Effects of Acquired Aphasia on the Recognition of Speech Under Energetic and Informational Masking Conditions

Sarah Villard[1] [ID] and Gerald Kidd, Jr.[1]

Abstract

Persons with aphasia (PWA) often report difficulty understanding spoken language in noisy environments that require listeners to identify and selectively attend to target speech while ignoring competing background sounds or "maskers." This study compared the performance of PWA and age-matched healthy controls (HC) on a masked speech identification task and examined the consequences of different types of masking on performance. Twelve PWA and 12 age-matched HC completed a speech identification task comprising three conditions designed to differentiate between the effects of energetic and informational masking on receptive speech processing. The target and masker speech materials were taken from a closed-set matrix-style corpus, and a forced-choice word identification task was used. Target and maskers were spatially separated from one another in order to simulate real-world listening environments and allow listeners to make use of binaural cues for source segregation. Individualized frequency-specific gain was applied to compensate for the effects of hearing loss. Although both groups showed similar susceptibility to the effects of energetic masking, PWA were more susceptible than age-matched HC to the effects of informational masking. Results indicate that this increased susceptibility cannot be attributed to age, hearing loss, or comprehension deficits and is therefore a consequence of acquired cognitive-linguistic impairments associated with aphasia. This finding suggests that aphasia may result in increased difficulty segregating target speech from masker speech, which in turn may have implications for the ability of PWA to comprehend target speech in multitalker environments, such as restaurants, family gatherings, and other everyday situations.

Keywords

aphasia, auditory masking, speech recognition, psychoacoustics, hearing loss

Received 22 July 2019; Revised 24 September 2019; accepted 1 October 2019

Introduction

Studies examining receptive language processing in persons with aphasia (PWA) typically conduct testing in quiet settings, under the reasonable assumption that any competing background noise could confound assessment of participants' comprehension abilities. However, conversations in daily life do not always take place in quiet rooms; rather, many unfold in busy restaurants and crowded stores, inside cars and buses, on city sidewalks, at family dinner tables and other social gatherings, or in the presence of background sounds from a television or radio. Therefore, while studies assessing receptive language processing in quiet provide valuable information about pure comprehension in aphasia, their results may not fully capture the ability of PWA to process speech in real-world situations. Indeed, many PWA anecdotally report difficulty understanding speech in noisy environments (Skelly, 1975).

The challenge of attending to a target talker in the presence of auditory maskers, known as the "cocktail party problem" (Cherry, 1953), is nearly unavoidable in daily life and has been studied extensively in the general population (for reviews, see Bronkhorst, 2015; Middlebrooks, Simon, Popper, & Fay, 2017;

[1]Department of Speech, Language & Hearing Sciences, Boston University, MA, USA

Corresponding Author:
Sarah Villard, Boston University, 635 Commonwealth Avenue, 3rd Floor, Boston, MA 02215, USA.
Email: sarah.n.villard@gmail.com

Yost, 1997). A notable finding from this literature is that some groups of listeners encounter more difficulty than others in identifying and processing a target speech stream. In particular, older listeners (Ezzatian, Li, Pichora-Fuller, & Schneider, 2015; Gifford, Bacon, & Williams, 2007; Helfer, Chevalier, & Freyman, 2010; McCoy et al., 2005) and listeners with hearing loss (HL; Best, Thompson, Mason, & Kidd, 2013; Festen & Plomp, 1990; Gallun, Diedesch, Kampel, & Jakien, 2013; Marrone, Mason, & Kidd, 2008a) have been found to exhibit poorer performance than controls on masked listening tasks.

These findings regarding age and HL are pertinent when considering how PWA may perform on similar listening tasks. Not only is stroke more common in older individuals, but aphasia is more likely to occur in older than in younger stroke patients (Ellis & Urban, 2016; Engelter et al., 2006); furthermore, many PWA demonstrate some degree of HL (Formby, Phillips, & Thomas, 1987; Silkes & Winterstein, 2017; Zhang et al., 2018). In addition to the effects of age and HL, there is evidence that PWA may experience further breakdowns in processing target speech due to their acquired cognitive-linguistic impairments (Rankin, Newton, Parker, & Bruce, 2014; Winchester & Hartman, 1955). However, further research is needed to better understand the nature of this added difficulty and to distinguish it from the effects of pure comprehension deficits as well as to identify the factors that may facilitate or hinder successful processing of target speech by PWA.

Gaining a better understanding of how PWA process speech in complex acoustic environments is a topic of both practical and theoretical significance. The ease/difficulty with which PWA are able to selectively attend to a conversational partner in a real-world acoustic environment or follow changes in talkers during conversation may have a direct impact on social engagement and community participation in this population. These consequences, in turn, may affect psychosocial well-being and quality of life (Cruice, Worrall, Hickson, & Murison, 2003; Hilari, Needle, & Harrison, 2012). In addition, gaining a better understanding of how PWA process speech under adverse conditions may shed light on the relationships between various cognitive-linguistic mechanisms in aphasia, such as selective attention, auditory processing, and language comprehension.

This article describes a study intended to identify the influence of acquired aphasia on performance on a speech processing task under masked listening conditions by systematically examining the effects of different types of masking, in PWA and in age-matched healthy controls (HC).

Energetic and Informational Masking

Work on the cocktail party problem in the general population has identified two distinct types of masking produced by nontarget sounds. When energy from target and masker sources reach the human ear simultaneously, overlap in the representations of the sounds on the basilar membrane and in the auditory nerve may occur, obscuring portions—or all—of the target sound and causing energetic masking (EM). However, under conditions where there remains a sufficient neural representation of the target sound to support identification or comprehension, additional masking often occurs that cannot be explained solely by the spectrotemporal overlap of excitation in the periphery (e.g., Arbogast, Mason, & Kidd, 2002; Brungart, Simpson, Ericson, & Scott, 2001; Freyman, Helfer, McCall, & Clifton, 1999). This second type of masking is known as informational masking (IM). EM and IM are associated with different stages of processing: While EM is the result of limitations in early-stage peripheral processing, IM is thought to be due to subsequent breakdowns in central processing including selective attention, working memory, and the linguistic processing of speech sounds. The amount of EM produced by a given target–masker combination can be predicted based on the spectrotemporal overlap of the two sounds; however, predicting the effects of IM can be more difficult and often involves complex functions such as the utilization of a priori knowledge and expectation (for a recent review of IM in speech recognition, see Kidd & Colburn, 2017).

In general, high levels of IM for the task of speech recognition are produced under listening conditions where maskers consist of other intelligible talkers that can distract or confuse the listener or that can make it for difficult for the listener to piece together the "glimpses" of target speech available in masked conditions (e.g., Kidd et al., 2016, p. 134). If the target is a stream of intelligible speech and the masker consists of steady-state noise, a listener is unlikely to encounter difficulty perceptually segregating these two sources. However, in cases where the target and masker both consist of intelligible speech, the sources may not be as easily separated, resulting in higher levels of IM, manifest in some cases by explicit confusions of masker words for target words (e.g., Arbogast et al., 2002; Brungart et al., 2001; Kidd, Arbogast, Mason, & Gallun, 2005). Even within speech-on-speech masking conditions, the level of meaningfulness or comprehensibility of the masker is often important. For example, a time-reversed speech masker produces less interference than a forward speech masker (Freyman, Balakrishnan, & Helfer, 2001; Kidd, Mason, Best, & Marrone, 2010; Marrone, Mason, & Kidd, 2008b). Similarly, a masker spoken in a language unknown to the listener produces

less uncertainty about the target source than a masker spoken in a known language (Calandruccio, Brouwer, Van Engen, Dhar, & Bradlow, 2013; Van Engen & Bradlow, 2007). In both of these comparison cases, EM remains relatively constant, meaning that these decreases in confusion/uncertainty result in a release from IM. However, this issue is complex, and the extent to which IM is due to linguistic factors—such as the semantic content of the masker—is currently an active topic of research. For example, it is clear that conformance to an expected syntax is important for maintaining the segregation/focus of attention on a stream of target words (Kidd, Mason, & Best, 2014). However, the role of semantic relations among words, including the extent to which the semantic strength of unattended maskers influences performance, is currently unclear (cf., Brouwer, Van Engen, Calandruccio, & Bradlow, 2012; Calandruccio, Buss, Bencheck, & Jett, 2018). Moreover, time-reversed speech maskers may produce large amounts of IM even though explicit word confusions do not occur (Kidd et al., 2016, 2019). In addition to linguistic factors such as syntax and semantics, lower level segregation cues such as talker differences (e.g., fundamental frequency) or spatial separation of target and masker can provide a release from IM (e.g., Kidd et al., 2016) while combining these cues may further reduce IM (Rennies, Best, Roverud, & Kidd, 2019). However, although the aforementioned factors can help predict the amount of IM that will be produced by a given target–masker combination, susceptibility to IM has still been observed to vary substantially from person to person (e.g., Clayton et al., 2016; Rennies et al., 2019; Swaminathan et al., 2015).

Motivation for Investigating EM and IM in Aphasia

An important motivation for this study is the documented existence of auditory selective attention deficits in PWA. While aphasia is defined by the presence of impaired language processing, PWA have also been shown to exhibit impaired attention abilities relative to HC (for reviews, see Kurland, 2011; Murray, 2012; Villard & Kiran, 2017). This finding is of particular interest when considering the cocktail party problem in PWA, as the task of identifying and processing target speech in the presence of distractions is essentially a task of selective attention (e.g., Broadbent, 1952). Selective attention may be particularly important under high-IM masking conditions: Individual differences in selective attention abilities have been found to be predictive of individual differences in the ability to attend to a target talker in a multitalker environment, even among healthy/normal-hearing listeners (Clayton et al., 2016; Oberfeld & Kloeckner-Nowotny, 2016).

Although impaired performance by PWA has been observed on a broad range of tasks spanning many attentional types and modalities, there is evidence that auditory selective attention may be particularly affected. An early study of selective attention in aphasia compared the performance by PWA and controls on both visual and auditory tasks and found that, although PWA exhibited more errors and slower performance than controls on all tasks, their performance was poorest on the auditory selective attention tasks (Kreindler & Fradis, 1968). Later studies have confirmed that extraneous auditory information is highly distracting to PWA even in the context of very simple tasks. One study used a nonlinguistic auditory target identification task in which participants were asked to listen for and identify a harmonic complex (Erickson, Goldinger, & LaPointe, 1996). When presented with this task under quiet listening conditions, the performance of PWA was equivalent to that of the control participants; however, when the target was interspersed with nontarget pure tones, the performance of PWA declined relative to the controls. Another study found that while PWA were able to achieve high performance on auditory semantic judgment and lexical decision tasks with no distractions present, the addition of competing auditory information caused the performance of PWA to decrease to a greater extent than control performance (Murray, Holland, & Beeson, 1997). In a similar vein, yet another study asked PWA and control participants to listen to a sentence and judge its "syntactic correctness" (Murray, 2018). Again, PWA were able to achieve high accuracy on this task when presented in isolation, but accuracy decreased significantly when sentence-length pure-tone distractors were superimposed on the sentences. Furthermore, although it has been established that PWA perform more poorly than controls even on nonlinguistic attention tasks, there is also evidence that PWA performance may decline to a greater extent when language processing demands are added (Hula, McNeil, & Sung, 2007; Kreindler & Fradis, 1968; Murray et al., 1997; Villard & Kiran, 2018).

Given these findings of poorer than normal performance on auditory selective attention tasks and further decreases in performance when language demands are involved, it is unsurprising that PWA typically report difficulty with the task of understanding speech in noisy environments, arguably one of the most common auditory selective attention tasks encountered in everyday life. However, despite this accumulation of evidence and intuition based on common experience, most studies of selective listening in aphasia have been limited to consideration of auditory attention in a broad sense without focusing on the issues that relate specifically to speech recognition under masked conditions. To our knowledge, only two previous studies have sought to examine

the ability of PWA to selectively process target speech under masked listening conditions. One early study compared the performance of PWA and age- and hearing-matched controls on a task requiring attention to speech in the presence of steady-state noise maskers, that is, a condition presumably involving high EM but relatively little IM, and found that PWA performance was poorer than that of controls (Winchester & Hartman, 1955). A more recent study examined the ability of PWA to identify and selectively attend to target speech while ignoring different types of competing auditory information, including both speech maskers and noise maskers (Rankin et al., 2014). The results from this second study indicated that PWA demonstrated a poorer ability to receptively process speech under masked conditions than did controls of similar age and hearing status and that this difference was present when either speech or noise maskers were used.

One limitation of the existing work on auditory selective attention and auditory masking in PWA is that, to the best of our knowledge, the role of spatial separation of sources—an important segregation cue in multiple-talker "cocktail party" listening environments (e.g., Kidd & Colburn, 2017)—has not been examined when presenting multiple auditory stimuli. Most natural communication takes place when target and masker sources are spatially separated (i.e., primarily separation in azimuth). For example, when listening to speech in a crowded room, the listener often is situated facing the target talker—so that vision may be directed toward the target—and masker talkers are thus typically located at different points toward the right or left of the target. This spatial separation of target and masker results in the availability of critical binaural cues, in the form of interaural time and level differences, that listeners can take advantage of to aid in the perceptual segregation of target and masker(s) (Dirks & Wilson, 1969; Freyman et al., 1999; Hawley, Litovsky, & Colburn, 1999; Hirsh, 1950). Separation of sources therefore is not only a natural component of many everyday listening environments, but it is key to solving multiple-source listening tasks. For these reasons, an important question in examining the cocktail party problem in aphasia is whether PWA are able to take advantage of spatial cues to reduce the effects of IM. On a related note, for individuals with unilateral brain damage, spatial separation along the horizontal plane may also interact with pathological unilateral neglect, which in some cases may result in partial or total auditory extinction of stimuli presented on the side of space contralateral to the affected hemisphere. It is not uncommon for individuals with right-hemisphere brain lesions to exhibit "left neglect" or reduced attention to stimuli presented on the left side of space; this neglect may include reduced processing of auditory stimuli presented on the left (Carlyon, Cusack, Foxton, &

Robertson, 2001). The converse of this phenomenon—right neglect in individuals with left-hemisphere damage and aphasia—is less common (Beis et al., 2004). However, despite this, a number of studies have noted some degree of decreased attention to stimuli presented on the right in PWA (Bouma & Ansink, 1988; Ihori, Kashiwagi, & Kashiwagi, 2015; Marshall, Basilakos, & Love-Myers, 2013; Petry, Crosson, Rothi, Bauer, & Schauer, 1994; Shisler, 2005).

In addition to the potential relevance of impaired attention in PWA to solving the cocktail party problem and the interest in using a paradigm that exploits the natural perceptual benefit of spatial separation of sources, another motivation for this study stems from the existing work on susceptibility to IM in other populations. Previous studies on auditory masking have demonstrated that identifying how EM and IM interact within a given task is key to understanding their respective contributions to breakdowns in listener performance, particularly when the target consists of intelligible speech (cf., Kidd & Colburn, 2017). Several studies have shown that the listener's cognitive-linguistic competence or maturity may interact with the type(s) of masking present concluding that, in particular, impaired or less developed cognitive-linguistic competence may be associated with increased susceptibility to IM. There is evidence from the developmental literature, for example, that children are more susceptible than adults to the effects of IM (Corbin, Bonino, Buss, & Leibold, 2016; Fallon, Trehub, & Schneider, 2000; Hall, Grose, Buss, & Dev, 2002; Leibold & Buss, 2013; Wightman & Kistler, 2005). Other work has revealed that individuals listening to target speech in a second, nonnative language may be more susceptible to IM than when the target speech is presented in their native language (Calandruccio, Van Engen, Dhar, & Bradlow, 2010; Kilman, Zekveld, Hällgren, & Rönnberg, 2014). In addition, on a more clinical note, there is evidence that military service members and veterans with a history of high-intensity blast exposure demonstrate impaired central auditory processing capabilities, including, for some individuals, increased difficulty recognizing speech in noise (Gallun et al., 2016). Such findings highlight the importance of the cognitive-linguistic capability of the listener in the susceptibility to IM under speech-on-speech masking conditions. However, the extent to which this observation may apply to PWA—another population with decreased cognitive-linguistic capabilities—is not yet known.

Ideal Time–Frequency Segregation

Because one goal of this study was to look specifically at IM in PWA, and because most speech masking conditions comprise some combination of EM and IM, it was

important to employ a methodology that allowed EM and IM to be determined separately. Although the effects of IM for a given target–masker combination often vary from listener to listener, these individual differences may be estimated through the use of a processing technique termed "ideal time–frequency segregation" (ITFS). This technique was first developed for use in computational auditory scene analysis (e.g., Cooke, 2006; Li & Loizou, 2007; Wang, 2005) and was subsequently adapted for quantifying IM in masked speech recognition tasks in human listeners by Brungart, Chang, Simpson, and Wang (2006). Application of ITFS entails analyzing individual time–frequency (T-F) units of the combined auditory signal (i.e., target speech and masker(s)) reaching the ear of a listener, determining the relative energy of the target and masker (signal-to-noise ratio [SNR]) in each T-F unit, and then identifying those T-F units in which the target energy exceeds a criterion SNR. The computation of SNR in each T-F unit assumes a priori knowledge of target and masker waveforms. T-F units that fail to meet a prespecified SNR (termed the level criterion [LC]) are removed, and the stimulus is then reconstructed, resulting in an ITFS-processed signal consisting only of the target-dominated glimpses of the original combined signal. The purpose of applying ITFS is to produce a version of the signal that emulates not only the effects of EM but also is free—or very nearly free—of IM. Listener performance (e.g., target-to-masker ratio [TMR] at threshold) for an ITFS-processed signal can then be subtracted from the same listener's performance when presented with the original signal to determine the amount of additional masking due to IM. It should be noted that ITFS relies on a number of assumptions, for example, that the values chosen (extent in frequency and time) for the T-F units are relatively consistent with the internal resolution of the human auditory system as well as that the masker energy in the retained units does not result in significant IM. The specific LC selected may influence the degree to which this second assumption is satisfied: For example, although an LC of 0 dB is often used, the LC may be set at any level, and this value could affect the estimates of EM and IM (cf., Brungart et al., 2006). Despite these caveats, however, ITFS is considered to be the best approach currently available for obtaining a reasonable estimate of IM in a speech-on-speech masking task for a wide range of target–masker combinations (e.g., Rennies et al., 2019).

The use of ITFS has yielded findings that support the conclusion that reductions in masking due to unintelligible and/or spatially separated maskers are almost entirely attributable, in many cases, to reductions in IM (e.g., Rennies et al., 2019). One study by Kidd et al. (2016) used ITFS to examine reductions in masking due to time reversal of masker speech and spatial

separation between target and masker (as well as the effect of different-sex talkers for target and masker) in normal-hearing listeners. Their findings suggested that, indeed, release from masking in all speech masker conditions was due primarily to decreased IM. A follow-up study examining EM versus IM in listeners with HL found that, while listeners with HL were more susceptible than normal-hearing listeners to both EM and IM, manipulation of the masker conditions listed earlier provided less benefit to listeners with HL (Kidd et al., 2019). Interestingly, when ITFS processing was applied, listeners with HL obtained smaller improvements in performance than normal-hearing listeners, a result that they speculated could be due to reduced audibility of the remaining glimpses in the ITFS-processed signal or to a reduction in the ability to form coherent streams of speech from the degraded representations of the processed stimuli. These findings suggest that listeners with HL may be less able to take advantage of contextual/situational cues that would normally provide a release from IM and that, on a practical level, the presence of HL complicates the determination of the extent of IM and EM in speech masking conditions. Their findings also serve to underscore the importance of separating the effects of HL from the effects of aphasia when examining auditory masking in PWA.

Motivation, Aims, and Key Considerations of This Study

The goal of this study was to assess the effect of acquired aphasia on the ability of a listener to selectively attend to target speech in complex acoustic environments. In particular, we aimed to distinguish between the effects of EM and IM by systematically manipulating masker characteristics while holding other experimental parameters—including target talker characteristics, syntactic characteristics, and spatial separation of sources—constant. The experiment described in the following sections of this article therefore utilized an approach thought to produce large amounts of IM (a matrix-style forced-choice speech identification procedure; cf., Brungart 2001; Kidd, Best, & Mason, 2008; Kollmeier et al., 2015) with the methods adapted for use with PWA. In addition, spatial separation between target and masker sources was incorporated into the design and implemented using head-related impulse responses (IRs) measured in our laboratory from the Knowles Electronic Manikin for Acoustic Research (KEMAR) manikin (e.g., Kidd et al., 2016). This approach provided listeners with access to important binaural cues similar to those present in many naturalistic listening environments and which are thought to be key in solving complex, multiple-source listening. Finally, in order to separate the effects of aphasia from the effects of age

and HL, we recruited a group of age-matched HC and compensated for the effects of HL for all participants by providing frequency-dependent linear gain according to an established hearing aid algorithm (National Acoustics Laboratories-Revised Profound, Byrne, Parkinson, & Newall, 1991).

Gaining a clearer understanding of the contribution of acquired cognitive-linguistic deficits to the ability of PWA to understand target speech under challenging listening conditions not only has the potential to add to the existing knowledge base on selective auditory attention in PWA but may also help to identify barriers to typical/everyday social interactions. Findings from this work could have implications for new approaches to diagnosis and rehabilitative and compensatory treatment of communication function in this population. The specific objectives of this study were as follows:

1. The first objective was to determine whether PWA experience greater difficulty than age-matched HC with speech recognition in listening situations dominated by IM. To accomplish this objective, we compared the masking observed on a receptive speech processing task for PWA and HC using two types of maskers thought to differ in the extent to which they produce EM and IM. Specifically, the two masker types consisted of (a) two concurrent, spatially separated (from the target and each other) speech-spectrum-shaped Gaussian noises that mimicked the broadband spectrotemporal properties of the speech maskers (high EM) and (b) two spatially separated, concurrent intelligible talkers uttering speech similar to the target speech (high IM). These masked conditions were tested while compensating for any HL that was present in either PWA or HC on an individual participant basis. The underlying hypothesis was that PWA would demonstrate higher susceptibility to IM, but not EM, than controls, as evidenced by higher masked thresholds for target speech (expressed as TMR in dB as discussed further later) in the speech masking condition. A secondary, related hypothesis was that the error patterns for some PWA would reflect spatial biases related to the hemisphere of their acquired brain lesion.
2. The second objective was to measure the effect of removing IM while emulating the effects of EM through the use of ITFS, in PWA and in age-matched HC. We hypothesized that the removal of IM would result in improved performance for both groups. We also sought to determine whether PWA would be able to successfully integrate the sparse glimpses of the target speech that remain after ITFS processing and that represent the reduced target information available in masked speech conditions.

Methods

The experimental task used in this study consisted of a closed-set, forced-choice speech identification paradigm that was specifically adapted for use with PWA. This approach has been used frequently in past studies involving individuals with normal hearing or with sensorineural HL (e.g., Kidd et al., 2008, 2016, 2019). A typical trial consists of the auditory presentation of a target sentence, followed by the visual display on a computer monitor of a series of columns of written words, at which point the participant would be expected to select the target words they had heard. This trial structure poses comprehension, verbal working memory, and reading demands, which, while likely negligible for unimpaired individuals, could create serious confounds for the interpretation of PWA results. Therefore, the procedure used in the studies cited earlier was modified for use with PWA so that a smaller closed set of simple auditory stimuli was used, with a corresponding set of pictures as response options. Our modified speech identification task required good visual perception as well as the ability to semantically map a spoken word to a picture (within a consistent four-item closed set); however, it removed many of the other demands often present in standard speech identification tasks. This task was used for both PWA and HC participants. Participants were required to demonstrate ceiling-level performance on the task in quiet before beginning the full set of conditions.

Participants

Twelve PWA participants (five females, mean age 60.8 years, range: 48–74) and 12 HC participants (four females, mean age 61.4 years, range: 49–70) completed the experiment. PWA participants were referred to us by the Aphasia Research Laboratory at Boston University. HC participants were recruited from the Boston community through word of mouth, posted flyers, and online recruitment postings. All PWA exhibited decreased language abilities as the result of a unilateral cerebral infarction or hemorrhage that had occurred at least 12 months prior to participation in this study. No participants reported history of dementia, Parkinson's, or traumatic brain injury. All participants demonstrated adequate vision for performing the experimental task (see Table 1 for additional demographic information about participants). All procedures were approved by the institutional review board at Boston University.

Hearing Testing

Immediately following the consent process, participants underwent a pure-tone audiometric hearing test including the following frequencies tested separately in the right and left ears: 250 Hz, 500 Hz, 1 kHz, 2 kHz, 3 kHz,

Table 1. Demographic and Audiological Information for All Participants.

PWA	Age	Sex	Handedness (premorbid for PWA)	4F-PTA (left ear)	4F-PTA (right ear)	3HF-PTA (left ear)	3HF-PTA (right ear)	SRT
PWA1	53	M	R	25.0	17.5	65.0	46.7	24.3
PWA2	56	M	R	16.3	12.5	28.3	28.3	25.7
PWA3	54	M	R	15.0	12.5	47.5	39.2	23.5
PWA4	61	F	R	9.4	8.8	25.0	15.0	16.7
PWA5	56	F	R	15.6	13.8	33.3	20.0	18.7
PWA6	74	F	L	23.8	26.3	45.0	53.3	28.0
PWA7	62	M	L	45.0	48.8	60.0	85.0	27.3
PWA8	65	M	L	10.6	11.3	45.0	28.3	15.5
PWA9	67	M	L	32.5	32.5	68.3	71.7	21.7
PWA10	64	F	R	9.4	7.5	16.7	11.7	18.8
PWA11	70	F	R	13.8	15.0	41.7	38.3	27.0
PWA12	48	M	R	9.4	10.0	21.7	23.3	17.7
Mean:	60.8			18.8	18.0	41.5	38.4	22.1
HC1	62	M	R	20.0	17.5	46.7	51.7	25.0
HC2	61	M	R	27.5	25.0	48.3	53.3	29.5
HC3	62	M	R	6.4	8.8	31.7	23.3	19.2
HC4	55	F	R	23.1	25.0	30.0	41.7	26.7
HC5	60	F	L	7.5	6.3	25.0	20.0	21.2
HC6	69	F	R	10.6	11.3	38.3	30.0	22.7
HC7	67	M	R	12.5	12.5	38.3	33.3	17.3
HC8	49	F	R	10.6	11.3	10.0	10.0	16.7
HC9	63	M	L	11.3	8.8	36.7	20.0	17.0
HC10	60	M	R	5.0	5.0	25.0	23.3	18.0
HC11	70	M	R	36.3	38.8	86.7	90.0	29.0
HC12	59	M	R	4.4	1.3	21.7	20.0	15.8
Mean:	61.4			14.6	14.3	36.5	34.7	21.5

Note. 4F-PTA = four-frequency pure-tone average hearing threshold in dB HL, for 500 Hz, 1 kHz, 2 kHz, and 4 kHz; 3HF-PTA = three high-frequency pure-tone average hearing threshold in dB HL, for 4 kHz, 6 kHz, and 8 kHz; SRT = speech reception threshold for experimental sentences in dB SPL (post-gain), based on the average estimated thresholds from two quiet adaptive tracks; PWA = persons with aphasia; HC = healthy control; M = male; F = female; L = left; R = right.

4 kHz, 6 kHz, and 8 kHz. One participant (PWA7) reported that he had undergone several ear surgeries in the past and owned hearing aids that he sometimes used. This individual participated in the hearing test and experiment without the use of hearing aids. No other participants reported any current use of hearing aids.

All participants demonstrated some degree of sensorineural HL; for most, this loss was limited to the higher frequencies (see Table 1 for four-frequency pure-tone averages [4F-PTAs, calculated from 500 Hz, 1 kHz, 2 kHz, and 4 kHz] as well as three high-frequency pure-tone averages [3HF-PTAs, calculated from 4 kHz, 6 kHz, and 8 kHz] in each ear for all participants). Both the mean 4F-PTAs and the 3HF-PTAs in the PWA group were slightly poorer than the respective averages in the HC group (approximately 4 dB poorer in both cases). However, the result of an independent samples *t* test comparing 4F-PTAs (averaged between the left and right ears) between groups was nonsignificant, as was a similar independent samples *t* test comparing 3HF-PTAs (also averaged between the left and right ears) between groups. Based on these comparisons, we concluded that hearing profiles were similar between the two groups.

Linguistic and Cognitive Testing

PWA underwent a battery of standardized tests in order to determine their aphasia types and better understand their cognitive-linguistic profiles.[1] For participants in this group, the presence of aphasia was confirmed through language testing along with the clinical judgment of the first author, a certified speech-language pathologist. Two language tests were administered to all PWA participants: Part 1 of the Western Aphasia Battery-Revised ([WAB-R] Kertesz, 2007) and the Boston Naming Test ([BNT] Goodglass, Kaplan, & Weintraub, 1983). WAB-R results indicated that four PWA exhibited Broca's aphasia, while the remaining eight exhibited Anomic aphasia. In addition, to assess cognitive abilities including attention, the Cognitive-Linguistic Quick Test ([CLQT]

Table 2. Stroke and Standardized Testing Information for PWA.

PWA	Hemisphere of cerebral lesion	MPO	Aphasia type	WAB-R AQ	WAB-R AC	BNT	CLQT Composite	CLQT Attention	TEA Auditory Elevator Counting With Distraction	TEA Map Search (2 min)
PWA1	Left	119	Anomic	0.96	0.98	0.95	0.90	0.90	0.80	44
PWA2	Left	>200[a]	Broca's	0.59	0.68	0.62	0.55	0.78	0.20	21
PWA3	Left	170	Broca's	0.63	0.62	0.90	0.80	0.87	0.90	74
PWA4	Left	81	Anomic	0.96	1.00	1.00	1.00	0.96	1.00	56
PWA5	Left	110	Anomic	0.98	0.95	1.00	1.00	0.95	0.20	35
PWA6	Left	138	Broca's	0.36	0.55	0.12	0.55	0.69	0.40	14
PWA7	Right	219	Anomic	0.84	0.98	0.80	0.95	0.87	0.60	44
PWA8	Left	31	Anomic	0.94	0.82	0.93	0.85	0.88	0.30	45
PWA9	Left	42	Anomic	0.90	0.89	0.87	0.85	0.91	0.60	23
PWA10	Left	18	Anomic	0.98	1.00	0.98	1.00	1.00	1.00	55
PWA11	Left	162	Broca's	0.58	0.75	0.08	0.91	0.80	0.30	35
PWA12	Left	47	Anomic	0.92	0.89	0.92	1.00	0.97	1.00	41

Note. Scores for standardized tests are reported as the fraction of points earned out of a total of 1.00, with the exception of the TEA Map Search, for which raw scores are reported. MPO = months post onset; WAB-R = Western Aphasia Battery; AQ = Aphasia Quotient; AC = Auditory Comprehension; BNT = Boston Naming Test; CLQT = Cognitive-Linguistic Quick Test; TEA = Test of Everyday Attention; PWA = persons with aphasia.
[a]More precise information unavailable.

Helm-Estabrooks, 2001) was administered, along with the following three subtests of the Test of Everyday Attention ([TEA]; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994): Map Search (1 and 2 min), Elevator Counting, and Elevator Counting with Distraction. For the two TEA auditory elevator tasks, in which individuals were asked to listen to a series of tones and then indicate the number of tones played, a visual number line was presented during the response period for each trial so that the participant had the option of pointing to their response rather than verbalizing it (see Table 2 for information about PWA participants' scores on the aforementioned measures).

Experimental Stimuli

Auditory stimuli. The experimental auditory stimuli included speech tokens and noise tokens. The speech tokens consisted of laboratory-produced audio recordings of 12 individually spoken single words drawn from the list of tokens comprising a small, closed-set speech testing corpus (the American English version of the Oldenburg matrix sentence test; cf., Kollmeier et al., 2015; see Table 3), each one spoken by each of eight different female talkers, for a total of 96 recordings. The recordings were produced by Sensimetric Corporation (Malden, MA). The four words in a given category were selected for their phonemic distinguishability from one another; additionally, the four objects were also selected for their ease of imageability and their roughly equal semantic distance from one another. Each word was produced individually with neutral inflection and subsequently concatenated into sentences following the procedures described by Kidd et al. (2008) and used

Table 3. Experimental Matrix.

Subject	Verb	Object
Nina	**Wants**	Chairs
Kathy	Gives	Rings
Lucy	Has	Spoons
Rachel	Sees	Toys

Note. Target subject and verb indicated in boldface.

in several other studies employing matrix-based closed-set speech identification testing (e.g., Best et al., 2013; Kidd et al., 2014, 2016, 2019).

Visual stimuli. The visual stimuli used in the experiment consisted of four laboratory-created black-and-white digital line drawings, adapted from open-source images available on the Internet. Each image depicted one of the four objects (*chairs*, *spoons*, *rings*, and *toys* see Figure 1). Although plural forms of the words were presented auditorily, only one exemplar of a chair, spoon, and ring was pictured in order to simplify the available options and minimize visual clutter. No participants reported, or appeared to encounter, any difficulty matching plural spoken words to single-object images. All visual stimuli were presented on a computer screen via a graphical user interface.

Experimental Task Conditions and Processing of Stimuli

The experimental listening task included three conditions: a speech masking condition, a noise masking
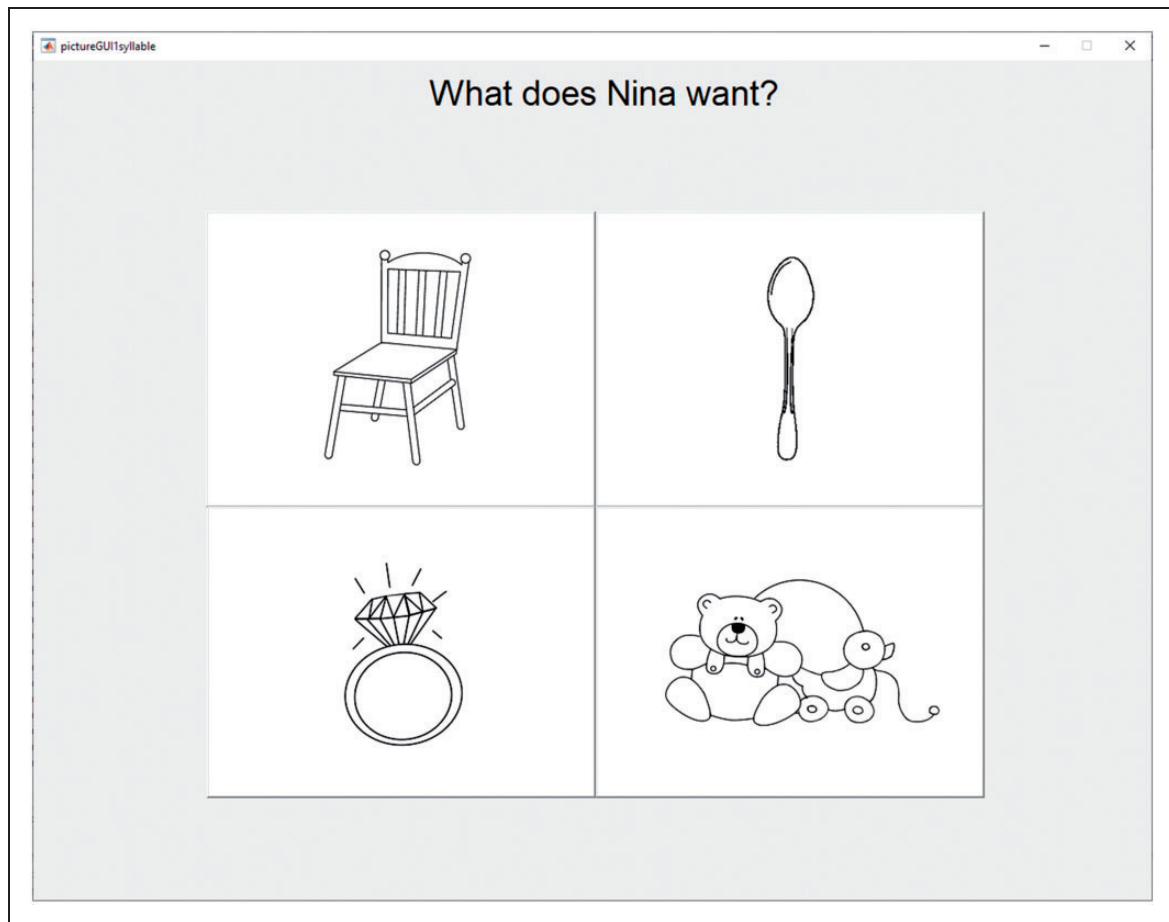
Figure 1. Visual response options provided to participants following presentation of auditory stimuli for each trial.

condition, and a glimpsed speech condition. The speech masking condition was intended to produce EM as well as high IM, the noise masking condition was intended to produce high EM with low IM, and the glimpsed speech condition was intended to retain the effects of EM while eliminating IM (cf., Rennies et al., 2019). Across conditions, participants were instructed to attend to a three-word target sentence drawn from the experimental matrix. In addition, two maskers were presented simultaneously with the target. Stimuli were convolved with IRs recorded in our laboratory using the KEMAR manikin situated at a distance of 5 feet from a loudspeaker array. IRs for the following source positions, which remained constant throughout the experiment, were used: 0° azimuth (straight ahead) for the target and ±45° azimuth for the two maskers, that is, 45° to the right and left of the target in the horizontal plane. The digital waveforms were D/A converted through an RME HDSP 9632 Audio Stream Input/Output (ASIO) 24-bit sound card (Audio AG, Haimhausen, Germany). A 44100 Hz sampling frequency was used. All signal processing and experimental control were implemented via MATLAB (MathWorks, Inc., Natick, MA).

Speech masking condition. In the speech masking condition, maskers consisted of two sentences drawn from the same matrix as the target. For each trial in this condition, three different talkers were randomly selected; one was designated as the target talker, while the other two were designated as the masker talkers. A subject-verb-object target sentence beginning with "Nina wants" and ending in a randomly selected object (e.g., "Nina wants spoons") was presented by concatenating the corresponding single-word audio files spoken by the target talker. Two masker subject-verb-object sentences, each one containing single words randomly chosen from the remaining subjects, verb, and objects (e.g., "Kathy sees rings"; "Rachel gives toys"), were also constructed. Each of these sentences was spoken by one of the masker talkers, again by concatenating the corresponding single-word audio files. All three sentences were presented synchronously, with the onsets of each successive triplet (subjects, verbs, and objects) temporally aligned.

Noise masking condition. In the noise masking condition, maskers consisted of two sentence-length tokens of speech-shaped, speech-envelope-modulated noise. For

each trial in this condition, procedures were identical to those in the speech masking condition, except that for each of the individual masker words, a noise file with the same long-term average spectral shape as the overall speech corpus and the same duration as the chosen masker word was created. This file was then modulated according to the broadband amplitude envelope of the chosen masker word and presented in its place. The result, for each of the two maskers, was a series of three noise samples temporally aligned with the three target words. The noise tokens were designed to provide similar spectral and envelope information as the speech tokens but were unintelligible and unrecognizable as speech.

Glimpsed speech condition. For the glimpsed speech condition, three different talkers and three different sentences again were selected on every trial, as described for the speech masking condition earlier. However, ITFS processing, using an LC of 0 dB, was then applied to each triplet of sentences, for a given ear. This meant that, for example, ITFS was applied to the combined signal from "Nina wants chairs" (originating from 0° azimuth), "Kathy gives rings" (originating from −45° azimuth), and "Rachel sees spoons" (originating from +45° azimuth) reaching the left ear; similarly, ITFS was applied to the combined signal from the same three sentences reaching the right ear. ITFS processing involved dividing the combined signal into a matrix of T-F units (also referred to as "tiles"), such that 128 frequency channels spanning 80 to 8000 Hz were analyzed, with 20-ms windows (sequential windows overlapping by 10 ms; cf., Brungart et al., 2006; Kidd et al., 2016). Each T-F unit in the matrix was assigned a value of either 1 or 0, where 1 indicated that the target energy was equal to or greater than the total masker energy in that T-F unit and 0 indicated that the masker energy exceeded the target energy in that T-F unit. Subsequently, all of the tiles in the combined signal that were designated 1 were retained, and all the tiles designated 0 were removed. The remaining target-dominated tiles were reassembled as the glimpsed target. The resulting "glimpsed" files were then presented to the appropriate ears. To match the long-term average spectrum of the glimpsed stimuli, we applied a high-pass 6th-order Butterworth filter at 80 Hz and a low-pass 14th order Butterworth filter at 8100 Hz to all nonglimpsed stimuli.

Talker/word randomization and processing were completed online during both the speech masking and noise masking conditions. However, because of the time needed for ITFS processing, 41,000 sets of glimpsed stimuli (1,000 for each TMR from −40 dB to 30 dB) were pregenerated and stored for playback prior to the experiment. For each trial during the glimpsed speech condition, one of the 1,000 files matching the current TMR was chosen at random and presented to the listener.

Frequency-Specific Gain

Across all three conditions, the last step of processing before presentation of the audio stimuli consisted of application of frequency-specific level gain, which was individualized for each participant's left and right ears based on their pure-tone hearing test results. The gain procedure of National Acoustics Laboratories-Revised Profound (Byrne et al., 1991) was used to create a linear filter for each participant which was applied to all stimuli throughout the experiment. While the application of gain does not fully compensate for HL, it provides a degree of amplification that takes into consideration both audibility and loudness. Hereafter, all intensity levels listed should be read as the level stated plus the individual participant's gain (e.g., "60 dB SPL" should be read as "60 dB SPL pregain").

Experimental Procedures

Experimental task setup. During the experimental task, participants were seated in a double-walled sound-treated Industrial Acoustic Corporation (North Aurora, IL) booth[2] in front of a computer monitor. Auditory stimuli were presented through Sennheiser HD280 Pro headphones (Sennheiser, Inc., Wedemark, Germany). Participants used a mouse to navigate through the experiment (e.g., to click "Start" or "Continue") as well as to indicate their responses. The only exception to this occurred with PWA6, who demonstrated substantial difficulty manipulating the mouse; this participant used a pointer to indicate their chosen response on the screen, and the experimenter clicked that response using the mouse. The mouse was configured for right-handed use by default; however, participants who were left-handed or who used their left hand due to right-sided hemiparesis were offered a choice of whether they would prefer to have the mouse reconfigured for left-handed use (no participants selected to have the mouse reconfigured, in most cases citing a familiarity with using the right-handed configuration).

Practice trials. Before beginning the experiment, each participant was first required to complete 10 practice trials in quiet. For each of the 10 practice trials, the procedures described earlier for the speech masking condition were implemented, except that the speech stimuli were presented in isolation, with no maskers present. All practice sentences were presented at 60 dB SPL. Following the auditory presentation of each sentence, a graphical user interface displaying the four response options appeared, and the participant was instructed to click the picture corresponding to "what Nina wants." To move forward

in the study, all participants were required to achieve 100% accuracy on 10 of 10 practice trials. If necessary, participants were reinstructed and practice trials were readministered until 100% accuracy was achieved. All participants were able to achieve 100% accuracy on a practice run within two attempts.

Determining speech reception threshold. Next, each participant completed two runs of an adaptive track designed to determine their speech reception threshold (SRT) for the stimuli used in the experiment. Like the practice trials described earlier, the SRT adaptive track consisted of sentences presented in isolation, with no maskers present. The first sentence in the SRT test was presented at a level of 70 dB SPL. The intensity levels of subsequent sentences were varied adaptively, according to a one-up, one-down procedure that estimates the 50% correct point on the psychometric function (Levitt, 1971; note that chance performance is 25% correct for the one-in-four forced-choice task). After each correct response, the level decreased by the specified step size, while after each incorrect response the level increased. Each point at which the direction of change reversed from decrease to increase (or vice versa) was termed a reversal. The step size, or amount by which the level was increased or decreased, began at 6 dB and switched to 2 dB after the third reversal. The adaptive track ended after nine reversals; the intensity levels at which the last six reversals occurred were averaged to determine the SRT. Each participant completed two adaptive tracks, and the SRTs for these two tracks were averaged to determine an overall SRT for each participant (see Table 1).

Determining uncomfortable loudness level. Next, an uncomfortable loudness level (UCL) was determined for each participant. It was explained to the participant that we would like to set a maximum loudness level for the experiment based on their comfort level and that we would like for them to listen to the noise at increasing loudness levels until they felt that they did not wish to hear anything louder. To determine each participant's UCL, a sequence of two simultaneous sentence-length tokens of speech-shaped, speech-modulated noise, with source positions of −45° and +45° azimuth, was used. The first trial in the sequence played each noise token at 68 dB SPL. The participant was instructed that if they wished, they could click a button on the screen to hear a sound that was a bit louder, and that they should let the experimenter know when the sounds were approaching the upper edge of their comfort level. The levels played during this track, in order, were as follows: 68, 72, 76, 78, 80, 82, 84, 86, and 88 dB SPL. When the participant indicated that they did not want to listen to any louder sounds, the track was discontinued, and the level of the last sound played was determined to be that participant's

UCL. If the participant did not choose to stop the track, the track was discontinued at 88 dB SPL, and this level was set as the UCL for the main experiment.

Masker familiarization and instructions. Following determination of SRTs and UCLs, and prior to beginning the experimental task, participants listened to two examples of trials in each of the three conditions; this was intended to familiarize them with what they would be listening to in the subsequent experiment. Following these examples, participants were given instructions for the experiment. Participants were reminded "to always listen to what Nina wants" and to ignore anything else that they heard. Also, if they were not sure of the correct answer, they should take their best guess. They were advised that the experiment was self-paced, that they should focus on getting each trial correct if possible, and that response time was not important. At no point were participants given any explicit information about the source positions of the target or maskers.

Experimental runs and adaptive tracks. Each participant completed five experimental runs, each consisting of three adaptive tracks, one in each experimental condition. The presentation level of the target sentence was held constant at 60 dB SPL throughout the experiment (although ITFS processing in the glimpsed speech condition frequently resulted in stimuli that were presented at a lower overall level postprocessing). For the first trial of a given adaptive track, each masker was presented at 30 dB SPL, corresponding to a TMR of 30 dB. The masker level (and thus the TMR) for each subsequent trial was then varied adaptively, according to a one-up, one-down procedure: After each correct response, the TMR decreased by a given step size, while after each incorrect response it increased. As during the unmasked SRT track, the step size, or amount by which the TMR was increased or decreased, began at 6 dB and switched to 2 dB after the third reversal. However, in cases where an adaptive track would have resulted in the presentation of maskers at a level above the listener's individual UCL, the track instead presented the maskers at the UCL. Similarly, in cases where an adaptive track would have played a set of glimpsed stimuli at a TMR below −40 dB (the lowest TMR of the pregenerated stimuli), the track instead used a TMR of −40 dB. In addition, no TMR ever exceeded 32 dB. The adaptive track ended after nine reversals; TMRs at the last six reversals were then averaged to determine a threshold estimate for that condition. The order of conditions was counterbalanced, both across the five runs and across participants. Participants were encouraged to take breaks between runs as needed and were permitted to complete the five runs either during a single study visit

or spread across two study visits, depending on their preference and level of fatigue.

Data Analysis

Calculation of Threshold Estimates

The five threshold estimates collected for each participant in each condition were averaged to produce an overall threshold estimate expressed in dB TMR (see Table 4). Because learning effects can occur during listening tasks, a standard deviation (*SD*) was also calculated to assess the variability across the five thresholds estimated within a given condition. If the *SD* exceeded 10 for any condition, the tracks were visually inspected. Only one *SD*, that of HC5 on the speech masking condition (*SD* = 18), met this criterion. In this case, visual inspection of the adaptive tracks showed that the speech masker threshold obtained during the first run varied substantially from those obtained during the subsequent runs. Data from the first run were therefore dropped for this participant, and they were asked to complete a sixth run. Data from Runs 2 through 6 (*SD* = 6.7) were used in all subsequent analyses; the mean from these estimates for HC5 is presented in Table 4. In addition, each participant's overall threshold for the glimpsed speech condition was subtracted from their overall threshold for the speech masking condition, resulting in a value representing the "additional masking" (which presumably provides an estimate of IM; cf., Brungart et al., 2006; Kidd et al., 2016) introduced by the speech masking condition (see Table 4).

Due to the predetermined ranges of TMRs available for use to each participant (based on their UCL as well as the closet set of pregenerated glimpsed stimuli), participants were sometimes presented with stimuli at a TMR that was higher than the adaptive track would otherwise have produced. The frequency at which TMRs were prevented from decreasing due to these constraints was examined, and, while most participants' tracks show that this happened at least once on at least one condition, it did not appear to be an overwhelming pattern that occurred across estimates for any participant. We speculate that some of these incidences may be attributable to the fact that the chance of getting any single trial correct from random guessing was 25%; therefore, it is likely that most participants occasionally found themselves at a TMR below their actual threshold due to a series of accidental correct guesses.

Right Versus Left Error Analysis

Finally, masker errors within the speech masking condition were computed on a trial-by-trial basis and summed across all five runs for each participant. A "masker

Table 4. Average Threshold Estimates and Additional Masking Levels for All Participants.

PWA	Speech masking threshold	Noise masking threshold	Glimpsed speech threshold	Additional masking (speech— glimpsed)
PWA1	−17.2	−14.9	−30.4	13.2
PWA2	9.9	−13.4	−24.0	33.9
PWA3	4.8	−19.2	−30.5	35.3
PWA4	−14.0	−13.9	−31.3	17.3
PWA5	−10.9	−15.7	−33.5	22.7
PWA6	−6.3	−12.1	−18.9	12.7
PWA7	6.4	−14.6	−28.5	34.9
PWA8	5.1	−14.3	−29.5	34.6
PWA9	0.4	−13.5	−24.3	24.7
PWA10	−6.1	−19.4	−30.4	24.3
PWA11	−3.3	−14.9	−20.5	17.2
PWA12	1.7	−14.5	−28.7	30.4
Mean:	−2.4	−15.0	−27.5	25.1
HC1	−18.5	−16.2	−27.1	8.7
HC2	4.8	−13.0	−26.0	30.8
HC3	−14.3	−16.1	−26.7	12.5
HC4	−14.0	−17.2	−34.0	20.0
HC5	−18.1	−19.0	−30.9	12.7
HC6	7.1	−13.5	−29.5	36.6
HC7	−17.5	−19.3	−30.5	13.0
HC8	−22.2	−19.2	−30.1	7.9
HC9	−19.1	−17.4	−32.9	13.9
HC10	−19.1	−20.9	−31.2	12.1
HC11	−15.4	−13.5	−29.8	14.4
HC12	−21.2	−19.3	−30.9	9.7
Mean:	−14.0	−17.1	−30.0	16.0

Note. "Additional masking" refers to the additional masking thought to be due to informational masking, as calculated by subtracting the glimpsed speech threshold from the speech masker threshold. PWA = persons with aphasia; HC = healthy control.

error" meant that the response for a given target word was one of the words presented on that trial in either of the two masker strings (all three words were mutually exclusive on a given trial). When the substitution was from the left masker, the error was classified as a left-biased error, and when the substitution was from the right masker, the error was classified as a right-biased error. A χ^2 test was performed on the errors for each participant to determine whether or not a pattern of left- or right-biased errors occurred (see Table 5).

Results

Effects of Group and Condition

A 2 × 3 (Group × Masker Type) repeated measures analysis of variance was performed to examine the effect of

Table 5. Results of Left-Biased Versus Right-Biased Error Comparisons.

	Left-biased errors	Right-biased errors	Chi-squared (χ^2) p value	Ratio lower/higher
PWA1	7	12	.251	0.583
PWA2	16	10	.239	0.625
PWA3	30	11	**.003**	0.367
PWA4	11	14	.549	0.786
PWA5	10	13	.532	0.769
PWA6	21	9	.028	0.429
PWA7	22	15	.25	0.682
PWA8	1	36	**8.71 e−09**	0.028
PWA9	24	17	.274	0.708
PWA10	19	9	.059	0.474
PWA11	5	20	**.003**	0.250
PWA12	21	9	.028	0.429
HC1	11	12	.835	0.917
HC2	18	24	.355	0.750
HC3	22	13	.128	0.591
HC4	14	19	.384	0.737
HC5	16	5	.016	0.313
HC6	9	15	.221	0.600
HC7	17	7	.041	0.412
HC8	10	16	.239	0.625
HC9	15	7	.088	0.467
HC10	18	12	.273	0.667
HC11	20	15	.398	0.750
HC12	6	14	.074	0.429

Note. Significant results at corrected α level in boldface. Ratio lower/higher refers to the ratio of R-to-L or L-to-R errors, depending on which was less/more frequent for that participant. PWA = persons with aphasia; HC = healthy control.

group and condition on participants' overall thresholds. Because Mauchly's test of sphericity indicated that the assumption of sphericity was not met in this data set, Greenhouse–Geisser corrected results are reported here where applicable. Results indicated a significant main effect of group, $F(1, 22) = 10.17$, $p < .001$; a significant main effect of condition, $F(1.22, 26.74) = 98.44$, $p < .01$; and a significant group by condition interaction effect, $F(1.22, 26.74) = 6.59$, $p < .05$.

As a follow-up to this analysis of variance, three independent samples t tests comparing PWA versus HC thresholds, one for each condition, were performed. The α level for significance was adjusted, using a Bonferroni correction, to .017 (0.05 divided by 3). For the speech masking condition, a significant difference was observed between PWA (mean = −2.44, SD = 8.62) and HC (mean = −13.96, SD = 9.63), $t(22) = 3.09$, $p < .017$. For the noise masking condition, the difference between PWA (mean = −15.03, SD = 2.19) and HC (mean = −17.05, SD = 2.63) did not reach significance,

$t(22) = 2.04$, $p = .054$. Similarly, for the glimpsed speech condition, the difference between PWA (mean = −27.55, SD = 4.56) and HC (mean = −29.97, SD = 2.39) was nonsignificant, $t(22) = 1.63$, $p = .117$ (see also Figure 2 for group-level thresholds for each condition).

In addition, several follow-up Pearson correlations were performed to determine whether speech masking conditions were associated with 3HF-PTAs or with noise masking thresholds. Correlations between 3HF-PTAs and speech masking thresholds were found to be nonsignificant for both groups. A significant correlation was found between noise masking thresholds and speech masking thresholds for HC ($r = .761$, $p < .01$) but not for PWA.

Additional Masking

Next, an independent samples t test comparing levels of additional masking (natural speech threshold—glimpsed speech threshold) revealed a significant difference in levels of additional masking between PWA (mean = 25.11, SD = 8.64) and HC (mean = 16.02, SD = 8.91), $t(22) = 2.54$, $p < .05$ (see also Figure 3 for a comparison of additional masking between groups).

Right Versus Left Error Analysis

Because each group contained 12 participants, the α level for significance for the χ^2 error analysis tests within the speech masking condition was adjusted, using a Bonferroni correction, to .00417 (0.05 divided by 12). Results were found to be significant at this adjusted α level for only three participants, all in the PWA group. PWA3 exhibited a significantly higher proportion of left-biased errors; conversely, PWA8 and PWA 11 both exhibited significantly higher proportions of right-biased errors. These results are presented in Table 5. Note that the total number of errors varied from participant to participant because the adaptive tracking procedure used to estimate threshold is based on a constant number of reversals in the track and not on a fixed/constant number of trials. In addition, "neutral" errors, or errors that matched neither the left nor the right masker object, are not represented here, as such errors were assumed to be the result of random guessing.

Relationships Between Experimental Results and Cognitive-Linguistic Testing

Finally, for PWA participants, a correlation matrix was created to examine whether any associations existed between speech masking thresholds and scores on tests of language or cognition that might reasonably be expected to tap into processes involved in selective attention to target speech. Scores included in this analysis were the following: WAB-R Aphasia Quotient, WAB-
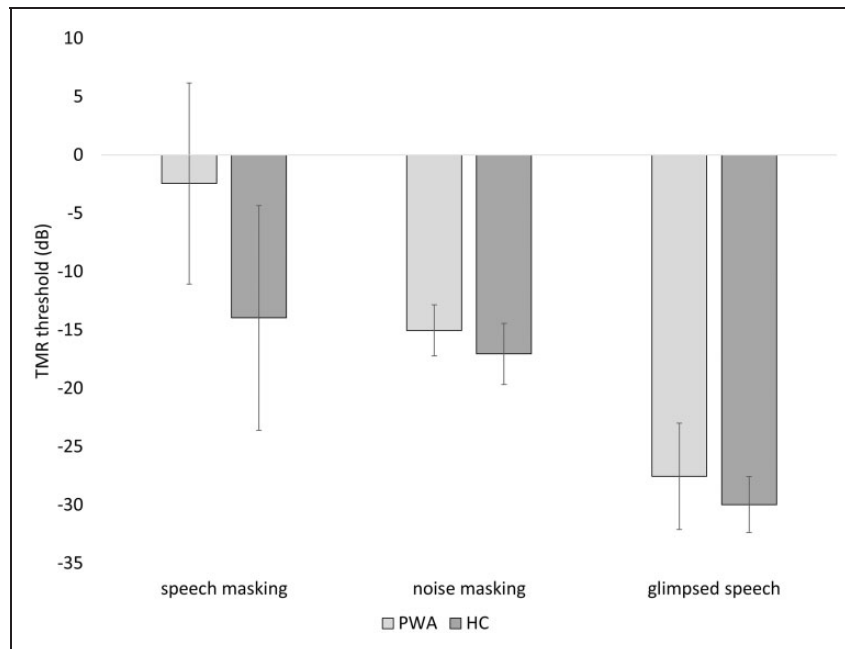
Figure 2. TMR thresholds for PWA and HC in each condition. Error bars indicate standard deviation. HC = healthy control; PWA = persons with aphasia; TMR = target-to-masker ratio.
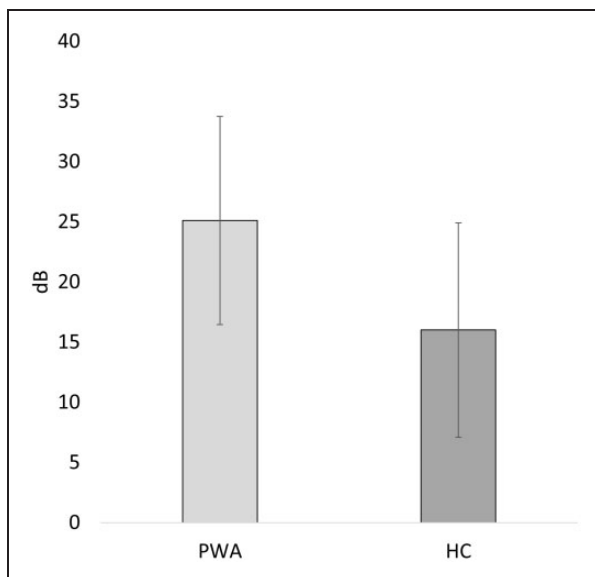

Figure 3. Additional masking for PWA and HC, computed by subtracting each participant's glimpsed speech threshold from their speech masking thresholds. Error bars indicate standard deviation. HC = healthy control; PWA = persons with aphasia.

R Auditory Comprehension, CLQT Attention, TEA-Elevator Counting with Distraction (a test of selective auditory attention), and TEA Map Search (2 min; a test of selective visual attention). The α level for this analysis was adjusted to .01 to correct for multiple comparisons. No significant correlations were noted. A similar

correlation matrix was created to examine any possible associations between additional masking and the same tests of language or cognition. Again, the α level for this analysis was set at .01. No significant correlations were noted. We also examined whether aphasia type appeared to be related to speech masking thresholds or additional masking (i.e., whether the PWA with notably high speech masking thresholds or notably high levels of additional masking all exhibited a particular aphasia type); no clear relationship was noted.

Finally, experimental results from PWA participants were divided into two groups based on aphasia type—Broca's aphasia versus Anomic aphasia—and compared to determine whether there was any clear relationship between speech masking thresholds and aphasia type or between additional masking and aphasia type. The mean speech masking threshold was 1.3 dB for PWA with Broca's aphasia and −4.3 dB for PWA with Anomic aphasia; the mean amount of additional masking was 24.8 dB for PWA with Broca's aphasia and 25.3 dB for PWA with Anomic aphasia. Because the group sizes were unequal, nonparametric Mann–Whitney *U* tests were used for both comparisons. In both cases, results were found to be nonsignificant.

Discussion

This study examined the ability of PWA and age-matched HC to process speech under masked listening conditions, with the particular goal of better

understanding and comparing the relative contributions of EM and IM to speech recognition within and between these two groups. To this end, sentences from a small closed-set matrix were presented auditorily under speech masking, noise masking, and glimpsed speech conditions. Notable features of the experimental paradigm included spatial separation of target and masker sources as well as the application of individualized frequency-specific gain to help compensate for the effects of HL. The results suggested that PWA and HC were similarly susceptible to the effects of EM, but that PWA were more susceptible than HC to the effects of IM.

Increased Susceptibility to IM in PWA

Our interpretation of the results from the speech masking condition, in which masker sentences were highly similar to and confusable with target sentences, is that this condition produced high levels of IM while causing much lower levels of EM. This conclusion is based on the TMRs at threshold obtained in that condition relative to the thresholds in the glimpsed and noise masking conditions (cf., Kidd et al., 2019; Rennies et al., 2019). These results showed that thresholds in this condition were higher (worse) for PWA than for HC, suggesting that PWA speech processing abilities break down more than those of HC as levels of competing speech are increased, which in turn suggests that PWA are more susceptible to the effects of IM than are HC. The strongest support for this conclusion was provided by the finding that PWA demonstrated significantly higher levels of additional masking than HC, as evidenced by subtracting each participant's glimpsed speech threshold from their speech masker threshold. Calculating this difference in thresholds is equivalent, in theory, to subtracting performance on an EM-only condition from performance on an EM-plus-IM condition, leaving only the effects of IM. Additional masking is a particularly meaningful metric because it compares each participant to themselves, thereby controlling for any potential influence of individual differences in performance which may be quite large in IM-dominated speech recognition tasks (cf., Clayton et al., 2016; Kidd & Colburn, 2017; Swaminathan et al., 2015) or to factors that may affect the ability to hear and use glimpses of masked speech (cf., Kidd et al., 2019). The higher levels of additional masking, in conjunction with equivalent group mean thresholds for the glimpsed speech and noise masking conditions, is consistent with the conclusion that PWA are especially adversely affected by competing talkers in "cocktail party" listening situations.

Separating Cognitive-Linguistic Deficits From Age and HL

An important question in the interpretation of these results is whether HL—and, as a result, the ability to detect target energy—may have influenced the group differences. As has been discussed in the literature on masked listening in older hearing-impaired listeners, the effects of age and HL are easily confounded (Frisina & Frisina, 1997; Pichora-Fuller & Souza, 2003); however, there is evidence that these two attributes differentially contribute to listeners' speech processing abilities in nonideal environments (Gallun et al., 2013). Age was easily controlled for in this study by comparing PWA to a group of age-matched HC. HL can be somewhat more difficult to control for, as the hearing profiles of individual listeners can be difficult to match precisely between groups. However, several points lend support for the conclusion that HL was adequately controlled for in this study. To begin with, while both the 4F-PTAs and the 3HF-PTAs were, on average, slightly poorer for PWA than for HC, between-group differences for both of these threshold averages were nonsignificant, suggesting that overall the two groups were well matched for auditory sensitivity. The second indication that differences in audibility between the two groups were not a concern was that SRTs were nearly identical between the two groups. These SRTs, which were assessed using target sentences from the experimental task, were presented with individualized frequency-specific gain applied separately to each ear of all participants (both PWA and HC). The fact that there was no difference in SRTs between the two groups suggests that the gain was effective in ensuring that the stimuli were equally audible to both groups of participants. (Incidentally, the fact that the PWA were able to achieve SRTs equal to those demonstrated by the HC also underscores the important fact that PWA did not encounter difficulty *comprehending* the experimental sentences when they were presented with no masking present.) Finally, thresholds for the glimpsed speech condition did not differ significantly between groups. The glimpsed speech condition is the condition during which any audibility issues would be most likely to surface, due to the fewer remaining glimpses and lower overall levels for lower TMR trials (despite the fact that the target level was fixed at 60 dB SPL). The absence of a group difference in this condition therefore indicates that the individualized gain effectively prevented reductions in audibility of stimuli due to HL. Given these findings, we believe it is reasonable to conclude that the observed group differences resulted not from peripheral factors but rather from acquired cognitive-linguistic deficits within the PWA group.

Validation of Our Methods in Light of Results From Previous Studies

Another important issue to consider when interpreting the results of this study is the use of a simplified experimental paradigm that required participants to report only the final word of a three-word sentence and to choose their response from a field of only four pictured options. As discussed earlier, this method was implemented to avoid potential confounds related to pure comprehension, reading, scanning, or verbal working memory that would likely have arisen for PWA in the context of a more typical psychoacoustic experiment involving longer sentences to listen to, more words to report, or lists of written words to choose from. We believe that this study's four-alternative forced-choice paradigm effectively removed these potential confounds; however, it also necessarily resulted in a higher chance level (25%) for each trial and could conceivably have affected threshold estimates in other, unintended ways. Therefore, we compared the threshold estimates obtained in this study for the speech masking and glimpsed speech conditions to the threshold estimates reported by previous studies in our laboratory that used roughly similar signal processing procedures but presented participants with more typical response configurations. Kidd et al. (2016; see also Rennies et al., 2019, for similar findings) examined speech masking and glimpsed speech conditions in young, normal-hearing adults, using a spatially separated two-talker masker paradigm; participants were asked to select a target word for each of five words in a sentence, given a field of eight written response options for each word. More recently, Kidd et al. (2019) used the same paradigm to test young hearing-impaired listeners. Individualized gain was applied in Kidd et al.'s previous study, as in this study. The thresholds obtained in the two previous studies for the speech masking condition were −19.6 dB for normal-hearing listeners (Kidd et al., 2016) and −8.5 dB for hearing-impaired listeners (Kidd et al., 2019). In this study—where participants' hearing profiles spanned a range from normal to mild–moderate impairment—the HC speech masking threshold was −13.95 dB, roughly in the middle of the two earlier estimates. For the glimpsed speech condition, the estimates were even closer: Kidd et al. (2016) observed a −29.4 dB threshold for normal-hearing listeners, Kidd et al. (2019) observed a −24 dB threshold for hearing-impaired listeners, and this study's glimpsed speech threshold for HC was −29.97. Although some of the specifics of these studies were different (e.g., different age groups and different degrees of spatial separation of the maskers), the fact that these previously published thresholds are roughly similar to the thresholds obtained in this study is consistent with the conclusion that the simplified methods used here did not distort the threshold estimates in any essential way.

Possible Reasons for Impaired Performance in PWA

While results from this study suggest that PWA are more susceptible to the effects of IM than are HC and that this increased susceptibility is likely to be due to breakdowns associated with aphasia, identifying the precise nature of these breakdowns is more difficult. Even when target speech is sufficiently audible to a listener, successful processing of that speech in the face of irrelevant auditory information still entails several steps. First, the listener must segregate the target from the maskers—which in this case included identifying the voice of the target talker based on hearing the word "Nina"—and must keep these sources separate throughout the presentation of the sentence, possibly using binaural cues to focus spatial attention on the target while ignoring sounds from other directions. The listener must also focus attention on the source they have selected, comprehend the attended message, and hold the comprehended information in memory during response selection.

It is impossible to definitively conclude where in this process breakdowns may have occurred in this study for PWA. However, some possibilities may be cautiously eliminated. There is no indication, for example, that PWA had difficulty recognizing the word "Nina" used to designate the target. Similarly, it seems unlikely that breakdowns were due to difficulties with voice recognition. The literature on voice recognition in unilateral stroke patients suggests that, unlike individuals with right-hemisphere stroke, individuals with left-hemisphere stroke and aphasia are generally able to distinguish between different voices as well as controls (Lang, Kneidl, Hielscher-Fastabend, & Heckmann, 2009; Van Lancker & Canter, 1982). This point could be tempered somewhat by the fact that our study included several premorbidly left-handed PWA and brain organization in left-handed individuals is not well understood; however, it seems unlikely that difficulties in voice recognition could explain the observed group difference. Nor is it plausible that PWA had difficulty holding the correct response in mind during response selection during the speech masking condition, as their performance was comparable to that of HC on the noise masking and glimpsed speech conditions, and the response demands across all three conditions were identical. And because all PWA demonstrated ceiling-level comprehension of the target sentences in quiet, comprehension seems unlikely to have been an issue. The remaining explanations for poorer PWA performance during the speech masking condition are that PWA experienced either lapses in source segregation or difficulty selectively attending to target stimuli or some combination of the

two. We know that PWA were able to segregate the target as well as HC during the noise masking condition, and we also know that PWA were able to use the available target glimpses as well as HC during the glimpsed speech condition when there was no competing masker. This strongly suggests that the problem is related to the speech masker or to the interaction of the speech masker and the segregation/selective attention cues available.

Exhibiting a particular difficulty with source segregation or selective attention under speech-on-speech masking conditions could potentially be related to a number of different issues. One possible explanation is that ignoring intelligible speech requires a greater amount of processing resources than ignoring noise, and that PWA have insufficient processing resources available for this task (and/or are unable to allocate their available resources effectively). It is also worth noting that speech processing, as a task that unfolds dynamically over time, may have posed particular problems for PWA. A number of studies have provided evidence of impaired temporal processing of auditory information in PWA (Edwards & Auger, 1965; Fink, Churan, & Wittman, 2006; Oron, Szymaszek, & Szelag, 2015). PWA have also been shown to exhibit increased time-based fluctuations in performance on an attention processing task relative to controls (Villard & Kiran, 2018). It is therefore also possible that delayed temporal processing or time-based fluctuations in attention during this study could have influenced PWA performance on the speech masking condition. However, the data from this study are insufficient to resolve this issue.

Finally, the data collected in this study do not provide strong support for the hypothesis that error patterns for some PWA would reflect spatial biases related to the hemisphere of their stroke. We found that three PWA—PWA3, PWA9, and PWA11, all of whom had a history of left-hemisphere damage—showed a significant asymmetry in masker substitution errors. Surprisingly, though, only one of these three showed an error bias in the expected direction. Typically, individuals with unilateral brain damage are less attentive to stimuli on the side of space contralateral to the lesion (if they show any unilateral spatial deficits)—and, as a result, would be expected to be *more* attentive to stimuli on the side of space ipsilateral to the lesion, presumably causing a higher proportion of ipsilateral masker word substitutions. However, only PWA3 showed a bias toward masker words presented on the left, while PWA8 and PWA11 showed a bias toward masker words presented on the right. These results are difficult to interpret given that the majority of PWA showed no significant bias toward either right or left masker words, and considering the mixed finding with respect to ipsilateral/contralateral errors. Further study of this issue

seems warranted before strong conclusions can be reached.

Individual Differences in Performance

The next important question, given the heterogeneity of the PWA group, is whether performance on the experimental task could be explained by participants' scores on cognitive-linguistic measures. It is first worth noting, however, that while PWA generally demonstrate substantial person-to-person variability in performance on almost any linguistic or cognitive measure, in this particular study, similar levels of person-to-person variability were observed within the PWA and HC groups. Such variability in the HC group was not unexpected; it is well known that even unimpaired individuals demonstrate differing levels of susceptibility to IM (e.g., Clayton et al., 2016; Oberfeld & Kloeckner-Nowotny, 2016; Swaminathan et al., 2015). It may be the case, therefore, that variability in PWA performance in this study was not driven by factors related to stroke or cognitive-linguistic deficits but rather to underlying individual differences. The lack of associations between the experimental results and the results of standardized testing in PWA do not contradict this hypothesis, nor does the lack of an observed relationship between experimental results and aphasia type. However, it could also be the case that the standardized tests administered in this study simply did not capture the cognitive-linguistic abilities most relevant to receptive speech processing under masking conditions in PWA. In particular, we did not include any standardized testing relating to spatial auditory information. It could also be the case that our sample size failed to capture the existing relationships between experimental results and cognitive-linguistic performance.

It is particularly interesting to note that even among the PWA who exhibited a very mild aphasia (PWA1, PWA4, PWA5, and PWA10), substantial person-to-person variability in levels of additional masking was observed. This is somewhat unexpected, given that these participants' performances on standardized auditory comprehension measures—which include items substantially more challenging than the experimental sentences used in this study—in quiet were near or at ceiling. The existence of this variability suggests that even PWA with intact or near-intact comprehension, who are expected to function at a high level in everyday communicative situations, may still vary significantly in their susceptibility to the effects of IM in multitalker environments.

Finally, the wide range of thresholds observed during the speech masking condition, including some thresholds above 0 dB, introduces the possibility that not all participants were utilizing the same cues to differentiate

target from maskers. When SNRs are positive (i.e., when the target is higher in level than the maskers), listeners are more likely to be able to rely on a simple level segregation cue to attend to the target, rendering utilization of voice-based and binaural cues unnecessary. In this study, six PWA and two HC participants demonstrated positive thresholds on the speech masking condition. While it is not possible to ascertain which cues these participants actually were using, it is possible that they had difficulty using more complex segregation cues and were only able to reliably identify or attend to the target when the level of the target exceeded that of the masker. The practical effect of this level segregation cue would be to limit the higher thresholds potentially underestimating IM.

Implications and Future Directions

The finding that PWA exhibit more difficulty in a speech-on-speech masked listening task than similar-age controls—and, in particular, that this difficulty cannot be explained by HL or by comprehension deficits per se—may have important implications for understanding how PWA process target speech in everyday complex acoustic environments. It is well known that PWA are likely to have fewer social connections and engage less frequently in the community than their unimpaired same-age counterparts (Cruice, Worrall, & Hickson, 2006; Davidson, Howe, Worrall, Hickson, & Togher, 2008). It is often assumed that this decreased social participation is due to problems with expressive language or perhaps in some cases to impaired comprehension. However, our results emphasize the possibility that difficulty attending to target speech in the presence of background sounds, particularly other intelligible speech, could also function as a barrier to social engagement and community participation. It is widely known that age-related hearing impairment is associated with decreased community participation (e.g., Mick, Kawachi, & Lin, 2014; Weinstein & Ventry, 1982) and decreased quality of life (e.g., Dalton et al., 2003). The possibility that PWA—with or without HL—may face a similar situation is worth further investigation. In addition, at least one previous study on expressive language abilities in aphasia has also shown that PWA have more difficulty producing spoken language when distracting auditory information is present (Murray, 2000).

It is easy to understand how clinicians working with PWA might not specifically evaluate or address patients' difficulties segregating target speech from auditory maskers. Clinicians who work with PWA in the chronic stage often see these individuals in quiet settings, such as structured conversation groups or one-on-one therapy settings in university or outpatient environments, where background talkers are usually not present.

However, after exiting the clinic, PWA may engage in communication on a busy street, in a crowded store, or at a family reunion, and in these environments, their ability to process the speech of a conversational partner may decline in ways that their standardized testing results might not predict. Fortunately, there is an increasing awareness of the importance of audiological assessment in PWA (Silkes & Winterstein, 2017; Zhang et al., 2018), although standard audiometric evaluations do not include speech masking conditions (Jakien, Kampel, Stansell, & Gallun, 2017). We hope that the results of this study will further encourage clinicians to consider the degree to which the presence—and type—of background noise might impact speech processing in PWA.

Finally, although this study examined masked listening abilities in PWA in the chronic stage of recovery, this line of research could potentially be extended to examine the same skills in PWA in earlier stages of recovery. Just as PWA in the chronic stage encounter adverse listening conditions in many real-world environments in home and community settings, PWA in the acute and subacute stages of recovery are likely to spend time in hospitals and rehabilitation facilities, which often involve notable levels of background noise (e.g., Pope, Gallun, & Kampel, 2013). Examining the ability of PWA in these earlier stages of recovery to selectively attend to target speech could therefore inform our understanding of recovery patterns and could potentially have implications for language treatment approaches in these settings.

Conclusion

In this study, we examined the effects of both EM and IM on receptive speech processing in individuals with aphasia and in age-matched controls. The application of individualized frequency-specific gain allowed us to examine the effects of aphasia on performance without the confound of differing levels of audibility due to HL. Results suggest that aphasia—even, in some cases, mild aphasia—may result in difficulties separating target speech from masker speech that cannot be accounted for by age, HL, or pure comprehension deficits. Although further work is needed to identify at precisely what point PWA abilities falter, as well as which cognitive-linguistic abilities may be predictive of the degree of this impaired processing in individual PWA, these findings demonstrate that this is an important issue in PWA. Gaining a better understanding of this topic may lead to a better understanding of how cognitive-linguistic impairments come to bear on everyday communication in nonideal listening environments for PWA and may eventually contribute to the development of strategies to minimize these barriers.

ORCID iD

Sarah Villard https://orcid.org/0000-0003-1181-9970

Notes

1. In some cases, scores on some of these measures were obtained from other laboratories with the participant's written authorization. All scores were recent within the past 6 months.
2. One PWA participant, PWA6, could not be tested in the double-walled soundbooth due to accessibility limitations; this participant was instead tested in a larger, single-walled Industrial Acoustic Corporation booth.

References

Arbogast, T. L., Mason, C. R., & Kidd, G., Jr. (2002). The effect of spatial separation on informational and energetic masking of speech. *The Journal of the Acoustical Society of America*, *112*(5), 2086–2098. doi:10.1121/1.1510141

Beis, J. M., Keller, C., Morin, N., Bartolomeo, P., Bernati, T., Chokron, S., . . . Perennou, D. (2004). Right spatial neglect after left hemisphere stroke: Qualitative and quantitative study. *Neurology*, *63*(9), 1600–1605. doi:10.1212/01. WNL.0000142967.60579.32

Best, V., Thompson, E. R., Mason, C. R., & Kidd, G. (2013). An energetic limit on spatial release from masking. *Journal of the Association for Research in Otolaryngology*, *14*(4), 603–610. doi:10.1007/s10162-013-0392-1

Bouma, A., & Ansink, B. J. (1988). Different mechanisms of ipsilateral and contralateral ear extinction in aphasic patients. *Journal of Clinical and Experimental Neuropsychology*, *10*(6), 709–726. doi:10.1080/ 01688638808402809

Broadbent, D. E. (1952). Failures of attention in selective listening. *Journal of Experimental Psychology*, *44*(6), 428–433. doi:10.1037/h0057163

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: Early processing and selection of multi-talker speech. *Attention, Perception, & Psychophysics*, *77*(5), 1465–1487. doi: 10.1044/1059-0889(2013/12-0072)

Brouwer, S., Van Engen, K. J., Calandruccio, L., & Bradlow, A. R. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. *The Journal of the Acoustical Society of America*, *131*(2), 1449–1464. doi:10.1121/1.3675943

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, *109*(3), 1101–1109. doi:10.1121/1.1345696

Brungart, D. S., Chang, P. S., Simpson, B. D., & Wang, D. (2006). Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation. *The Journal of the Acoustical Society of America*, *120*(6), 4007–4018. doi:10.1121/1.2363929

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), 2527–2538. doi:10.1121/1.1408946

Byrne, D. J., Parkinson, A., and Newall, P. (1991). Modified hearing aid selection procedures for severe-profound hearing losses. In G. A. Studebaker, F. H. Bess, & L. B. Beck (Eds.), *The Vanderbilt hearing aid report II* (pp. 295–300). Parkton, MD: York Press.

Calandruccio, L., Brouwer, S., Van Engen, K. J., Dhar, S., & Bradlow, A. R. (2013). Masking release due to linguistic and phonetic dissimilarity between the target and masker speech. *American Journal of Audiology*, *22*(1), 157–164.

Calandruccio, L., Buss, E., Bencheck, P., & Jett, B. (2018). Does the semantic content or syntactic regularity of masker speech affect speech-on-speech recognition?. *The Journal of the Acoustical Society of America*, *144*(6), 3289–3302. doi:10.1121/1.5081679

Calandruccio, L., Van Engen, K., Dhar, S., & Bradlow, A. R. (2010). The effectiveness of clear speech as a masker. *Journal of Speech, Language, and Hearing Research*, *53*, 1458–1471. doi:10.1044/1092-4388(2010/09-0210)

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 115. doi:10.1037/0096-1523.27.1.115

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979. doi:10.1121/1.1907229

Clayton, K.K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A. D., & Kidd, G., Jr. (2016) Executive function, visual attention and the cocktail party problem in musicians and non-musicians. *PLoS One*, *11*(7), 1–17. doi:10.1371/ journal.pone.0157638

Cooke, M. (2006). A glimpsing model of speech perception in noise. *Journal of the Acoustical Society of America, 119*, 1562–1573. doi:10.1121/1.2166600

Corbin, N. E., Bonino, A. Y., Buss, E., & Leibold, L. J. (2016). Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. *Ear and Hearing, 37*(1), 55. doi:10.1097/AUD.0000000000000201

Cruice, M., Worrall, L., & Hickson, L. (2006). Quantifying aphasic people's social lives in the context of non-aphasic peers. *Aphasiology, 20*(12), 1210–1225. doi:10.1080/02687030600790136

Cruice, M., Worrall, L., Hickson, L., & Murison, R. (2003). Finding a focus for quality of life with aphasia: Social and emotional health, and psychological well-being. *Aphasiology, 17*(4), 333–353. doi:10.1080/02687030244000707

Dalton, D. S., Cruickshanks, K. J., Klein, B. E., Klein, R., Wiley, T. L., & Nondahl, D. M. (2003). The impact of hearing loss on quality of life in older adults. *The Gerontologist, 43*(5), 661–668. doi:10.1093/geront/43.5.661

Davidson, B., Howe, T., Worrall, L., Hickson, L., & Togher, L. (2008). Social participation for older people with aphasia: The impact of communication disability on friendships. *Topics in Stroke Rehabilitation, 15*(4), 325–340. doi:10.1310/tsr1504-325

Dirks, D. D., & Wilson, R. H. (1969). The effect of spatially separated sound sources on speech intelligibility. *Journal of Speech and Hearing Research, 12*(1), 5–38. doi:10.1044/jshr.1201.05

Edwards, A. E., & Auger, R. (1965). The effect of aphasia on the perception of precedence. In American Psychological Association (Ed.), *Proceedings of the annual convention of the American Psychological Association* (pp. 207–208). Washington, DC: American Psychological Association.

Ellis, C., & Urban, S. (2016). Age and aphasia: A review of presence, type, recovery and clinical outcomes. *Topics in Stroke Rehabilitation, 23*(6), 430–439. doi:10.1080/10749357.2016.1150412

Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., . . . Lyrer, P. A. (2006). Epidemiology of aphasia attributable to first ischemic stroke: Incidence, severity, fluency, etiology, and thrombolysis. *Stroke, 37*(6), 1379–1384. doi:10.1161/01.STR.0000221815.64093.8c

Erickson, R. J., Goldinger, S. D., & LaPointe, L. L. (1996). Auditory vigilance in aphasic individuals: Detecting nonlinguistic stimuli with full or divided attention. *Brain and Cognition, 30*(2), 244–253. doi:10.1006/brcg.1996.0016

Ezzatian, P., Li, L., Pichora-Fuller, K., & Schneider, B. A. (2015). Delayed stream segregation in older adults: More than just informational masking. *Ear and Hearing, 36*(4), 482–484. doi:10.1097/AUD.0000000000000139

Fallon, M., Trehub, S. E., & Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *Journal of the Acoustical Society of America, 108*(6), 3023–3029. doi:10.1121/1.1323233

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America, 88*(4), 1725–1736. doi:10.1121/1.400247

Fink, M., Churan, J., & Wittmann, M. (2006). Temporal processing and context dependency of phoneme discrimination in patients with aphasia. *Brain and Language, 98*(1), 1–11. doi:10.1016/j.bandl.2005.12.005

Formby, C., Phillips, D. E., & Thomas, R. G. (1987). Hearing loss among stroke patients. *Ear and Hearing, 8*(6), 326–332. doi:10.1097/00003446-198712000-00007

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America, 109*(5), 2112–2122. doi:10.1121/1.1354984

Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *The Journal of the Acoustical Society of America, 106*(6), 3578–3588. doi:10.1121/1.428211

Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hearing Research, 106*(1–2), 95–104. doi:10.1016/S0378-5955(97)00006-3

Gallun, F. J., Diedesch, A. C., Kampel, S. D., & Jakien, K. M. (2013). Independent impacts of age and hearing loss on spatial release in a complex auditory environment. *Frontiers in Neuroscience, 7*, 252. doi:10.3389/fnins.2013.00252

Gallun, F. J., Lewis, S. M., Folmer, R. L., Hutter, M., Papesh, M., Belding, H., & Leek, M. R. (2016). Chronic effects of exposure to high-intensity blasts: Results of tests of central auditory processing. *Journal of Rehabilitation Research & Development, 53*(6), 705–720. doi:10.1682/JRRD.2014.12.0313

Gifford, R. H., Bacon, S. P., & Williams, E. J. (2007). An examination of speech recognition in a modulated background and of forward masking in younger and older listeners. *Journal of Speech, Language, and Hearing Research, 50*, 857–864. doi:10.1044/1092-4388(2007/060)

Goodglass, H., Kaplan, E., & Weintraub, S. (1983). *Boston naming test*. Philadelphia, PA: Lea & Febiger.

Hall, J. W., III, Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing, 23*(2), 159–165. doi:10.1097/00003446-200204000-00008

Hawley, M. L., Litovsky, R. Y., & Colburn, H. S. (1999). Speech intelligibility and localization in a multi-source environment. *The Journal of the Acoustical Society of America, 105*(6), 3436–3448. doi:10.1121/1.424670

Helfer, K. S., Chevalier, J., & Freyman, R. L. (2010). Aging, spatial cues, and single-versus dual-task performance in competing speech perception. *The Journal of the Acoustical Society of America, 128*(6), 3625–3633. doi:10.1121/1.3502462

Helm-Estabrooks, N. (2001). *Cognitive linguistic quick test*. New York, NY: The Psychological Corporation.

Hilari, K., Needle, J. J., & Harrison, K. L. (2012). What are the important factors in health-related quality of life for people with aphasia? A systematic review. *Archives of Physical Medicine and Rehabilitation, 93*(1), S86–S95. doi:10.1016/j.apmr.2011.05.028

Hirsh, I. J. (1950). The relation between localization and intelligibility. *The Journal of the Acoustical Society of America*, *22*(2), 196–200. doi:10.1121/1.1906588

Hula, W. D., McNeil, M. R., & Sung, J. E. (2007). Is there an impairment of language-specific attentional processing in aphasia? *Brain and Language*, *103*(1–2), 240–241. doi:10.1016/j.bandl.2007.07.023

Ihori, N., Kashiwagi, A., & Kashiwagi, T. (2015). Right unilateral spatial neglect in aphasic patients. *Brain and Language*, *147*, 21–29. doi:10.1016/j.bandl.2015.05.001

Jakien, K. M., Kampel, S. D., Stansell, M. M., & Gallun, F. J. (2017). Validating a rapid, automated test of spatial release from masking. *American Journal of Audiology*, *26*(4), 507–518. doi:10.1044/2017_AJA-17-0013

Kertesz, A. (2007). *WAB-R: Western Aphasia Battery-Revised.* New York, NY: Grune & Stratton.

Kidd, G., Jr., Arbogast, T. L., Mason, C. R., & Gallun, F. J. (2005). The advantage of knowing where to listen. *The Journal of the Acoustical Society of America*, *118*(6), 3804–3815. doi:10.1121/1.2109187

Kidd, G., Jr., Best, V., & Mason, C. R. (2008). Listening to every other word: Examining the strength of linkage variables in forming streams of speech. *The Journal of the Acoustical Society of America*, *124*(6), 3793–3802. doi:10.1121/1.2998980

Kidd, G., Jr., & Colburn, H. S. (2017). Informational masking in speech recognition. In J. C. Middlebrooks, J. Z. Simon, A. N. Popper, & R. R. Fay (Eds.), *The auditory system at the cocktail party* (pp. 75–109). New York, NY: Springer. doi:10.1007/978-3-319-51662-2_4

Kidd, G., Jr., Mason, C. R., & Best, V. (2014). The role of syntax in maintaining the integrity of streams of speech. *The Journal of the Acoustical Society of America*, *135*(2), 766–777. doi:10.1121/1.4861354

Kidd, G., Jr., Mason, C. R., Best, V., & Marrone, N. (2010). Stimulus factors influencing spatial release from speech-on-speech masking. *The Journal of the Acoustical Society of America*, *128*(4), 1965–1978. doi:10.1121/1.3478781

Kidd, G., Jr., Mason, C. R., Best, V., Roverud, E., Swaminathan, J., Jennings, T., . . . Steven Colburn, H. (2019). Determining the energetic and informational components of speech-on-speech masking in listeners with sensorineural hearing loss. *The Journal of the Acoustical Society of America*, *145*(1), 440–457. doi:10.1121/1.5087555

Kidd, G., Jr., Mason, C. R., Swaminathan, J., Roverud, E., Clayton, K. K., & Best, V. (2016). Determining the energetic and informational components of speech-on-speech masking. *The Journal of the Acoustical Society of America*, *140*(1), 132–144. doi:10.1121/1.4954748

Kilman, L., Zekveld, A., Hällgren, M., & Rönnberg, J. (2014). The influence of non-native language proficiency on speech perception performance. *Frontiers in Psychology*, *5*, 651. doi:10.3389/fpsyg.2014.00651

Kollmeier, B., Warzybok, A., Hochmuth, S., Zokoll, M. A., Uslar, V., Brand, T., & Wagener, K. C. (2015). The multilingual matrix test: Principles, applications, and comparison across languages: A review. *International Journal of Audiology*, *54*, 3–16. doi:10.3109/14992027.2015.1020971

Kreindler, A., & Fradis, A. (1968). *Performances in aphasia: A neurodynamic diagnostic and psychological study.* Paris, France: Gauthier-Villars.

Kurland, J. (2011). The role that attention plays in language processing. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*, *21*(2), 47–54. doi:10.1044/nnsld21.2.47

Lang, C. J., Kneidl, O., Hielscher-Fastabend, M., & Heckmann, J. G. (2009). Voice recognition in aphasic and non-aphasic stroke patients. *Journal of Neurology*, *256*(8), 1303–1306. doi:10.1007/s00415-009-5118-2

Leibold, L. J., & Buss, E. (2013). Children's identification of consonants in a speech-shaped noise or a two-talker masker. *Journal of Speech, Language, and Hearing Research*, *56*, 1144–1155. doi:10.1044/1092-4388(2012/12-0011)

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477. doi: 10.1121/1.1912375

Li, N., & Loizou, P. (2007) Factors influencing glimpsing of speech in noise. *Journal of the Acoustical Society of America*, *122*, 1165–1172. doi:10.1121/1.2749454

Marrone, N., Mason, C. R., & Kidd, G., Jr. (2008a). The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *The Journal of the Acoustical Society of America*, *124*(5), 3064–3075. doi:10.1121/1.2980441

Marrone, N., Mason, C. R., & Kidd, G., Jr. (2008b). Tuning in the spatial dimension: Evidence from a masked speech identification task. *The Journal of the Acoustical Society of America*, *124*(2), 1146–1158. doi:10.1121/1.2945710

Marshall, R. S., Basilakos, A., & Love-Myers, K. (2013). Further evidence of auditory extinction in aphasia. *Journal of Speech, Language, and Hearing Research*, *56*(1), 236–249. doi:10.1044/1092-4388(2012/11-0191)

McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *Quarterly Journal of Experimental Psychology*, *58*, 22–33. doi:10.1080/02724980443000151

Mick, P., Kawachi, I., & Lin, F. R. (2014). The association between hearing loss and social isolation in older adults. *Otolaryngology—Head and Neck Surgery*, *150*(3), 378–384. doi:10.1177/0194599813518021

Middlebrooks, J. C., Simon, J. Z., Popper, A. N., & Fay, R. R. (2017). *The auditory system at the cocktail party*. New York, NY: Springer. doi:10.1007/978-3-319-51662-2

Murray, L. L. (2000). The effects of varying attentional demands on the word retrieval skills of adults with aphasia, right hemisphere brain damage, or no brain damage. *Brain and Language*, *72*(1), 40–72. doi:10.1006/brln.1999.2281

Murray, L. L. (2012). Attention and other cognitive deficits in aphasia: Presence and relation to language and communication measures. *American Journal of Speech-Language Pathology*, *21*, S51–S64. doi:10.1044/1058-0360(2012/11-0067)

Murray, L. L. (2018). Sentence processing in aphasia: An examination of material-specific and general cognitive

factors. *Journal of Neurolinguistics*, *48*, 26–46. doi:10.1016/j. jneuroling.2018.03.007

Murray, L. L., Holland, A. L., & Beeson, P. M. (1997). Auditory processing in individuals with mild aphasia: A study of resource allocation. *Journal of Speech, Language, and Hearing Research*, *40*(4), 792–808. doi:10.1044/jslhr.4004.792

Oberfeld, D., & Kloeckner-Nowotny, F. (2016). Individual differences in selective attention predict speech identification at a cocktail party. *Elife*, *5*, e16747. doi:10.7554/eLife.16747

Oron, A., Szymaszek, A., & Szelag, E. (2015). Temporal information processing as a basis for auditory comprehension: Clinical evidence from aphasic patients. *International Journal of Language & Communication Disorders*, *50*(5), 604–615. doi:10.1111/1460-6984.12160

Petry, M. C., Crosson, B., Rothi, L. J. G., Bauer, R. M., & Schauer, C. A. (1994). Selective attention and aphasia in adults: Preliminary findings. *Neuropsychologia*, *32*(11), 1397–1408. doi:10.1016/0028-3932(94)00072-7

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International Journal of Audiology*, *42*(sup2), 11–16. doi:10.3109/14992020309074638

Pope, D. S., Gallun, F. J., & Kampel, S. (2013). Effect of hospital noise on patients' ability to hear, understand, and recall speech. *Research in Nursing & Health*, *36*(3), 228–241. doi:10.1002/nur.21540

Rankin, E., Newton, C., Parker, A., & Bruce, C. (2014). Hearing loss and auditory processing ability in people with aphasia. *Aphasiology*, *28*(5), 576–595. doi:10.1080/02687038.2013.878452

Rennies, J., Best, V., Roverud, E., & Kidd, G., Jr. (2019). Energetic and informational components of speech-on-speech masking in binaural speech intelligibility and perceived listening effort. *Trends in Hearing*, *23*, 1–21. doi:10.1177/2331216519854597

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *The Test of Everyday Attention: TEA*. Bury St. Edmunds, England: Thames Valley Test Company.

Shisler, R. (2005). Aphasia and auditory extinction: Preliminary evidence of binding. *Aphasiology*, *19*(7), 633–650. doi:10.1080/02687030444000930

Silkes, J. P., & Winterstein, K. (2017). Speech-language pathologists' use of hearing screening for clients with aphasia: Challenges, potential solutions, and future directions. *American Journal of Speech-Language Pathology*, *26*(1), 11–28. doi:10.1044/2016_AJSLP-14-0181

Skelly, M. (1975). Rethinking stroke: Aphasic patients talk back. *The American Journal of Nursing*, *75*(7), 1140–1142. doi: 10.2307/3423493

Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Kidd, G., Jr., & Patel, A. D. (2015). Musical training, individual differences and the cocktail party problem. *Scientific Reports*, *5*, 11628. doi:10.1038/srep11628

Van Engen, K. J., & Bradlow, A. R. (2007). Sentence recognition in native-and foreign-language multi-talker background noise. *The Journal of the Acoustical Society of America*, *121*(1), 519–526. doi:10.1121/1.2400666

Van Lancker, D. R., & Canter, G. J. (1982). Impairment of voice and face recognition in patients with hemispheric damage. *Brain and Cognition*, *1*(2), 185–195. doi:10.1016/0278-2626(82)90016-1

Villard, S., & Kiran, S. (2017). To what extent does attention underlie language in aphasia? *Aphasiology*, *31*(10), 1226–1245. doi:10.1080/02687038.2016.1242711

Villard, S., & Kiran, S. (2018). Between-session and within-session intra-individual variability in attention in aphasia. *Neuropsychologia*, *109*, 95–106. doi:10.1016/j. neuropsychologia.2017.12.005

Wang, D. (2005). On ideal binary mask as the computational goal of auditory scene analysis. In P. Divenyi (Ed.), *Speech separation by humans and machines* (pp. 181–197). Boston, MA: Springer. doi:10.1007/0-387-22794-6_12

Weinstein, B. E., & Ventry, I. M. (1982). Hearing impairment and social isolation in the elderly. *Journal of Speech, Language, and Hearing Research*, *25*(4), 593–599. doi:10.1044/jshr.2504.593

Wightman, F. L., & Kistler, D. J. (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *The Journal of the Acoustical Society of America*, *118*(5), 3164–3176. doi:10.1121/1.2082567

Winchester, R. A., & Hartman, B. T. (1955). Auditory dedifferentiation in the dysphasic. *Journal of Speech and Hearing Disorders*, *20*(2), 178–182. doi:10.1044/jshd.2002.178

Yost, W. A. (1997). The cocktail party problem: Forty years later. In R. Gilkey & T. R. Anderson (Eds.), *Binaural and spatial hearing in real and virtual environments* (pp. 349–347). Mahwah, NJ: Lawrence Erlbaum Associates.

Zhang, M., Pratt, S. R., Doyle, P. J., McNeil, M. R., Durrant, J. D., Roxberg, J., & Ortmann, A. (2018). Audiological assessment of word recognition skills in persons with aphasia. *American Journal of Audiology*, *27*(1), 1–18. doi:10.1044/2017_AJA-17-0041