’ situations (Chandresekaran et al., 2015; Keidser, et al., 2015). In addition, this difficulty could exacerbate the aforementioned interpersonal difficulties that individuals living with depression experience. A common complaint from older adults is that they are able to hear speech but have difficulties understanding what is being said. This experience can be challenging from a communication standpoint but can also relate to mental health status (Keidser et al., 2015). If you are unable to understand what is going on around you, this may lead to feelings of isolation and thus contribute to depression. Research supports that self-reported hearing difficulties may be more closely associated with mental health than hearing loss assessed by pure-tone audiometry (Keidser et al., 2015). Deficits in selective attention may contribute to this common difficulty in older adults.

Aside from attentional deficits, other factors may affect speech-in-noise recognition in depressed individuals. Bower (1981) put forth an ‘affect priming theory,’ in which the mood of an individual influences their activation of information, such that mood-congruous information is processed more quickly than mood-incongruous information. According to this hypothesis, depressed individuals may have a pre-activation of negative affect nodes in their semantic network, and thus a stronger response to negative events in daily life (Cohen, Gunthert, Butler, O’Neill, & Tolpin, 2005). In the case of SIN, information carrying a negative valence may be more easily processed by depressive individuals compared to neutral- or positively-valenced information.
Alternatively, the emotional-context insensitivity hypothesis (Rottenberg, Gross, & Gotlib, 2005) posits that reactivity to emotional material in depression is regulated and corresponds to the current mood of the individual. This hypothesis has been supported in a semantic priming study that compared the priming effects of words with positive, negative, and neutral valence; whereby in depressed individuals, the emotional valence of the word had an effect on priming depending on the participants’ mood ratings (Sass et al., 2014).

Several studies have shown that some vocal emotions can be more salient than others in noisy environments, leading to better comprehension (Dupuis & Pichora-Fuller, 2008; Dupuis & Pichora-Fuller, 2014). Incongruence between semantic and vocal emotion in spoken phrases can elicit neuro-electric mismatch negativity-like responses in listeners, suggesting that this difference can be noticed automatically (Schirmer & Kotz, 2003). However, the effects of semantic emotion on SIN comprehension have not been studied as intensively as they have been in clear speech, with one study looking at ratings of valence and arousal of words used in a common speech test compared to their detection in noise (Dupuis & Pichora-Fuller, 2008).

Although the above results suggested that a neutral tone of voice taken for semantically emotional content may have altered the perceived valence and/or arousal in listeners, whether the salience of positively or negatively valenced words enhances or detracts from identification and understanding of speech in challenging listening environments has not previously been explicitly tested. This listening sensitivity to the emotional valence of words in individuals with depression may also be moderated by age. In contrast to younger adults, older adults have demonstrated a ‘positivity bias’ in attention and memory for emotional information (Mather & Carstensen, 2005). Older adults have been shown to memorize positive information better than negative information, and the reverse was found for younger adults, who memorized negative information better (Thomas & Hasher, 2006).
However, in a study examining attentional biases by measuring reaction time in a digit parity task, older adults did not show a particular positive or negative bias, while younger adults showed a bias towards negative information (Thomas & Hasher, 2006). Though older adults attended equally to all stimuli in this study, older adults still showed reliable recognition only for positive words, and not for neutral or negative words (Thomas & Hasher, 2006). In a study tracking gazing patterns at emotionally-valenced faces, younger adults displayed mood-congruent biases while older adults showed a mood-invariant positive bias (Isaacowitz, Toner, Goren, & Wilson, 2008). Other work has shown no difference between younger and older adults in automatic processing, instead showing a negativity bias for both groups in a rapid serial visual presentation task (Mickley Steinmetz, Muscatell, & Kensinger, 2010).

Late-life onset of depression has been tied to impaired executive function, which also declines with age; one proposed consequence is that older adults who demonstrate more decline of cognitive control would show less of a valence-related bias (Mather & Carstensen, 2005). Therefore, a lack of valence-related biasing of emotional word identification in noise might be seen in a population of older adults with depression compared to age-matched controls. This would fit the emotional-context insensitivity hypothesis (Rottenberg et al., 2005). On the other hand, a meta-analysis of performance on emotional Stroop and dot probe tasks revealed an attentional bias towards negative information for individuals with depression that was not moderated by age, supporting affect priming models (Peckham, McHugh, & Otto, 2010).

The current study aimed to examine the separate and combined effects of emotional valence of words and diagnoses of clinical depression on SIN comprehension in older adults. This study included participants with current MDD, remitted MDD, and a group of healthy controls. Groups in either current or remitted MDD met criteria for a clinical diagnosis of depression. An experimental word-in-noise task (EWIN) comprised of words with neutral (e.g., kettle), positive
(e.g., hug), and negative (e.g., failure) valence was administered in a single visit. As a measure of speech comprehension, participants completed this approximately 45-minute EWIN task by listening to words spoken with varying levels of background noise and repeated out loud the words they heard. Background noise consisted of multi-talker babble noise which was presented with two levels of signal-to-noise ratio (SNR) (i.e., SNR +5 and SNR 0), and three levels of speakers (i.e., 2-, 4-, or 8-speakers in the background noise).

**Hypothesis:** As supported from previous research (Mather & Cartensen, 2005; Isaacowitz et al., 2008), it was predicted that older healthy adults would show a positive bias towards understanding positively valenced spoken words. It was hypothesized that older adults with depression would show a reduction in the positivity bias observed in healthy older adults on the speech comprehension task (EWIN). Rather, clinically depressed older adults may exhibit a negative bias in understanding of speech. This would suggest that older adults with depression would perform similarly to young healthy adults, possibly showing a similar negative bias for perceiving negatively valenced spoken words (Mather & Carstensen, 2005). Such findings may shed light on what individuals with depression might have more ease or difficulty hearing in challenging listening conditions and may point to certain factors that underlie both depressive symptomatology and SIN perception ability.

**Method**

**Participants:**

Twenty-eight participants were recruited from the Baycrest participant research database and from referrals made by the research team of psychiatrist Dr. Linda Mah. Four additional participants completed testing but were excluded from analysis: one participant completed only 2 of 4 blocks of the EWIN task and was excluded due to an incomplete data set, one participant had SNR dB loss to the extent that this participant was unable to complete the EWIN task (SNR dB
loss 17.7 as assessed by QuickSIN), and two participants had pure-tone hearing thresholds that exceeded maximum values for the study (69.67 and 59.67 dB HL, respectively). Twenty-eight participants completed testing and were included in analysis for this study; all were older adults at least 60 years of age ($M_{age} = 72.80, SD = 6.17, 10$ males, $18$ females) without known dementia or memory impairment. Figure 1 presents the pure-tone audiogram results for study participants; all had hearing thresholds within normal range with respect to their age ($\leq 45$ dB HL average for all frequencies tested between 250 Hz to 8000Hz). A one-way analysis of variance (ANOVA) revealed no significant differences between groups on hearing thresholds, ($F(2, 27) = 0.36, p = .733, \eta^2_p = .03$). This study was approved by the Research Ethics Board at Baycrest. All participants gave written informed consent prior to beginning the experiment and were compensated for their time at an hourly rate.

**Figure 1.** Average pure-tone audiometric thresholds by group for frequencies tested between 250 Hz and 8000 Hz.
Older adult participants were placed into one of three groups: 9 currently depressed, 9 remitted depressed, and 10 healthy controls. Currently depressed older adults had a lifetime history of DSM-V MDD with a current major depressive episode (MDE). Remitted depressed older adults had a lifetime history of DSM-V MDD but who were in remission from depression (i.e., did not have current MDE). Presence of MDD was established by Patient Health Questionnaire (PHQ-9) score ≥ 10. All 9 participants in the current depression group had PHQ-9 scores of ≥ 10 at time of referral; four had PHQ-9 scores of 10 at time of testing and two had PHQ-9 scores of 9 at time of testing. Three additional participants were included in the current group who had Beck Depression Inventory-II (BDI-II) scores ≥ 15. The PHQ-9 and BDI-II scores were positively correlated ($r = .633$) and it was determined from interviews in addition to data collected that these three participants also met criteria for current MDE. The third participant group was comprised of age- and demographically- matched older healthy controls. Demographic information for participants is summarized in Table 1.
Table 1. Participant Demographics

<table>
<thead>
<tr>
<th>Group</th>
<th>Dep. Current</th>
<th>Dep. Remitted</th>
<th>Healthy Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (Years)</strong></td>
<td>63-84 (M = 73.89, SD = 6.72)</td>
<td>67-78 (M = 71.44, SD = 3.94)</td>
<td>61-82 (M = 72.90, SD = 6.97)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>2 males, 7 females</td>
<td>3 males, 6 females</td>
<td>5 males, 5 females</td>
</tr>
<tr>
<td><strong>Education (Years)</strong></td>
<td>12-20 (M = 15.44, SD = 2.35)</td>
<td>8-16 (M = 14.11, SD = 3.10)</td>
<td>14-24 (M = 17.50, SD = 2.88)</td>
</tr>
<tr>
<td><strong>MoCA scores (out of 30)</strong></td>
<td>20-30 (M = 26.22, SD = 3.35)</td>
<td>18-30 (M = 25.78, SD = 3.53)</td>
<td>25-30 (M = 27.80, SD = 1.99)</td>
</tr>
<tr>
<td><strong>PHQ-9</strong></td>
<td>4-27 (M = 12.22, SD = 8.32)</td>
<td>1-7 (M = 3.56, SD = 1.88)</td>
<td>0-3 (M = 1.10, SD = 1.10)</td>
</tr>
<tr>
<td><strong>BDI-II</strong></td>
<td>7-23 (M = 14.67, SD = 5.45)</td>
<td>0-11 (M = 6.67, SD = 3.20)</td>
<td>1-7 (M = 3.30, SD = 2.21)</td>
</tr>
</tbody>
</table>

*PHQ-9 = Patient Health Questionnaire 9, BDI-II = Beck Depression Inventory II; MoCA = Montreal Cognitive Assessment.

The inclusion of patient groups comprised of both currently depressed older adults and older adults currently in remission enabled us to distinguish between underlying cognitive vulnerabilities and any transient effects of acute depression state. Participants with MDD were outpatients seeking treatment at Baycrest Hospital in Toronto, Ontario, Canada or from the Toronto community recruited through newspaper advertisements. All participants in either the currently depressed or remission groups met DSM-V diagnostic criteria for MDD as established by a psychiatrist (i.e., Dr. Mah) or a qualified delegate. The majority of participants were monolingual, and 10 were bilingual. All participants were fluent in English and either spoke English as a first language or spoke English as a second language but learned English by 5 years of age.
Exclusion criteria specific to participants with MDD were as follows: a) serious suicidal ideation or behavior, and b) inability to provide informed consent. Exclusions specific to controls included having a history of neurologic or other DSM-V major psychiatric disorder including drug or alcohol abuse (simple phobia would not be excluded).

Exclusion criteria for all groups included: a) a history of drug (including benzodiazepines) or alcohol abuse within the previous 12 months, or a lifetime history of alcohol or drug dependence (as assessed using DSM-V-criteria); b) significant self-reported medical or neurocognitive conditions and interventions that may significantly impact cognitive function (e.g., hepatitis C, electroconvulsive therapy within the past 6 months, epilepsy, development delay, attention deficit disorder); c) neurological illness/disease including history of head trauma with loss of consciousness > 20 minutes, stroke; and d) English language acquired after age 5.

Sample size was estimated based on a calculation of effect size from Sass et al. (2014), a study that looked at the effect of semantic emotion on priming and the differences between individuals with and without depression. The effect was the Group by Valence interaction in response time, $\eta^2_p = 0.13$. Using this effect size and a 3 x 3 design in G*Power 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007), with the conventional values of $\alpha = 0.05$ and $(1-\beta) = 0.8$, estimated sample size required was $n = 80$. An additional estimation was conducted on a correlation ($r = -0.44$) from Chandrasekaran et al. (2015) of ratings of depressive symptoms and performance on a speech-in-noise test in a group of participants classed as high-depressive: in a two-tailed test with conventional values of $\alpha = 0.05$ and $(1-\beta) = 0.8$, estimated sample size required was $n = 24$, suggesting that our estimated sample size should also be able to capture any within-group correlations between level of depression and task performance.
EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION

Protocol:

All participants were screened by phone and/or email to ensure study eligibility criteria were met. The entire experiment took place in a single session over a 2-hour testing period. Participants had the option of completing testing in two separate sessions on separate days within one week of each other, however all participants opted to complete testing in a single session. Participants were contacted by phone and/or email to confirm eligibility in the study. Participant demographics, including a health history and depression history, were collected during phone conversations and in-person during the testing session. All testing took place at the Rotman Research Institute at Baycrest Health Sciences in Toronto. During part 1 of testing, participants were seated in a sound-attenuating booth to complete the experimental word-in-noise task. Three participants were regular hearing aid users for mild hearing loss; one in each of the above subject groups. Participants who used hearing aids did not use them for the pure-tone threshold testing, QuickSIN, BKB, or EWIN. All participants heard the EWIN at the same volume levels, and no corrections were made to EWIN stimuli for participants who used hearing aids.

Experimental Word-in-Noise (EWIN) Task

Stimuli

All words used in the main experimental task were selected using data from the Affective Norms for English Words (ANEW; Bradley & Lang, 1999), a normative set of ratings of emotional affect for a large number of English words. One hundred and twenty words from each of three specific ranges were selected for three valence categories: positive (valence rating > 7), negative (valence rating < 3), and neutral (valence rating 4-6), for a total of 360 words. Words selected for each valence category were matched in frequency, length, and arousal using the English Lexicon Project database (Balota et al., 2007). Words selected were balanced on both valence and arousal (see Appendix A for statistics on words selected).
EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION

All words were recorded by a female speaker of North American English in a sound-attenuated booth at Baycrest Health Sciences using a Shure KSM44 microphone. The stimulus set was finalized offline by editing words at zero-crossings and was equalized for RMS amplitude. Stimuli was mixed with multi-talked babble noise at two signal-to-noise ratios (SNRs) offline. Each stimulus was embedded in either: 2-talker, 4-talker, or 8-talker babble noise at an SNR of +5 or 0. Stimulus presentation and response recording was conducted using Presentation software (Neurobehavioral Systems, Albany, CA, USA) at a level of 75dB sound pressure level (SPL), measured with a Larson Davis SPL meter (Model 824) using a 2-cc sampler. Regardless of pure-tone audiometric assessments, all participants listened to stimuli from the main experimental task at the same volume (75dB HL). Participants were presented words in varying levels of noise through Etymotic ER-3A insert earphones (Etymotic Research, Elk Grove, IL, USA) and asked to repeat them out loud. All audio files were labeled only by participant ID only to maintain participant confidentiality.

Procedure

Each participant listened to four blocks of stimuli, each containing 90 words for a total of 360 words as described above. Order of words considering valence, number of speakers, and SNR ratio was randomized for each participant individually. Participants heard one word at a time presented simultaneously with background noise in short audio segments and were instructed to listen for the target word (spoken by the same female speaker over trials, and who was not one of the background speakers). Prior to starting the task, participants were able to hear a sample of the experimental stimuli with responses to familiarize themselves with the procedure and expectations. Instructions appeared on a computer monitor placed in front of the participant prior to the start of each block. Once the experiment began, a small ‘t’ would appear on the computer screen during presentation of each short sound clip, and the word ‘respond’ would then appear
when the sound clip ended, prompting the participant to say out loud what they heard. If a participant could not decipher what word was spoken amidst the background noise (who were either 2-, 4-, or 8-talker babble noise), they were instructed to say ‘skip.’ The word ‘skip’ was not one of the target words and allowed researchers to create an initial transcription of responses. Upon repeating what they heard out loud, participants would press the space bar on a keyboard placed in front of them to move to the next trial. The EWIN task was self-paced; participants controlled how quickly they moved through the four blocks of stimuli, therefore there were individual differences in time to complete the task. Between each block of 90 words participants were given a short break. All participants completed all four blocks of 90 words, for a total of 360 words heard by each participant.

**Additional Testing**

After the EWIN task, a series of assessments was administered to characterize patient and control populations along the axes of hearing ability, mood, and cognition. These assessments included:

1. **A pure-tone audiometric threshold test.** This test assesses sensorineural hearing ability, identifying hearing thresholds and assessing hearing loss. Participants were presented with a series of tones and instructed to press a button each time they detected a tone. Pure tone thresholds were collected using the standard clinical procedure for each octave between frequencies 250 to 8000 Hz binaurally using an audiometer (model GSI-16).

2. **The Quick Speech-in-Noise task (QuickSIN)** (Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004). The QuickSIN task measures SNR loss (signal-to-noise ratio loss), which indicates the ability of participants to hear speech in noise. Participants listened to short, pre-recorded sentences in the presence of multi-talker babble noise and were asked to repeat the words they heard. The sound was presented binaurally at 70 dB SLP, and the
SNR ratio was lowered in 5-dB steps from 25 dB to 0 dB, progressing from normal to severe impairment in performance. The background noise was initially very low and increased progressively, making it more difficult for participants to decipher the recorded sentences. Each participant was presented with five lists of six sentences, each containing five key words per sentence. Participants were scored one point for each of the correct five words per sentence. The SNR loss was determined by subtracting the total number of words correct from 25.5, which represents the SNR required for participants to correctly identify 50% of the key words (Killion et al., 1997).

3. **The Bamford-Kowal-Bench (BKB) sentence tests** (Bench, Kowal, & Bamford, 1979). The BKB sentence test also measures SNR loss, indicating the ability of participants to hear speech in noise. Similar to the QuickSIN, participants listened to short pre-recorded sentences in the presence of multi-talker babble noise and were asked to repeat back what they heard. Sound was presented binaurally at 70 dB SLP, and the SNR ratio was lowered in 3-dB steps from 21 dB to -6 dB. Each participant was presented 2 lists, split into A and B sections, with 10 sentences per list. This task was initially quite easy but became progressively more difficult as the SNR ratio decreased. The first sentence of each of the 4 list sections contained four key words, and each subsequent sentence contained three key words. Participants were scored one point for each of the correct key words per sentence. The SNR loss for the BKB was determined by subtracting the total number of words correct from 23.5.

4. **The 9-item Patient Health Questionnaire** (PHQ-9; Kurt, L., & W., 2001). The PHQ-9 is a 9-question assessment to detect the presence and severity of depression symptoms. This questionnaire is used to establish depression diagnosis and is primarily used in primary healthcare settings. It is a short, written assessment that asks participants to rate how often
they have been bothered by specific problems (e.g., feeling tired or having little energy) over the past two weeks.

5. **The Positive and Negative Affective Schedule** (PANAS; Watson, Clark, & Tellegen, 1988). This short mood scale consists of twenty words that describe different feelings and emotion. The twenty items are formed by combining two 10-item scales, one with positive affect words (e.g., proud) and the other with negative affect words (e.g., ashamed). Participants indicated on paper, using a 1 to 5 Likert scale, the extent to which they felt a particular feeling or emotion at the time of completion of the PANAS.

6. **The Beck Depression Inventory** (BDI-II; Beck, Steer, & Brown, 1996). This is a validated measure of depression symptoms in clinical and healthy populations. The BDI-II consists of twenty-one items, which each are comprised of a group of statements describing feelings and beliefs. Participants rate on a 4-point scale (ranging from 0 to 3) the statement that best represents how they have felt in the past two weeks, including the day of testing. Responses are tallied and scores from page 1 and page 2 are added to calculate the total BDI-II score.

7. **The Montreal Cognitive Assessment** (MoCA; Nasreddine et al., 2005). The MoCA is a brief 30-question test used to screen for mild cognitive impairment (MCI) and dementia. The MoCA evaluates different types of cognitive abilities, including: orientation, short-term memory/delayed recall, executive function/visuospatial ability, language abilities, abstraction, animal naming, attention, and a clock drawing test. Participants completed this assessment in a quiet testing suite with a researcher or research assistant trained in MoCA administration.
Data Analysis:

Statistical analyses were conducted using R Version 3.4.1 (R Core Team, 2017), package “ez” (Lawrence, 2016) and using SPSS Version 24.0 (IBM Corp, 2016). For analysis of the EWIN task results, R Version 3.4.1 package was used. For each participant, a raw score for accuracy on the EWIN task was calculated. Each participant heard 360 words across 18 possible combinations of SNR (i.e., 5 dB or 0dB), number of speakers in background noise (i.e., 2, 4, or 8), and valence (i.e., positive, negative, or neutral). Therefore, the maximum raw score per combination of the above three factors was 20. Accuracy percentages for the EWIN task were calculated for each participant by dividing each raw score by 20 and multiplying by 100. A dataset was compiled with participant ID, group, SNR, valence of words, number of speakers, accuracy raw scores on the EWIN task, and accuracy percentage on the EWIN task. A linear mixed model was used to analyze the accuracy percentages for all participants on the EWIN task. Fixed effects in the model included SNR, number of speakers, and valence. Fixed effects were treated as categorical variables (Chandrasekaran et al., 2015). A linear mixed model with random slopes for the variables of speaker, SNR, and valence, and a random intercept for each participant resulted in a singular fit. A model comparison between a linear mixed model with and without the random slopes of speaker, SNR, and valence, which both included a random intercept for each participant, revealed no significant differences in fit between the models ($\chi^2(20) = 23.47, p = .266$). Therefore, a fixed model with a random intercept for each participant was chosen for analysis of the accuracy percentage on the EWIN task.

Results

QuickSIN

The average dB SNR loss measured by QuickSIN performance is presented by group in Figure 1. A one-way analysis of variance (ANOVA) revealed no significant differences across
groups on QuickSIN scores \((F(2, 27) = 0.81, p = .456, \eta_p^2 = .06)\). Older adults with current MDD \((M = 1.77, SD = 2.50)\) did not significantly differ from those with remitted MDD \((M = 2.88, SD = 1.64, p = .281)\) or from those in the control group \((M = 2.86, SD = 2.19, p = .277)\).

![Figure 2. Average SNR performance on QuickSIN task by group.](image)

**Figure 2.** Average SNR performance on QuickSIN task by group.

**Bamford-Kowal-Bench**

A one-way ANOVA revealed no significant differences across groups on BKB scores \((F(2, 27) = 0.08, p = .927, \eta_p^2 = .01)\). Older adults with current MDD \((M = 0.00, SD = 0.99)\) did not significantly differ from those with remitted MDD \((M = -0.04, SD = 1.38, p = .775)\) or from those in the control group \((M = 0.05, SD = 1.29, p = .940)\).

**Positive and Negative Affectivity Scale**

A one-way ANOVA analyzed the average scores of the positive and negative components of the Positive and Negative Affectivity Scale (PANAS) separately. There was a significant effect of group for scores associated with positive emotions \((F(2, 27) = 5.74 , p = .009, \eta_p^2 = .32)\).

Planned comparisons revealed that older adults in the current MDD group significantly differed in positive feelings and emotions felt at time of testing compared to the control group. Individuals in the currently depressed group \((M = 2.76, SD = 0.39)\) identified less with positive feelings and
EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION

emotions at time of testing compared to the control group ($M = 3.68$, $SD = 0.72$, $p = .002$). Figure 3a presents the average positive scores on the PANAS across groups. There was no significant difference in average positive scores between the control and remission groups ($p = .165$), and older adults in remission from depression had a trend towards higher average positive PANAS scores compared to the currently depressed group ($p = .069$).

![Figure 3a](image1.png)

**Figure 3a.** Average positive scores on the PANAS by group.

![Figure 3b](image2.png)

**Figure 3b.** Average negative scores on the PANAS by group.
A one-way ANOVA revealed a trend level effect of group on average negative PANAS scores ($F(2, 27) = 2.36$, $p = .115$, $\eta^2_p = .16$). Planned comparisons revealed that the remitted depression group did not significantly differ in reporting negative emotions at time of testing from the current depression group ($p = .215$) or from the control group ($p = .398$). Further planned comparisons revealed that the current depression group reported significantly higher average scores for negative feelings and emotions compared to the control group ($p = .040$). Figure 3b presents average negative scores on the PANAS across groups.

To assess the relationship between depression status and scores on the PANAS, a Pearson product-moment correlation was calculated for (a) scores on the BDI-II and positive PANAS values and (b) BDI-II scores and negative PANAS values. Across groups, BDI-II scores had a moderate negative correlation with positive PANAS scores, $r = -.548$, $n = 28$, $p = .003$. There was also a small positive correlation between BDI-II scores and negative PANAS scores, $r = .372$, $n = 28$, $p = .051$.

**Experimental Word-in-Noise (EWIN) Task**

An analysis of deviance (Type II Wald chi-square tests) found no significant main effects of group on accuracy percentage of the WIN task ($\chi^2(2) = 0.0578$, $p = .972$). There were no significant group differences between currently depressed individuals, individuals with remitted depression, and healthy controls on the accuracy percentage of word responses. Figure 4 presents the accuracy (%) of words correctly detected in the EWIN task by group.
Figure 4. Accuracy (%) on EWIN task by group.

For subsequent analysis, results were averaged over the levels of group, with confidence level 0.95. There was a significant main effect of SNR on accuracy percentage on the EWIN task ($\chi^2(1) = 1329.46, p = 2.2 \times 10^{-16}$). Across groups, accuracy percentage on the EWIN task was significantly different between stimuli presented at SNR +5 compared to SNR 0. There was also a small main effect of valence on accuracy percentage across groups, ($\chi^2(2) = 6.72, p = .035$).

Accuracy percentage on the EWIN task for older adults significantly differed between target words presented at different valences (i.e., positive, negative, or neutral). Pairwise comparisons revealed that older adults had higher accuracy on the EWIN task for positively-valenced words compared to negatively-valenced words ($p = .036$). There was no significant difference found in accuracy on the EWIN task for positive words compared to neutral words ($p = .923$), and no significant difference found in accuracy for neutral compared to negative words ($p = .091$).

There was a significant interaction between number of speakers and valence of words ($\chi^2(4) = 12.87, p = .0119$). Planned contrasts revealed that at the 4-speaker level only, older adults
had higher scores of EWIN comprehension for positive words than for negative words ($p = .0061$). At the 4-speaker level only, older adults also had higher accuracy in EWIN comprehension for neutral words compared to negative words ($p = .002$). At the 4-speaker level, there was no significant difference in accuracy in identifying positive words compared to neutral words in the EWIN task ($p = .943$). No significant differences were found in accuracy percentage at the 2- or 8-speaker levels for any valence comparison. Results for planned contrasts were averaged over the levels of group and SNR for the interaction between speakers and valence. Figure 5 presents the accuracy (%) of words correctly detected in the EWIN task by number of speakers and valence of words, across all participants.

![Figure 5](image_url)

**Figure 5.** Accuracy (%) on EWIN task by number of speakers and valence of words across groups.

There was also a significant interaction between SNR and valence of words on accuracy percentage on the EWIN task ($\chi^2 (2) = 9.28, p = .00967$). Figure 6 presents the accuracy (%) correct on the EWIN task at SNR +5 and SNR 0 across groups. Planned contrasts revealed that for stimuli presented at SNR +5, older adults had significantly higher accuracy in identifying
positively-valenced words compared to both negatively-valenced words ($p = .0266$) and neutrally-valenced words ($p = .0386$). No significant difference was found between accuracy percentage at SNR +5 between negative and neutral words ($p = .989$). Further planned contrasts revealed that older adults had significantly higher accuracy for neutral words compared to negative words at SNR 0 ($p = .0135$). When stimuli in the EWIN task were presented at SNR 0, older adults better identified neutral words compared to negative words. Results for planned contrasts were averaged over the levels of group and speaker for the interaction between SNR and valence.

![Figure 6](image)

**Figure 6.** Accuracy (%) on EWIN task by SNR and valence of words across groups.

Additionally, Pearson correlations were calculated between accuracy scores (%) on the EWIN task and average SNR loss for the QuickSIN. As anticipated, there was a strong negative correlation between accuracy on EWIN and SNR loss, whereby individuals across groups with lower SNR dB loss (i.e., lower QuickSIN scores and therefore better SIN comprehension), had significantly higher accuracy on the EWIN task ($r = -.802, n = 28, p = .000$). A bivariate
correlation was computed to assess the relationship between QuickSIN scores and BKB scores, which had a moderate to high positive correlation ($r = .70, n = 28, p = .000$).

**Discussion**

Major depressive disorder is a mental illness that is marked by consistent depressed mood and loss of interest in formerly pleasurable activities, and individuals diagnoses with MDD are at a higher risk for self-harm or suicide (APA, 2013). Interpersonal interaction is an important human activity that can help mitigate feelings of depression and isolation, but older adults often struggle with understanding speech in the presence of background noise. Considering the difference in profile of both listening ability and emotional processing of individuals with depression, listening sensitivity in those with depression may vary according to the emotional valence of words. This study is one of the first to examine the influence of clinical depression and emotional semantic valence on SIN comprehension abilities. More specifically, it is believed to be the first to examine this specifically in older adults, who face additional challenges in understanding SIN compared to younger adults.

Older adults completed two speech in noise tasks (i.e., the QuickSIN & BKB), the PANAS, and a main experimental word-in-noise task. No disadvantage in performance on the QuickSIN or the BKB was found for the current depression or remitted depressed groups compared to the healthy control group. Older adults with current MDD had lower scores of positive emotions on the PANAS compared to both remitted MDD and healthy control groups. Older adults with current depression identified with positive feelings less, and identified slightly more with negative feelings at time of testing, as measured on the PANAS. While there was no main effect of group on the EWIN, across groups there was a positive bias shown in results, whereby older adults accurately identified positively-valenced words better, but only when the background noise contained 4-speakers. Findings partially replicated the ‘positive bias’ in
attention and memory for emotional information (Mather & Cartensen, 2005), however results must be interpreted cautiously due to conflict with previous research.

The QuickSIN task assessed older adults’ SIN comprehension abilities. A higher score on the QuickSIN indicates lower SIN comprehension abilities, and a lower score indicates higher speech comprehension abilities. Current MDD, remitted MDD, and healthy control groups all performed similarly on this task. These results contradict with growing literature on the relationship between depression and SIN perception, suggesting that individuals with depressive symptoms show impairments in listening to other speakers in noisy environments, despite no findings of sensorineural hearing loss (Chandrasekaran et al., 2015; Keidser, Seeto, Rudner, Hygge, & Rönnber, 2015).

The BKB task also assessed older adults’ SIN comprehension abilities, and similar results were found, with no differences between groups on average SNR loss (i.e., no difference in performance, or ability to understanding speech in noise). The present study found that older adults with current MDD were comparable in SIN comprehension ability to those with remitted MDD, and to healthy controls. One notable difference between participants in the present study compared to previous research is that participants included had a DSM-V diagnosis of MDD, while previous research has primarily focused on non-clinical samples with depressive symptoms. One recent study may be the first to examine SIN comprehension ability using a clinical sample of individuals with diagnosed MDD (Xie et al., 2019). This study also found that individuals with current MDD had lower performance on assessments of SIN comprehension ability, which was demonstrated on the BKB (Xie et al., 2019). This study included mainly younger adults, therefore differences in SIN comprehension may be affected by age. The present study included groups of individuals who all met criteria for clinical diagnosis of MDD at time of referral, who were then sorted into current MDD and remitted MDD. Many of the participants included in the present
study were undergoing current psychological treatment, and some were taking psychotropic medications to help with symptoms. Therefore, it may be the case that individuals in the current MDD group improved in symptoms from time of referral to time of testing and this influenced the null finding of group differences. A recent study by Xie et al. (2019) had a considerably larger sample size and therefore further studies with greater sample size and power are needed to determine whether the effect of MDD on the ability to understand SIN exists across the lifespan.

The PANAS was administered as a means of assessing participants’ current mood. Participants rated how much they felt a given positive or negative emotion in the present moment. Ratings of positive and negative emotions were scored and analyzed separately. For positive emotions, there was a significant effect of group. Older adults in the current MDD group reported lower ratings of positive emotions felt at time of testing compared to the control group. As anticipated, this means that currently depressed older adults identified less strongly with positive feelings compared to healthy controls. Although the difference in positive PANAS scores between older adults with current MDD and remitted MDD was not significant, those with remitted MDD also showed a trend of identifying more strongly with positive emotions than those with current MDD.

For the negative PANAS scores, there were no significant differences between groups on ratings of negative emotions felt at time of testing. There was a trend that older adults with current MDD identified most strongly with negative emotions, next to older adults with remitted MDD, and healthy controls identifying least with negative emotions. These results were necessary to evaluate the emotional-context sensitivity hypothesis (Rottenberg et al., 2005), which states that reactivity to emotional information in depression is regulated, and this corresponds to the current mood of an individual. In contrast, Bower (1981)’s ‘affect priming theory’ posits that the mood of an individual influences their activation of information, such that mood-congruous
information is processed more quickly than mood-incongruous information. The ‘affect priming theory’ does not speak to the current mood of the individual, as much as it evaluates general mood. In the context of understanding SIN, it was hypothesized that depressed individuals would more easily process negatively-valenced words compared to neutral- or positively-valenced words.

The EWIN, however, did not support this hypothesis. No significant group differences were found on the EWIN task. This was an unexpected finding, and possible explanations for a lack of differences between groups will be discussed in the context of current literature and limitations of the study. Since no group differences were found on accuracy on the EWIN task, the levels of group were averaged for subsequent analyses. Across groups, there was a significant main effect of SNR on accuracy percentage. Accuracy in detecting the target word on the EWIN was significantly higher for SNR +5 compared to SNR 0. Stimuli presented at SNR +5 are easier to detect, even in background noise, compared to stimuli presented at SNR 0. There was also a main effect of valence on accuracy percentage across groups. Older adults, in general, significantly differed in accuracy on the EWIN task depending on the valence of the target word presented (i.e., positive, negative, or neutral). Given that this finding is not group-specific, this does not clearly support Bower (1981)’s ‘affect priming theory’ or Rottenberg et al. (2005)’s emotional-context insensitivity hypothesis. In the present study, results from the PANAS revealed group differences in positive emotions currently experienced, but no significant group differences between groups for negative emotions, despite a trend for current MDD to feel more negative emotions. There are significant group differences on current mood of individuals, but this does not appear to affect accuracy in detecting words of difference valences at the group level.

Across all groups, there was a significant interaction between the number of speakers and valence of words in the EWIN task. Older adults had higher accuracy on the EWIN task for
positive words compared to negative words, which is consistent with a ‘positivity bias’ in attention and memory for emotional information demonstrated by Mather and Carstensen (2005). Previous research has shown support for a positivity bias in older adults (Thomas & Hasher, 2006), and one study found this positive bias in older adults to be mood-invariant (Isaacowitz et al., 2008). The present study supports this finding: older adults, despite differences between groups in current mood, overall correctly identified positive words more than negative words on the EWIN task. Further analysis revealed that older adults only showed higher accuracy for positive over negative words at the 4-speaker level of background noise. Across groups, older adults also had lower accuracy for negative words compared to neutral words at the 4-speaker level, which supports the positivity bias in older adults, but does not support differentiation among groups in this study.

Chandrasekaran et al. (2015) proposed that impairment in executive functioning, more specifically the ability to pay attention to one auditory stream among competing streams, may add to difficulty reported by depressed individuals in SIN comprehension. Variations in the number of background speakers in the present study aimed to address this difficulty, but results could not speak to differences in the ability to pay attention to one auditory stream among competing streams as no group differences were found. Why specifically the positive over negative bias was found at the 4-speaker level compared to the 2-speaker or 8-speaker level of background speakers is not known. It is surprising that a difference was found at the 4-speaker level but not at the 2- or 8-speaker levels of background noise. Fewer speakers in the background noise could increase difficulty in attending to one stream over another single stream, while more speakers in background noise may reduce the executive functioning required to segregate one stream, however further research is needed in this area. Neither of these explanations was supported in the findings.
A significant interaction was also found between SNR and valence on accuracy percentage on the EWIN task. At SNR +5, older adults had higher accuracy in identifying positively-valenced words compared to both negatively-valenced words and neutrally-valenced words. This further supports a positivity bias in older adults. The aforementioned differences in older adults in accuracy of detecting words of difference valences (i.e., positive, negative, neutral) contributes to research examining the effects of semantic emotion on SIN comprehension (Dupuis & Pichora-Fuller, 2008). These effects have been studied more in-depth in clear speech, but interest is growing in understanding the effects of depression and age on SIN comprehension (Dupuis & Pichora-Fuller, 2008; Xie et al., 2019). At SNR 0, older adults had significantly higher accuracy for detecting neutral words compared to negative words on the EWIN. This means that older adults better identified neutral words in background noise than negative words, by why this was found exclusively at SNR 0 is unclear.

Importantly, overall findings of the impact of depression on speech comprehension in noise ability were not consistent with a very recent study by Xie et al. (2019). This study, unlike most previous research, included a clinical sample of individuals who were clinically diagnosed with MDD. This study assessed SIN comprehension using sentence recognition tasks across three conditions with maskers varying the extent of linguistic content (Xie et al., 2019). They analyzed speech recognition errors, finding higher errors for individuals with MDD compared to neurotypical controls. Xie et al. (2019) suggested that the depression-related deficit in speech perception may be related to increased distractibility caused by linguistic interference in background talkers. This study included primarily young adults, and therefore results may differ for older adults as executive functioning and hearing ability declines with age.

There are two important limitations to note within this study. Firstly, the sample size included was low, and a greater sample size is required for statistical power. The recent study by
Xie et al. (2019) included a larger sample size and found significant differences between currently depressed individuals and healthy controls; therefore it is possible that group differences could not be found in the current study even if they were present due to low power. Ongoing testing will increase sample size for the present study to evaluate this possibility. Secondly, although individuals were screened for MDD, results of the QuickSIN, BKB, and EWIN suggest minimal differences between groups that have been shown in previous research. It is possible that the individuals in the current MDD group suffer primarily from chronic depression that fluctuates in intensity, compared to more acute diagnoses of MDD. Additionally, individuals included in the current MDD group based on PHQ-9 and BDI-II scores would not meet criteria for “severe” depression, but rather more mild and moderate cases of depression. It is possible that acute cases of MDD would better represent current MDE. Future studies will need to more carefully classify and segregate groups based on depressive symptoms, DSM-V diagnoses, as well as other factors. Findings from the present study should be interpreted with caution, as depression and emotional valence of words may still impact SIN comprehension ability in older adults. Despite these limitations, this study is, to the best of our knowledge, the first of its kind to examine the impact of clinical depression and emotional semantic valence on SIN comprehension abilities in older adults. In the event that the current MDD group better represents individuals with remitted MDD as depressive symptoms as measured by PHQ-9 and BDI-II were not ‘severe’, the findings of this study are still important in contributing to aging and depression research, and understanding SIN comprehension abilities in older adults.

**Conclusion**

This study examined the effects of clinical depression and emotional semantic valence on the ability to understand SIN in older adults. Based on the findings from the EWIN, there are no differences in SIN comprehension ability between older adults with current MDD, remitted
EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION

MDD, and healthy controls. However, due to low sample size and group classification, older adults with clinical depression may still face additional barriers in understanding SIN, and valence of words may influence what they hear best. The present study found that across all groups (i.e., older adults with current MDD, remitted MDD, and health controls), older adults are better able to identify positively-valenced words in 4-speaker background noise. Findings partially support a positivity bias in older adults for remembering positively-valenced information better than negatively-valenced information, which appears to be mood-invariant (Isaacowitz et al., 2008). Findings also support previous research on the influence of valence on SIN comprehension (Dupuis & Pichora-Fuller, 2008). Contrary to hypotheses, older adults with current MDD did not show a reduction in the positivity bias, as this bias was observed across all groups of older adults. Study findings contribute to a growing body of literature on the influence of depression on speech comprehension abilities, specifically SIN comprehension in older adults. Future studies should include larger sample sizes and more rigorously segregate groups by depression status.
EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION

References


EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION


EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION


EMOTION AND DEPRESSION IN SPEECH-IN-NOISE PERCEPTION


Appendix

Table 1

*Summary statistics for the words used in Study 1*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Negative</th>
<th>Neutral</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of syllables</td>
<td>1.92 (0.71)</td>
<td>1.88 (0.72)</td>
<td>1.98 (0.73)</td>
</tr>
<tr>
<td>Log frequency (HAL)</td>
<td>8.68 (1.27)</td>
<td>8.81 (1.38)</td>
<td>8.72 (1.34)</td>
</tr>
<tr>
<td>Valence</td>
<td>2.18 (0.35)</td>
<td>5.19 (0.35)</td>
<td>7.85 (0.38)</td>
</tr>
<tr>
<td>Arousal</td>
<td>5.84 (0.97)</td>
<td>4.26 (0.81)</td>
<td>5.74 (1.01)</td>
</tr>
</tbody>
</table>

*Note.* Values presented are the mean, and values in parentheses are the standard deviation. HAL = Hyperspace Analogue to Language frequency norms (Lund & Burgess, 1996).