Measuring Inner Speech Objectively and Subjectively in Aphasia

Julianne M. Alexander[1]*, Peter Langland-Hassan[2], and Brielle C. Stark[1]

[1] Department of Speech, Language and Hearing Sciences and Program in Neuroscience, Indiana University, Bloomington, Indiana, USA

[2] Department of Philosophy, University of Cincinnati, Cincinnati, Ohio, USA

Abstract

Background

Many people with aphasia and people without brain injury talk to themselves in their heads, i.e., have "inner speech." Inner speech may be better than speech out loud for some people with aphasia and may a serve a variety of functions (e.g., emotion regulation), which motivates us to provide a high-fidelity characterization of it. Researchers have used multiple methods to measure this internal phenomenon in the past, which we combine here for the first time in a single study.

Aims

We compare performance on inner speech tasks that assess inner speech "in-the-moment" to general subjective impressions of inner speech abilities in individuals with and without aphasia and tease apart the relationship of aphasia severity to inner speech.

Methods & Procedures

Twenty people with mild-moderate aphasia and twenty neurotypical controls completed several inner speech tasks, including objective silent rhyme judgements (picture, written, and auditory), subjective reports of inner speech during naming, and subjective rating scales about inner speech experience more generally.

Outcomes & Results

Inner speech during silent rhyming tasks was associated with aphasia severity only for more complex tasks (e.g., picture and written rhyming but not auditory rhyming). In-the-moment inner speech reports during silent naming were not associated with aphasia severity, and nor were the subjective ratings about general inner speech experience. Individuals with and without aphasia demonstrated a variety of subjective general inner speech experiences, demonstrating heterogeneity of this phenomenon more broadly.

Conclusions

Methods of measuring inner speech complement each other and get at different facets of the inner speech phenomenon, and clinicians and researchers must carefully choose the method(s) that will provide the information about inner speech that they desire.

Keywords

Aphasia, inner speech, methods, severity

Introduction

Inner speech is the familiar experience of a "little voice in the head" (Perrone-Bertolotti et al., 2014). Research in neurotypical individuals and individuals with aphasia suggests that inner speech retains lexical and phonological properties (Oppenheim & Dell, 2008) and is associated with brain areas governing language production (Fama et al., 2017; Geva & Fernyhough, 2019) and with cognitive processes such as phonological working memory (Fama, Henderson, et al., 2019; Stark et al., 2017). Compellingly, we and others have shown that inner speech can be comparatively preserved alongside impoverished overt speech in individuals with aphasia (Fama et al., 2017; Geva, Bennett, et al., 2011; Stark et al., 2017), which has led to the hypothesis that inner speech has a consequential role in language recovery in aphasia (Fama & Turkeltaub, 2020; Stark et al., 2017).

Inner speech can be understood as having several components. One method of conceptualizing inner speech's multifaceted nature is the ConDialInt model, which describes three dimensions of inner speech: condensation, dialogality, and intentionality (Grandchamp et al., 2019). On the condensation dimension, inner speech ranges from condensed (without all acoustic, phonological, and syntactic information) to expanded (including articulatory and auditory properties) forms. Dialogality dimension refers to whether the inner speech captures just one voice/perspective, like a monologue or inner soliloquy, or whether it is dialogical (e.g., an inner verbal conversation in one's own voice or with the addition of other voices). For the intentionality dimension, inner speech can be purposeful, as when intentionally rehearsing information or manipulating it silently, or it can be unintentional/spontaneous, as in daydreaming or mind wandering (Grandchamp et al., 2019). Therefore, different methods for measuring inner speech also target different facets of inner speech.

Inner speech has been conceptualized as "inner ear" and "inner voice" (Hubbard, 2010; Smith et al., 1995), where the inner ear is a more passive process, involving just auditory information, with the "inner voice" being a more active process which includes articulatory information. Another group makes a similar distinction between "inner speaking" and "inner hearing" (Hurlburt et al., 2013), emphasizing the difference between whether a person is truly saying something in their head or if they are simply hearing it, as if played back on a recording. There are likely different underlying processes involved in inner voice/speaking versus inner ear/hearing (Hubbard, 2010). The "inner voice" is considered more language-dependent and complex than the "inner ear," and so might be more affected in people with more severe aphasia.

Finally, there are two definitions of inner speech most used in aphasiology, differing based on whether inner speech involves active manipulation of the word form in the head (as in Geva, Bennett, et al., 2011; Stark et al., 2017), or whether the interest is in access to the word form without the need for manipulation (as in Fama, Snider, et al., 2019). The two distinct definitions of inner speech involve different underlying cognitive processes. For example, inner speech involving manipulation has been measured objectively by silent rhyming decisions whereas inner speech involving only access and not manipulation has been measured mostly by subjective perception of whether inner speech is present. Both are legitimate means of evaluating inner speech, but require slightly different cognitive components, e.g., the former requiring more phonological working memory. Even within the objective measurements, cognitive complexity differs. For example, aural judgement (e.g., involving the "inner ear") is the cognitively simplest of these tasks, relying on phonological access and awareness and the phonological loop (Baddeley & Hitch, 2019). Written judgement (orthographic, e.g., involving the "inner voice") is cognitively more complex, relying on orthographic-to-phonological conversion, especially in instances where two words rhyme but have dissimilar orthography (e.g., bear, chair). Rhyme judgement using pictures, rather than silent reading of words and non-words (Geva & Warburton, 2019; Langland-Hassan et al., 2015) is generally thought to be the most difficult because it requires more cognitive steps (accessing visual-object information → accessing lexical-semantic information → accessing phonological information and retaining in working memory). A benefit is that these methods assure that the participants are completing the tasks asked of them (e.g., by responding via button press to a rhyme judgement) and because the experimenters better control the situation (e.g., by providing stimuli that range in orthographic match and by monitoring the subject to ensure that they are not moving their lips or tongue) (Geva, Jones, et al., 2011). Objective methods like rhyme judgment require the extra step of manipulation beyond access to the word, making it difficult to separate inner speech access from cognitive abilities such as working memory.

Notably, both methods (i.e., objective rhyming and subjective awareness of inner speech during naming) evaluate only inner speech in the moment, i.e., they isolate the experience of inner speech during a specific task (e.g., naming).

The in-the-moment evaluation of inner speech, described above, stands in contrast to evaluating inner speech as it occurs "in the wild"—that is, in the many unplanned and unprompted daily contexts in which it often occurs. Therefore, they are inherently limited in that they do not evaluate the function or presence of inner speech in its natural environment, nor are they sensitive to how inner speech may differ because of context. The general perception of inner speech, and its functions, have been measured by a variety of subjective surveys (Alderson-Day et al., 2018; Brinthaupt et al., 2009; Racy et al., 2020), descriptive experience sampling (Hurlburt et al., 2016) and thought listing (Racy et al., 2020), largely in populations without aphasia. It is natural to expect that the subjective perception of inner speech in general differs from inner speech as assessed in the moment during isolated tasks, given that research in non-brain-damaged adults suggests inner speech frequency ranges from 0-75% of sampled instances (Hurlburt et al., 2016); is used for a variety of functions, including problem-solving (Wallace et al., 2017), self-awareness (Morin, 2011b, 2011a) and emotion regulation (Morin & Michaud, 2007), and varies by context (e.g., may be more present during cognitively difficult scenarios) (Morin et al., 2018; Racy et al., 2020).

In this study, for what we believe is the first time, we analyse various methods for measuring the multifaceted inner speech variable in individuals with and without aphasia, and we specifically explore the extent to which aphasia severity impacts inner speech.

Research Questions

The current research evaluates three primary research questions:

1) Is performance on objective and subjective inner speech tasks mediated by aphasia severity?
2) Are cognitively less complex inner speech tasks related to aphasia severity differently than cognitively more complex tasks?
3) How do subjective measures and objective measures relate to one another?

If the ways of measuring inner speech do not relate to aphasia severity similarly, it is possible that they are measuring different facets or experiences of inner speech, as described in the ConDialInt model (Grandchamp et al., 2019). For example, participants may have differing perceptions of inner speech beyond the skills used to complete objective inner rhyme judgement tasks. The inner speech in these rhyme judgement tasks would be expanded, monologue, and intentional according to the ConDialInt model. In subjective ratings, participants may be drawing on experiences of inner speech that may be more condensed or less intentional. Additionally, if the methods that we use for measuring inner speech (objective vs subjective) are measuring similar inner speech facets, we would expect them to be similarly related to aphasia severity.

Method

Participants

Twenty-three participants with chronic post-stroke aphasia (PWA) were recruited by author PLH for a related study (details can be found in Langland-Hassan et al. (2021)), which was approved by the University of Cincinnati's ethical review board (IRB #2012-4185). The original study collected the Western Aphasia Battery-Revised (WAB-R) (Kertesz, 2006) to establish aphasia severity and type. To ensure adequate comprehension of experimental tasks, the original study included participants only if their auditory-verbal comprehension score on the WAB-R was ≥4 (out of 10). To provide a general overview of cognitive functioning, the original study also had PWA complete the non-linguistic subtests of the Cognitive Linguistic Quick Task (CLQT) (Helm, 2003), and the participants were subsequently classified as within normal limits, mildly impaired, moderately impaired, or severely impaired. If a participant obtained ratings of "within normal limits" and/or "mild" impairment on the non-linguistic subtests of the CLQT, this was taken as evidence of auditory and visual perception adequate to complete the experimental tasks and participants were included in the study.

N = 20 PWA were included in the following analyses as in Langland-Hassan et al. (2021), in which they excluded three of the original 23 PWA due to difficulty understanding task instructions. They were aged 35-72 years (M=56.4, SD=9.09), and all had chronic aphasia (1.5-31 years with aphasia, M=8.85±8.09). Four types of aphasia were represented: anomic (N=7), Broca's (N=6), conduction (N=6), and transcortical motor (N=1). Only one participant was classified as having severe aphasia according to the WAB AQ (an AQ score of <50), while all others were classified as mild-moderate (M=73.1±13.9). The average level of education was approximately equivalent to a Bachelor's degree (M=15.9±1.79 years).

As part of the original study, a sample of prospectively matched non-brain damaged individuals served as a control group. This group was at least 18 years of age, had no history of language impairments, had no history of major psychotic or neurologic episodes or intractable substance abuse, had at least a high school education, and were native speakers of American English. Twenty adults were included (7M/13F, M=55±9.4 years old [range 41-70], M=17±1.8 years of education) (Table 1).

Between the aphasia and control groups, there were no significant differences in age (t = -

0.42, p = 0.68), gender/sex ($\chi^2 = 0.38$, p = 0.54), or education (did not meet normality assumptions; W = 239.5, p = 0.27). Table 1 includes demographics for all included participants.

[TABLE 1 HERE]

Materials and Procedure

An inner speech battery, described below, was presented on a 21" touchscreen computer by means of a JavaScript application stored on a secure external web server. There are two major sections to the battery. The first section evaluates "in-the-moment inner speech," which is further divided into subjective and objective methods, whilst the second section evaluates how participants generally, subjectively interpret their inner speech.

Section 1: In-the-moment inner speech

Section 1 was further subdivided into Objective inner speech tasks (S1.1) and Subjective inner speech tasks (S1.2).

Section 1.1: Objective tasks.

The objective portion of the inner speech battery (see Figure 1 for a visual representation) included the following three subtests:

1. Picture Rhyme Judgement, in which the participants saw two pictures of objects and were asked if the words for them rhyme (as in Langland-Hassan et al. (2015)). This task is considered cognitively complex, as it requires access to semantic information via visual routes, then access to lexical and phonological information. It is commonly referred to as an "inner voice" task;
2. Written Rhyme Judgement, in which the participants saw two written words, some of which varied in orthography (e.g., Bear / Chair), and were asked if they rhyme. This task is considered cognitively complex, as it requires access to orthographic and phonological

information. It is commonly referred to as an “inner voice” task;

3. Auditory Rhyme Judgement, in which the participants heard two words and were asked if they rhyme. Participants could hear the words an unlimited number of times. This task is considered cognitively less complex than the other two rhyme tasks, as it requires access to only phonological representations and working memory. It is commonly referred to as an “inner ear” task.

The materials for the picture rhyming task were adapted from the materials of Geva & Warburton (2019) and were provided to author PLH by Geva. However, the sets used were slightly different from the published version of Geva & Warburton's materials (those materials were normed on British English-speaking adults) and so cannot be considered strictly equivalent.

The words used in each of the three inner rhyming tasks did not differ significantly in terms of their frequency, familiarity, or imageability, based on t-tests ($p > 0.3$ for all comparisons) of ratings from the online MRC Psycholinguistic Database (Coltheart, 1981). For all tasks, participants were instructed to give a yes/no response via button press on the touchscreen and to complete the task in silence and without moving their mouth or tongue. A research assistant monitored performance. There were 38 trials involving picture rhyming, 40 trials for written word rhyming, and 40 trials for auditory rhyme judgements with no time limits for each trial and no instructions about speed of answering.

[FIGURE 1 HERE]

Section 1.2: Subjective task: Naming Successful Inner Speech (SIS)

Participants also completed subjective reports of inner speech during a silent naming task in which they were shown one picture at a time and were asked to indicate whether they were able to say the word for the object silently in their head and have it sound right (“naming SIS”).

Again, they responded with a yes/no button press to the prompt. The subjective report of single word inner speech is considered a less complex cognitive task, as it requires only access to phonological information, but not manipulation of phonological information.

In order to compare inner speech to overt speech ability, participants completed overt (out loud) naming of the same items. They were shown one picture at a time and asked to name it out loud. The experimenter recorded their response in English orthography.

Section 2: General, subjective interpretation of inner speech

This section included four subjective rating questions, which were presented on the screen and read aloud by the experimenter. In contrast to the in-the-moment objective and subjective tasks above, these ratings were about participant's experience in general, not just during the task. The subjective rating questions used a visual rating scale (continuous arrow with numbers and descriptions) from 1-5 with verbal descriptions and are as follows:

1. <u>Idea without Word (IwW)</u>, "How often do you see something and know what it is but can't say the word for it out loud?" with 1 being almost never and 5 being almost always;
2. <u>Post-Stroke,</u> given to aphasia group only, "Do you think you talk to yourself in your head less now, more now, or about the same as before your stroke?" with 1 being much less now and 5 being much more now;
3. <u>Successful Inner Speech (SIS)</u>, "A few minutes ago we talked about how we sometimes see something and know what it is but can't say the word for it out loud. When that happens to you, how often can you say the right word for it in your head, and have it sound right?" with 1 being almost never and 5 being almost always. Note that two members of the control group were missing data for this question.

4. <u>Inner vs. Overt Speech</u>, "In general, how do you think your abilities to speak in your head compare to your abilities to speak out loud?" with 1 being much better in head and 5 being much better out loud.

Two of the subjective rating questions (IwW and SIS) were adapted from Fama's work (e.g., Fama et al., 2019, Fama & Turkeltaub, 2020). As noted above, the subjective responses were on a 5-point scale, which changed based on the question, and were about general abilities rather than subjective reports about specific single words. Therefore, they served as a complement to the subjective single word inner speech data also collected. Participants selected their answer on a touchscreen. PWA answered all four subjective questions, while the control group answered all but the post-stroke question.

To summarize, we distinguish the inner speech tasks in two ways: whether the method is objective or subjective, and whether the measurement is happening "in-the-moment" of inner speech or is a general interpretation of the overall inner speech experience.

Section Ordering

The subjective ratings described in Section 2 were always presented to participants first, and the order of the subjective questions was not randomized (they were presented in the order described above). Naming SIS (Section 1.2) came after the general subjective ratings from Section 2. Next, participants completed the rhyme judgement tasks (Section 1.1). The questions from Section 1.1 were presented in pseudo-randomized order by participant, alternating among six different orderings of the tasks. Overt naming was always last.

General Procedures

Participants with aphasia attended two sessions: a screening session in which they completed neuropsychological assessments (e.g., WAB-R), and a second session in which they completed the inner speech tasks and a categorization task relevant to a larger abstract thought study [which will not be discussed here and is available in another paper] (Langland-Hassan et al., 2021). Neurotypical controls attended only one session and completed the inner speech tasks and categorization task in that session.

Data Analysis

All data analysis was completed in R version 4.1.1 with RStudio Desktop 2022.07.1+554. The data analysis was designed to answer the following research questions:

1) How does inner speech relate to aphasia severity?
2) For the "in-the-moment" inner speech section tasks for the aphasia group only (Section 1), are cognitively less complex tasks related to aphasia severity differently than cognitively more complex tasks?
3) How does inner speech compare between persons with aphasia and neurotypical controls?

Neurotypical Controls compared to Participants with Aphasia

To ensure that the objective inner speech tasks were valid and to replicate findings from Geva & Warburton (2019) (i.e., that neurotypical adults could perform them with high accuracy), we will compute descriptive statistics for the objective inner speech tasks. We will also compare the neurotypical controls to the PWA on the three subjective inner speech rating scales that both groups completed (i.e., excluding the subjective post-stroke question).

Relationship between Aphasia Severity and Inner Speech

We will compute correlations between severity (as measured by the WAB-R Aphasia Quotient) and each of the objective measures (picture, auditory, and written). In addition to computing correlations with the continuous metric of aphasia severity, we will separate the aphasia group into two severity groups based on *a priori* severity groups identified by the WAB and analyse the extent to which inner speech scores differ by severity group. We will split the groups based on the Aphasia Quotient on the WAB, which has cut-off scores of "mild" being 76 and over (one person scored a 94, which could be considered "latent", i.e., non-clinical aphasia) and moderate being 51-75 (although we included one person who scored 47.8). We will compare the inner speech data between severity groups using the Wilcoxon rank sum test with continuity correction, given that the data did not satisfy normality assumptions (Meek et al., 2007). We had originally planned to also evaluate the impact of aphasia type (fluent vs. non-fluent) on inner speech, but we ultimately decided that due to the scoring criteria on the WAB (in which severity has a large bearing on type of aphasia), evaluating by type would not add sufficient additional information to our study and would be conflated with aphasia severity.

Relationship between Cognitive Complexity and Performance

We will use the correlations and comparisons from above along with hypothesized differences in cognitive complexity to analyse the relationships between the different tasks. We will compare performance on each of the Section 1.1 tasks with each other with Wilcoxon rank sum tests. Additionally, we will investigate the relationship between tasks from Section 1.2 (successful inner speech in-the-moment and overt naming). Finally, we will evaluate the relationship between Section 1.1 and Section 1.2 tasks, specifically the successful inner speech in-the-moment (section 1.2) and the objective picture rhyming task (section 1.1), because tasks require inner speech achieved through visual semantic access, then lexical and phonological access, but only picture rhyme judgment further requires the manipulation of information to make a judgment on rhyme, thus making it (in theory) cognitively more complex.

Relationship Between Different Methods of Measuring Inner Speech

We will compute Spearman's rank correlations to analyse the relationships between objective picture rhyme judgement and subjective ratings scales.

Results

Descriptive statistics and aphasia vs. control comparisons

Section 1 Tasks

Controls' data on objective tasks confirmed that the tasks were valid (i.e., performance was high/ at ceiling for all tasks) (see Table 2). As expected, the control group significantly outperformed the PWA group on all Section 1 tasks (all $p < .001$), except for Naming SIS, as both groups were near ceiling.

[TABLE 2 HERE]

Section 2 Tasks

For the subjective questions that asked more generally about inner speech that both groups answered, we saw similar answers for the controls and PWA (IwW p-value = 0.31, SIS p-value = 0.92), excepting that PWA more commonly reported that their inner speech was better or much better than their speech out loud, while the controls sometimes reported better speech out loud than in their heads ($W = 74$, p-value = 0.0003). Figure 2 represents the subjective answers from the PWA group and the control group, visually demonstrating that PWA and control groups only differ on their answers of whether they speak better in their head vs. out loud. Additionally, the PWA and control groups' answers vary considerably within groups.

[FIGURE 2 HERE]

Research Question 1: Evaluating the impact of aphasia severity on inner speech

People with milder aphasia (based on the WAB AQ) tended to perform better on the objective inner speech tasks (the rhyming tasks from Section 1.1), especially picture ($r^2 = .31$, $p = .01$) and written ($r^2 = .23$, $p = .03$) tasks, but not significantly different from individuals with more severe aphasia for the auditory task ($r^2 = .17$, $p = .06$). As WAB AQ increases (indicating milder aphasia severity), performance on the picture and written tasks also increases, but more of the participants at all aphasia severity levels are at or near ceiling for the auditory task (Figure 3).

[FIGURE 3 HERE]

When we grouped the PWA into Mild and Moderate aphasia severity groups based on their WAB AQ, the results were similar, with the mild group performing more accurately on the picture ($W = 89$, $p = 0.003$) and written ($W = 77$, $p = 0.04$) inner speech rhyme judgement tasks, but not on the auditory rhyme judgement task ($W = 71$, $p = 0.11$). As seen in Figure 4, this clinically relevant distinction appears related to inner speech abilities, as the mild group outperformed the moderate group on the tasks which required inner “voice” (picture, written), but not inner “ear” (auditory). Aphasia severity was not associated with Naming SIS reports, however ($r^2 = 0.02$, $p = 0.60$).

[FIGURE 4 HERE]

Aphasia severity groups answered similarly on all the general subjective questions about inner speech (Section 2) based on the Wilcoxon rank sum test with continuity correction (all $p>0.20$). Individuals from the different severity groups reported similar frequency of successful inner speech ($W = 48.5$, $p = 0.97$) and idea without word ($W = 37.5$, $p = 0.36$). Both participant groups reported using inner speech more often now or at the same frequency as they did before their stroke, with no significant differences between the groups ($W = 38$, $p = 0.36$). Both groups reported equal or better inner speech than overt speech, with no significant differences between the groups ($W = 33.5$, $p = 0.20$). As seen in Figure 5, subjective ratings were similar between the two severity groups, with intra- and inter-group variability.

[FIGURE 5 HERE]

Research Question 2: Relationship between cognitive complexity and performance (aphasia group only)

To examine the role of cognitive complexity, including its relationship with aphasia severity, we compared performance on the objective rhyme judgement tasks. For Section 1.1 objective rhyme tasks, PWA performed better on the auditory task (M = 0.76, SD = 0.17) than the written (M = 0.63, SD =0.17) and picture (M = 0.64, SD = 0.17) tasks, but there was no difference in performance between the written and picture tasks (Auditory vs picture: W = 279.5, p-value = 0.03; Auditory vs Written: W = 286, p-value = 0.02; Picture vs Written: W = 202.5, p-value = 0.96). It is also of interest how PWA do when we parse apart the residuals from each inner speech task, i.e., evaluate performance on picture inner speech when considering variance from written and auditory inner speech performances, and so on. For example, given that the picture task is thought to be the most cognitively complex due to the multiple cognitive components necessary to complete the task (e.g., visual access to semantic information, then lexical selection, then phonological selection), and comprises many of the same cognitive processes as the written and auditory, considering variance from the other objective tasks will enable us to evaluate performance more clearly for each task. When accounting for variance from each of the other objective tasks, a one-way ANOVA showed that PWA did not perform significantly differently on any task (three paired t-tests, all $p>0.99$). This suggests that, when evaluating each objective task as comprised of unique cognitive components, PWA performed similarly on audio, written, and picture inner speech tasks. The similar performance when accounting for variance suggests that the poor performance on picture and written rhyme tasks by the PWA group as a whole (no regression of other tasks) is due to the increased cognitive complexity of those tasks, rather than by their unique cognitive components.

When looking for group differences between mild and moderate severity, we also evaluated performance on the objective tasks in their regressed states (i.e., picture performance residuals when written and audio performance was regressed) and identified no significant differences for audio ($p=0.12$) or written ($p=0.36$) performance by severity group. We did identify a significant difference for picture performance (residuals) by severity group ($p=0.02$), which was no longer significant after Bonferroni correction (three comparisons, new significance $p<0.16$). This trend suggests that, even when controlling for variance from audio and written rhyme judgment, the picture rhyme judgment task was more difficult for those with more severe aphasia. The increased difficulty suggests that the unique cognitive components comprising this picture task, beyond those in the audio and written tasks, is influenced by aphasia severity. An example of this non-shared cognitive component may be lexical-semantic access via the visual stream.

We then compared all the tasks which would require lexical-semantic access with pictures, including overt picture naming (Section 1.2), silent naming with subjective reports of successful inner speech ("naming SIS" from Section 1.2), and picture rhyme judgement (from Section 1.1). Importantly, all PWA reported high successful inner speech during silent naming (all > 80%, refer to Table 2). Silent naming and overt naming share that they are picture naming tasks, both requiring access to the word without manipulation, with the difference between them being motor planning, programming, and execution needed for overt speech but not for inner speech. The relationship between Naming SIS and overt naming was not significant ($r_s = -0.17$, $p = 0.47$, Figure 6a). Naming SIS and picture rhyme judgement share that they are both inner speech, without the need for overt output, but picture rhyme judgement requires access to **and** manipulation of the word forms (to make a rhyme judgment), while silent naming requires only phonological access. The relationship between silent naming and picture rhyme judgement was also not significant (rs = 0.14, p = 0.55, Figure 6b). When visually examining the trends based on aphasia severity (Figure 6), there were some overall patterns. Those with mild aphasia severity appeared to have better correspondence between Naming SIS and other picture-based lexical access variables than those with moderate aphasia, who appeared to have more variability and individual differences, which we discuss below.

[FIGURE 6 HERE]

Looking at individual participants illuminates the differences in abilities and reports of inner speech. For example, in participants with moderate aphasia, there are varying strengths and weaknesses on tasks requiring lexical-semantic access with pictures (Table 3). First consider participant 1008, who made more phonemic naming errors (i.e., similar speech sounds, e.g., "bowl" for "bow") rather than semantic errors (i.e., similar meaning, e.g., "bowl" for "cup"). Other errors were either unrelated to the target or no response. Participant 1008 endorsed having 100% naming SIS, but overt naming score for the same items was 46%. They also scored poorly on the objective picture task (50%). On the rating scales for general subjective inner speech questions (Section 2), participant 1008 marked a "5" for how often they experience SIS, suggesting agreement between Naming SIS and general SIS experiences. Participant 1011 scored 50% on picture rhyme judgement, 13.5% on overt naming (with more semantic than phonemic errors), but 100% on reports of successful inner speech for the same pictures and rated the frequency of successful inner speech as 5 (almost always). In contrast, participant 1003, who also has moderate aphasia, scored 55% on picture rhyme judgement but 92% on overt naming (with no phonemic errors and some semantic) and 95% on reports of successful inner speech and rated the frequency of successful inner speech as lower (3, sometimes). It is possible that participants have difficulty with different cognitive processes required for picture rhyme judgement (semantic and phonological access, phonological awareness, working memory) even though they performed similarly on that task (~50%), and aphasia severity does not accurately convey the underlying complexity. Those differences appear in overt naming scores and their perceptions of inner speech.

Another contrast is from participant 1016, who performed well on the picture rhyme judgement task (82%) but poorly on overt naming (54%, with more semantic than phonemic errors), while rating the frequency of SIS in cases of anomia out loud as 2 (rarely). We postulate that participant 1016's primary impairments are in articulatory and motor components, but if they cannot find a word, it is because they usually do not have the semantic or phonetic access. Participant 1008 may have a different issue because they make more phonemic errors, with difficulty in phonetic access which makes both the picture rhyming and overt naming challenging.

[TABLE 3 HERE]

Research Question 3: Relationship between methods of measuring inner speech

When evaluating the relationships between objective rhyme judgement tasks (Section 1.1) and the subjective ratings about inner speech (Section 2), we chose to focus on the picture rhyme judgement task, which is an inner voice task with cognitive complexity similar to the written rhyme judgement task. We also chose the picture rhyme task because the PWA group demonstrated a wide range of performance, enabling us to evaluate variation.

We identified a single significant relationship between the picture rhyme judgement task and the subjective ratings from Section 2, which was a relationship with the Post-Stroke question. The lower the picture rhyme inner speech score, the more likely people were to say that they used their inner speech more now than pre-stroke (r_s= -0.53, p-value = 0.02). Since aphasia severity may mediate this relationship, we performed a partial correlation between picture rhyme score and subjective post-stroke rating with WAB AQ, finding that the relationship between picture rhyme score and subjective post-stroke rating remained significant (r=-0.48 p=0.04). Relationships between the picture rhyme task and SIS (r_s = 0.04, p-value = 0.86), IwW (r_s = -0.31, p-value = 0.19), and Inner vs. Outer (r_s = -0.38, p-value = 0.09) were not significant.

Figure 7 shows a comprehensive representation of correlations between general subjective inner speech questions (Section 2) and inner speech in-the-moment tasks (Section 1) as well as the demographic variables of age, years with aphasia, education (in years), and aphasia severity (WAB AQ). As presented above, aphasia severity is correlated with the objective picture and written tasks, but not any of the subjective tasks. The rhyme tasks from Section 1.1 are, understandably, correlated with one another. Some of the subjective ratings were related to one another, i.e., IwW is significantly related to Subjective: Post-Stroke and Subjective: SIS (both r_s=0.50, p<.05). Unsurprisingly, given that anomia is a hallmark of aphasia and weighted heavily in aphasia standardized batteries like the WAB-R, overt naming performance is related to aphasia severity. Overt naming is also related to the two objective rhyming judgements which require the “inner voice,” picture and written. Naming SIS was not associated with any other variables as shown in Figure 7.

[FIGURE 7 HERE]

Discussion

Inner speech is difficult to measure by nature, as it occurs entirely in one’s head.

Nevertheless, it has been measured using both subjective and objective paradigms, with individuals demonstrating its validity by showing that perception of inner speech relates to expected brain areas (Fama et al., 2017). Understanding how inner speech relates to severity when it is measured differently has downstream implications for more comprehensive study of inner speech in aphasia. It is important to evaluate the relationship of aphasia severity and inner speech, as aphasia severity is related to quality of life and participation (Williamson et al., 2011), and inner speech is likely related to quality of life factors, such as self-awareness (Morin, 2009, 2011; Morin & Michaud, 2007) and psychological distress (Heavey & Hurlburt, 2008). By evaluating how aphasia severity is related to inner speech, we can gain insight into the mechanisms behind changes in self-awareness and psychological health with aphasia. With the many methods of measuring inner speech that have been utilized in prior research, it is important to investigate the methods, revealing what information we can gain from each method and how the methods complement each other, measuring different facets of the inner speech phenomenon.

Is inner speech mediated by aphasia severity and how does this relate to task complexity?

First and foremost, individuals with aphasia performed better on the auditory task than the written or picture tasks. This was expected, given that individuals with aphasia performed worse on objective rhyme methods that required more complex cognitive-linguistic processes (e.g., picture, written rhyming, i.e., inner voice tasks) than simple processes (e.g., aural decisions, i.e., inner ear tasks). This difference in performance depends on whether the task involves inner voice or inner ear, which was shown by Tree & Playfoot (2019) in an individual with conduction aphasia and impaired phonological short-term memory and Fama, Henderson et al. (2019) and Langland-Hassan et al. (2015) in multiple PWA.

Individuals with milder aphasia performed better on the objective inner speech tasks that evaluated inner speech in-the-moment, especially the more complex picture and written tasks. We did not identify a significant relationship between severity and performance on the auditory task, likely because the performance was near ceiling for nearly all participants. Indeed, individuals with aphasia performed worst on inner speech elicited via pictures compared with inner speech elicited orthographically or aurally (Fama, Henderson, et al., 2019; Langland-Hassan et al., 2015; Tree & Playfoot, 2019). Altogether, this aligns with what we expected: that these objective tasks draw upon complex cognitive-linguistic resources, which are often impacted in more severe aphasia, and this impact on language abilities means that the inner voice is more impacted than the inner ear. This leads us to conclude that objective methods that require several steps for accessing inner speech (e.g., picture rhyme, written rhyme) are more liable to be testing the language production-like component of inner voice and more likely to be related to aphasia severity. The subjective task of Naming SIS was not related to aphasia severity for our sample, even though it is also an in-the-moment inner speech task. This is likely due to the difference in cognitive resources required for access alone vs. access and manipulation.

Notably, aphasia severity groups did not demonstrate significantly different responses for any of the subjective questions about how they generally experience inner speech. There was high within-group variability for most subjective questions (in both PWA and control groups), and for the PWA group, we did not find that ratings were related to severity. Given that the control group also had high within-group variability, this suggests that people have a range of experiences of inner speech, regardless of history of brain injury, which has been shown before in neurotypical adults (Alderson-Day et al., 2018; Racy et al., 2020). In fact, the frequency of inner speech ranged from 0-75% in young adults (Hurlburt et al., 2016).

We identified some trends suggesting that severity of aphasia may play a role in general subjective perception of inner speech if evaluated in a larger sample: for example, participants with moderate aphasia reported using inner speech more often now than they did before their stroke, while participants with milder aphasia tended to report having the same frequency of inner speech use now as they did prior to their stroke. This could be because people with moderate aphasia are intentionally using inner speech more often due to their more impaired overt speech. Both groups reported equal or better inner speech than overt speech. This finding strongly encourages further investigation into the nature and uses of inner speech across a diverse sampling of individuals with aphasia.

Figure 5 is an example of heterogeneity on the general subjective perceptions about inner speech in the mild and moderate aphasia groups. When asked about having an Idea Without Word ("IwW"), there was wide variation within the moderate aphasia group (all scale ratings used), whereas the mild group used only three of five of the scale ratings. This suggests greater homogeneity of experience in the mild group as it relates to Idea Without Word. For Successful Inner Speech frequency, both the mild and moderate groups showed high intra-group heterogeneity, using all scale options (moderate) or four out of five scale options (mild). Both the mild and moderate groups trended in feeling that their inner speech was better than their overt speech, demonstrating intra-group homogeneity. For the frequency of inner speech pre-stroke, the mild group tended to be most homogeneous, with nearly all members selecting option three ("about the same as before my stroke") whereas the moderate group tended to show higher intra-group heterogeneity but with a trend of choosing a higher scale option ("experience inner speech more now"). While these subjective scales did not significantly relate to aphasia severity or demonstrate significant differences between participant groups (control, PWA), they serve to clearly illustrate individual differences in the experience of inner speech. Some of these individual differences were likely present pre-stroke, because as we saw, the control group also had heterogeneity in their answers. Individual differences in inner speech have also been shown in PWA in past literature (Fama, Henderson, et al., 2019; Geva, Bennett, et al., 2011; Langland-Hassan et al., 2015; Stark et al., 2017).

How do in-the-moment inner speech (measured objectively and subjectively) and general perception of inner speech (measured subjectively) relate to one another?

There are two ways that we can examine this relationship: through direct comparison and through how each of them relates to aphasia severity. When compared directly through Spearman's correlations, picture rhyme judgement accuracy only related to one general, subjective inner speech rating for PWA: Post-Stroke. It is possible that the relationship seen between lower accuracy on the objective picture task and reporting using inner speech more now than prior to their stroke is because these individuals are relying on their inner speech more due to impoverished output but still have difficulty with working memory tasks and/or lexical retrieval (which are required for the inner rhyme judgement).

The picture and written rhyming performance related to aphasia severity, in that people with milder aphasia tended to have higher accuracy on the objective inner speech tasks. For the general subjective ratings, there were no significant relationships with aphasia severity. This difference, combined with the direct comparisons between objective rhyme judgements and subjective inner speech ratings also not showing many significant relationships, likely indicates that in the moment inner speech methods (objective, subjective) and methods evaluating general inner speech experience (subjective) capturing different aspects of the phenomenon of inner speech.

For the objective measures of evaluating in the moment inner speech, the auditory inner rhyme judgement is likely using the "inner ear," while the other two rhyme judgements (picture and written) use the "inner voice." Picture and written rhyme judgements likely use expanded monologue, and intentional inner speech, based on the ConDialInt model (Grandchamp et al., 2019) because the rhyme judgements require phonetic information, are not conversational, and are purposeful, and the auditory task may share these qualities, although it may be less "intentional." While the subjective rating questions about general inner speech were designed to also get at perception of inner speech, it is possible that participants were thinking beyond inner speech used for isolated tasks (e.g., rhyming, naming/word-finding) and more about their experience of an inner narrative, which may be more condensed and less intentional but likely still monologue. Our group defines inner narrative as talking to oneself internally, in silence, using words throughout daily life, which relates to work by Morin and colleagues on the relationship between inner narrative and extra-linguistic processes such as self-regulation and problem-solving (Morin, 2011a; Morin & Michaud, 2007), which is likely less related to aphasia severity. More condensed inner speech (i.e., in inner narrative) could mean that the inner speech is a few words rather than full sentences or only partial phonetic access, and less intentional inner speech would be more spontaneous. The differences in the condensation and intentionality dimensions may contribute to the results which indicate that there are not strong relationships between the objective and subjective methods of measuring inner speech and how they related to aphasia severity. For example, condensed inner speech would not require total phonetic access, so participants with difficulty accessing the phonemes would not necessarily report worse inner speech if they are thinking of inner narrative (condensed).

Limitations

Not all aphasia severity levels were represented in our participants in the PWA group, so relationships with severity may be stronger or appear differently if people with more severe aphasia as well as more mild aphasia (e.g., latent aphasia) participate in future studies. At a case study level, we explored the relationship between phonemic and semantic errors in overt naming and inner speech abilities, but this may benefit from a more detailed analysis with a larger set of naming data.

Some of the subjective questions about general inner speech may not have been answered in the same way by the two groups, as some members of the control group answered that they sometimes or often cannot say a word out loud but have successful inner speech (SIS), meaning that they had access to the word and its phonemes in their head. This seems unlikely for people who do not have a speech or language disorder, except possibly if someone were extremely exhausted or recently had oral surgery. Because the experience of anomia is likely uncommon for the control group, they may have assumed that the question was addressing a different experience, such as having partial phonetic access (tip-of-the-tongue phenomenon).

Future Directions

Future work should focus not only on inner speech in the moment for isolated tasks (e.g., rhyming, naming), but as a part of the daily experience of people with aphasia. The focus may also shift to include more clinical applications of relatively strong inner speech compared to overt speech, such as how successful inner speech may predict therapy outcomes. We are presently evaluating both of these directions in our lab as part of a National Institute on Disability, Independent Living and Rehabilitation Research (NIDILRR)-funded grant to Stark.

Conclusion and Clinical Implications

Aphasia severity relates to in-the-moment inner speech tasks that are cognitively more complex, such as picture and written rhyme judgments, but not to cognitively less complex tasks, including subjective inner speech during naming and auditory rhyme judgment. Aphasia severity was also not related to subjective interpretation of the general frequency/use of inner speech. We therefore suggest that the difference may be due to the cognitive-linguistic processes involved and alternative interpretations of the questions on the subjective rating scales.

Individuals with aphasia may retain inner speech despite impoverished lexical or phonological abilities (Fama et al., 2017; Fama, 2019; Fama, Snider, et al., 2019; Fama & Turkeltaub, 2020; Geva, Bennett, et al., 2011; Stark et al., 2017). It has long been understood that inner speech plays a role in aphasia, e.g., Brumfitt (1993) states that one must understand that the 'impact on the aphasic person's experience may be the lost ability to talk to oneself'. Moreover, in brain damaged patients who eventually recover from their trauma, self-awareness often returns in parallel with inner speech (Ojemann et al., 1996). The importance of inner speech in aphasia recovery is emphasized by Dr. Jill Bolte Taylor, who in her book, *My Stroke of Insight*, describes her experience recovering from a left hemispheric stroke (Bolte Taylor, 2006), recalling a total lack of inner speech: "The dramatic silence that had taken residency inside my head" (pp. 75–76), and emphasizing that this silence greatly impacted her self-awareness and sense of self (p. 67).

Various methods of measuring inner speech provide meaningful information about the experience of inner speech in both neurotypical controls and people with aphasia. These methods are likely capturing unique facets of inner speech. Because there is the possibility of relatively strong inner speech abilities in comparison to overt speech in PWA (Geva, Bennett, et al., 2011; Stark et al., 2017), characterizing inner speech in this population has clinical implications. It is important, however, to consider the specific facets of inner speech that are relevant to the clinical situation. For example, if a clinician is interested in the underlying naming abilities of their client, they may be more likely to use in the moment methods of measuring inner speech to assess the accuracy of access to the word (Fama, Snider, et al., 2019). The clinician could then gain valuable information about where their client was having difficulty in the process of speech production (i.e., difficulty with access alone vs. difficulty with access and manipulation of language information). If a clinician wanted to know more about the frequency of inner speech in daily life, they may use subjective rating scales and adapt them to their desired questions. Finally, if the clinician were interested in the contents and functions of their client's inner speech, they could opt for open-ended subjective questions or adapt pre-existing surveys (Racy et al., 2020) for individuals with aphasia (our lab is presently working on this). As always, clinicians and researchers must carefully select their methods of measuring a phenomenon based on the desired outcome.

Data Availability Statement

The data that support the findings of this study are openly available in The Open Science Framework at https://osf.io/8sjny/, reference number DOI 10.17605/OSF.IO/8SJNY.

References

Alderson-Day, B., Mitrenga, K., Wilkinson, S., McCarthy-Jones, S., & Fernyhough, C. (2018). The varieties of inner speech questionnaire – Revised (VISQ-R): Replicating and refining links between inner speech and psychopathology. *Consciousness and Cognition*, *65*, 48–58. https://doi.org/10.1016/j.concog.2018.07.001

Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex*, *112*, 91–106. https://doi.org/10.1016/j.cortex.2018.05.015

Bolte Taylor, J. (2006). *My stroke of insight: A brain scientist's personal journey*.

Brinthaupt, T. M., Hein, M. B., & Kramer, T. E. (2009). The Self-Talk Scale: Development, Factor Analysis, and Validation. *Journal of Personality Assessment*, *91*(1), 82–92. https://doi.org/10.1080/00223890802484498

Brumfitt, S. (1993). Losing your sense of self: What aphasia can do. *Aphasiology*, *7*(6), 569–575. https://doi.org/10.1080/02687039308248631

Coltheart, M. (1981). The MRC Psycholinguistic Database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497–505. https://doi.org/10.1080/14640748108400805

Fama, M. E. (2019). Self-reported inner speech relates to phonological retrieval ability in people with aphasia. *Consciousness and Cognition*, 12.

Fama, M. E., Hayward, W., Snider, S. F., Friedman, R. B., & Turkeltaub, P. E. (2017). Subjective experience of inner speech in aphasia: Preliminary behavioral relationships and neural correlates. *Brain and Language*, *164*, 32–42. https://doi.org/10.1016/j.bandl.2016.09.009

Fama, M. E., Henderson, M. P., Snider, S. F., Hayward, W., Friedman, R. B., & Turkeltaub, P. E. (2019). Self-reported inner speech relates to phonological retrieval ability in people with aphasia. *Consciousness and Cognition*, *71*, 18–29. https://doi.org/10.1016/j.concog.2019.03.005

Fama, M. E., Snider, S. F., Henderson, M. P., Hayward, W., Friedman, R. B., & Turkeltaub, P. E. (2019). The Subjective Experience of Inner Speech in Aphasia Is a Meaningful Reflection of Lexical Retrieval. *Journal of Speech, Language, and Hearing Research*, *62*(1), 106–122. https://doi.org/10.1044/2018_JSLHR-L-18-0222

Fama, M. E., & Turkeltaub, P. E. (2020). Inner Speech in Aphasia: Current Evidence, Clinical Implications, and Future Directions. *American Journal of Speech-Language Pathology*, *29*(1), 560–573. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0212

Geva, S., Bennett, S., Warburton, E. a., & Patterson, K. (2011). Discrepancy between inner and overt speech: Implications for post-stroke aphasia and normal language processing. *Aphasiology*, *25*(3), 323–343. https://doi.org/10.1080/02687038.2010.511236

Geva, S., & Fernyhough, C. (2019). A penny for your thoughts: Children's inner speech and its neuro-development. *Frontiers in Psychology*, *10*(JULY), 1708. https://doi.org/10.3389/fpsyg.2019.01708

Geva, S., Jones, P. S., Crinion, J. T., Price, C. J., Baron, J.-C., & Warburton, E. A. (2011). The neural correlates of inner speech defined by voxel-based lesion-symptom mapping. *Brain*, *134*(10), 3071–3082. https://doi.org/10.1093/brain/awr232

Geva, S., & Warburton, E. A. (2019). A Test Battery for Inner Speech Functions. *Archives of Clinical Neuropsychology*, *34*(1), 97–113. https://doi.org/10.1093/arclin/acy018

Grandchamp, R., Rapin, L., Perrone-Bertolotti, M., Pichat, C., Haldin, C., Cousin, E., Lachaux, J.-P., Dohen, M., Perrier, P., Garnier, M., Baciu, M., & Lœvenbruck, H. (2019). The ConDialInt Model: Condensation, Dialogality, and Intentionality Dimensions of Inner Speech Within a Hierarchical Predictive Control Framework. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02019

Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, *136*(2), 302. https://doi.org/10.1037/a0018436

Hurlburt, R. T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016). Exploring the Ecological Validity of Thinking on Demand: Neural Correlates of Elicited vs. Spontaneously Occurring Inner Speech. *PLOS ONE*, *11*(2), e0147932. https://doi.org/10.1371/journal.pone.0147932

Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition*, *22*(4), 1477–1494. https://doi.org/10.1016/j.concog.2013.10.003

Langland-Hassan, P., Faries, F. R., Gatyas, M., Dietz, A., & Richardson, M. J. (2021). Assessing abstract thought and its relation to language with a new nonverbal paradigm: Evidence from aphasia. *Cognition*, *211*. https://doi.org/10.1016/j.cognition.2021.104622

Langland-Hassan, P., Faries, F. R., Richardson, M. J., & Dietz, A. (2015). Inner speech deficits in people with aphasia. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00528

Meek, G., Ozgur, C., & Dunning, K. (2007). Comparison of the t vs. Wilcoxon Signed-Rank Test for Likert Scale Data and Small Samples. *Journal of Modern Applied Statistical Methods*, *6*(1). https://doi.org/10.22237/jmasm/1177992540

Morin, A. (2011a). Self-Awareness Part 1: Definition, Measures, Effects, Functions, and Antecedents. *Social and Personality Psychology Compass*, *5*(10), 807–823. https://doi.org/10.1111/j.1751-9004.2011.00387.x

Morin, A. (2011b). Self-Awareness Part 2: Neuroanatomy and Importance of Inner Speech. *Social and Personality Psychology Compass*, *5*(12), 1004–1017. https://doi.org/10.1111/j.1751-9004.2011.00410.x

Morin, A., Duhnych, C., & Racy, F. (2018). Self-reported inner speech use in university students. *Applied Cognitive Psychology*, *32*(3), 376–382. https://doi.org/10.1002/acp.3404

Morin, A., & Michaud, J. (2007). Self-awareness and the left inferior frontal gyrus: Inner speech use during self-related processing. *Brain Research Bulletin*, *74*(6), 387–396.

https://doi.org/10.1016/j.brainresbull.2007.06.013

Ojemann, J. G., John W., M., & Daniel L., S. (1996). Preserved Function in Brain Invaded by Tumor. *Neurosurgery*, *39*(2), 253–259. https://doi.org/10.1097/00006123-199608000-00003

Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition*, *106*(1), 528–537.

Perrone-Bertolotti, M., Rapin, L., Lachaux, J.-P., Baciu, M., & Lœvenbruck, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research*, *261*, 220–239. https://doi.org/10.1016/j.bbr.2013.12.034

Racy, F., Morin, A., & Duhnych, C. (2020). Using a Thought Listing Procedure to Construct the General Inner Speech Questionnaire: An Ecological Approach. *Journal of Constructivist Psychology*, *33*(4), 385–405. https://doi.org/10.1080/10720537.2019.1633572

Smith, J. D., Wilson, M., & Reisberg, D. (1995). The role of subvocalization in auditory imagery. *Neuropsychologia*, *33*(11), 1433–1454. https://doi.org/10.1016/0028-3932(95)00074-D

Stark, B. C., Geva, S., & Warburton, E. A. (2017). Inner speech's relationship with overt speech in poststroke aphasia. *Journal of Speech, Language, and Hearing Research*, *60*(9). https://doi.org/10.1044/2017_JSLHR-S-16-0270

Tree, J. J., & Playfoot, D. (2019). How to get by with half a loop – An investigation of visual and auditory codes in a case of impaired phonological short-term memory (pSTM). *Cortex*, *112*, 23–36. https://doi.org/10.1016/j.cortex.2018.08.008

Wallace, G. L., Peng, C. S., & Williams, D. (2017). Interfering With Inner Speech Selectively

Disrupts Problem Solving and Is Linked With Real-World Executive Functioning. *Journal of Speech, Language, and Hearing Research*, *60*(12), 3456–3460. https://doi.org/10.1044/2017_JSLHR-S-16-0376

Appendix: Tables and Figures

Table 1

Table 1: Demographic information of all included participants

	Aphasia (N=20)	Control (N=20)	Overall (N=40)
Age			
Mean (SD)	56.4 (9.09)	55.2 (9.66)	55.8 (9.28)
Median [Min, Max]	55.5 [35.0, 72.0]	55.0 [41.0, 70.0]	55.0 [35.0, 72.0]
Sex			
F	11 (55.0%)	13 (65.0%)	24 (60.0%)
M	9 (45.0%)	7 (35.0%)	16 (40.0%)
Years Education			
Mean (SD)	15.9 (1.79)	16.5 (1.82)	16.2 (1.81)
Median [Min, Max]	16.0 [12.0, 18.0]	18.0 [14.0, 18.0]	16.0 [12.0, 18.0]
Years with Aphasia			
Mean (SD)	8.85 (8.09)	NA (NA)	8.85 (8.09)
Median [Min, Max]	5.00 [1.50, 31.0]	NA [NA, NA]	5.00 [1.50, 31.0]
Aphasia Type			
Anomic	7 (35.0%)	0 (0%)	7 (17.5%)
Broca's	6 (30.0%)	0 (0%)	6 (15.0%)
Conduction	6 (30.0%)	0 (0%)	6 (15.0%)
Transcortical Motor	1 (5.0%)	0 (0%)	1 (2.5%)
WAB Aphasia Quotient			
Mean (SD)	73.1 (13.9)	NA (NA)	73.1 (13.9)
Median [Min, Max]	70.9 [47.8, 94.0]	NA [NA, NA]	70.9 [47.8, 94.0]

Table 2

Table 2: Descriptive Statistics for PWA and Controls on Objective and Subjective Inner Speech Tasks, with Section 1 as proportions and Section 2 as ratings from 1-5.

Section	Task	Aphasia (N=20)	Control (N=20)	Overall (N=40)
Sect	**Inner Picture Rhyming**			

Section	Task	Aphasia (N=20)	Control (N=20)	Overall (N=40)
	Mean (SD)	0.643 (0.166)	0.946 (0.0646)	0.795 (0.197)
	Median [Min, Max]	0.592 [0.368, 0.947]	0.970 [0.790, 1.00]	0.841 [0.368, 1.00]
	Inner Written Rhyming			
	Mean (SD)	0.634 (0.169)	0.948 (0.0409)	0.791 (0.200)
	Median [Min, Max]	0.625 [0.350, 0.950]	0.950 [0.880, 1.00]	0.895 [0.350, 1.00]
	Inner Auditory Rhyming			
	Mean (SD)	0.755 (0.148)	0.961 (0.0341)	0.858 (0.149)
	Median [Min, Max]	0.750 [0.450, 1.00]	0.950 [0.880, 1.00]	0.915 [0.450, 1.00]
Section 1.2	**Naming SIS**			
	Mean (SD)	0.972 (0.0397)	0.993 (0.0133)	0.972 (0.0397)
	Median [Min, Max]	0.973 [0.838, 1.00]	1.00 [0.970, 1.00]	0.973 [0.838, 1.00]
	Overt Naming			
	Mean (SD)	0.786 (0.217)	0.992 (0.0170)	0.889 (0.184)
	Median [Min, Max]	0.865 [0.135, 1.00]	1.00 [0.950, 1.00]	0.970 [0.135, 1.00]
Section 2	**Idea without Word**			
	Mean (SD)	2.70 (1.08)	2.35 (0.933)	2.53 (1.01)
	Median [Min, Max]	3.00 [1.00, 5.00]	2.00 [1.00, 4.00]	3.00 [1.00, 5.00]
	Successful Inner Speech			
	Mean (SD)	3.55 (1.32)	3.39 (1.61)	3.47 (1.45)
	Median [Min, Max]	4.00 [1.00, 5.00]	4.00 [1.00, 5.00]	4.00 [1.00, 5.00]
	Inner vs. Outer			
	Mean (SD)	1.90 (0.718)	2.95 (0.826)	2.43 (0.931)
	Median [Min, Max]	2.00 [1.00, 3.00]	3.00 [2.00, 5.00]	2.00 [1.00, 5.00]
	Post-Stroke			
	Mean (SD)	3.25 (1.02)	NA (NA)	3.25 (1.02)
	Median [Min, Max]	3.00 [1.00, 5.00]	NA [NA, NA]	3.00 [1.00, 5.00]

Table 3

Table 3: Scores on tasks and ratings requiring lexical-semantic access, highlighting the individual differences in participants with moderate aphasia.

Participant	Aphasia Severity (AQ)	Overt Naming Errors (%)	Overt Naming Score (%)	Naming SIS (%)	Rating- SIS (Scale 1-5)	Objective-Picture Rhyming (%)
1003	Moderate (67.6)	Phonemic: 0 Semantic: 50 Other: 50	92	95	3 - sometimes	55
1008	Moderate (47.8)	Phonemic: 28 Semantic: 0 Other: 72	46	100	5 - almost always	50
1011	Moderate (60.2)	Phonemic: 3 Semantic: 22 Other: 75	14	100	5 - almost always	50
1016	Moderate (52.4)	Phonemic: 0 Semantic: 24 Other: 76	54	100	2 - rarely	82

Figure 1

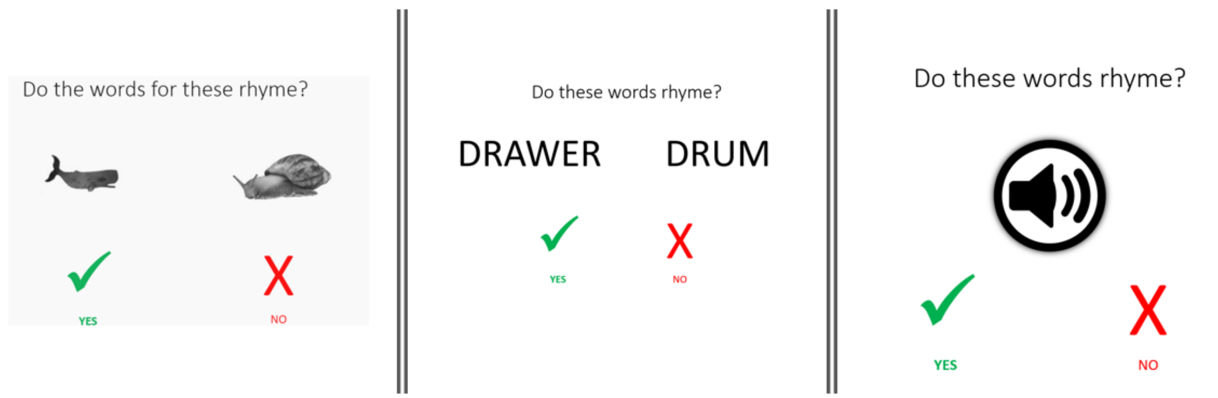

Figure 1: Examples of displays for picture rhyme judgement, written rhyme judgement, and auditory rhyme judgement. Participants responded by touching the green check mark or the red x.

Figure 2

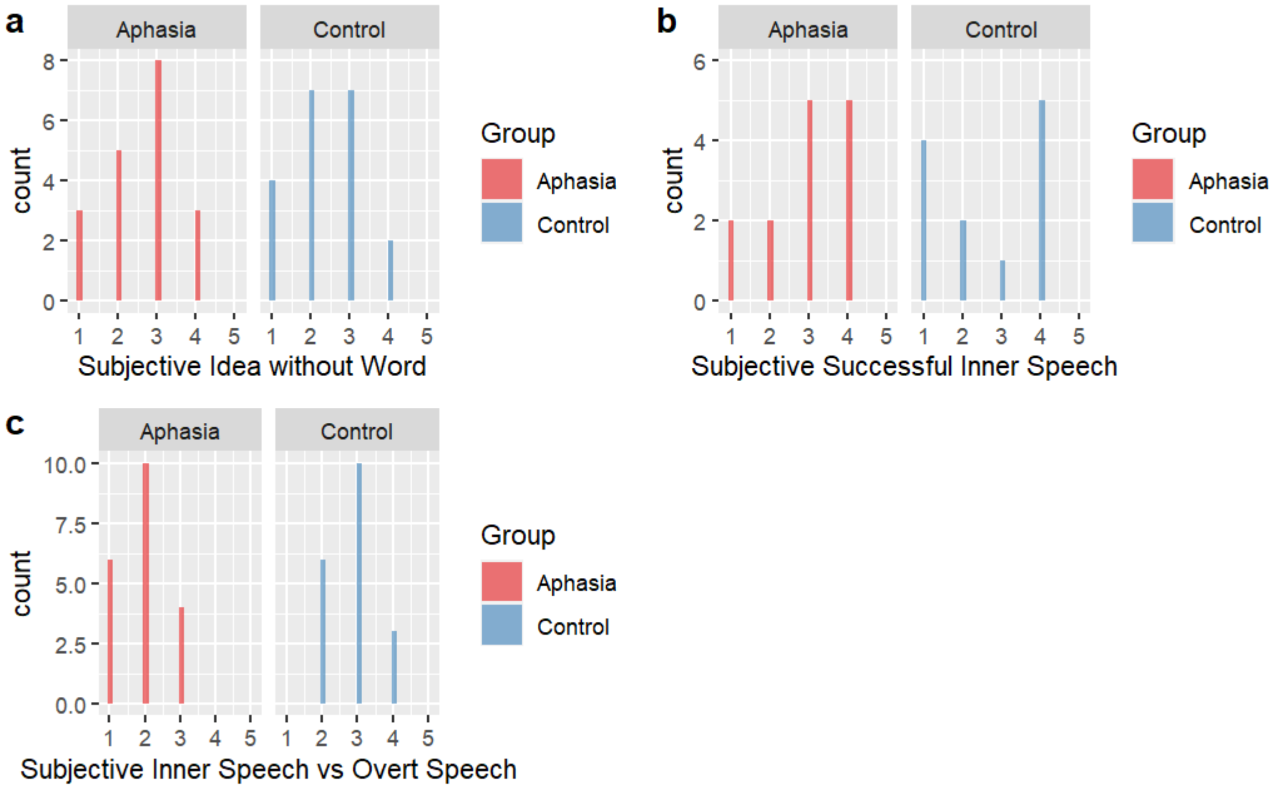

Figure 2: Subjective Ratings with persons with aphasia (PWA) and neurotypical controls comparison.
<u>Idea without Word (IwW)</u>, "How often do you see something and know what it is but can't say the word for it out loud?" with 1 being almost never and 5 being almost always; <u>Successful Inner Speech (SIS)</u>, "A few minutes ago we talked about how we sometimes see something and know what it is but can't say the word for it out loud. When that happens to you, how often can you say the right word for it in your head, and have it sound right?" with 1 being almost never and 5 being almost always; <u>Inner vs. Overt Speech</u>, "In general, how do you think your abilities to speak in your head compare to your abilities to speak out loud?" with 1 being much better in head and 5 being much better out loud

Figure 3

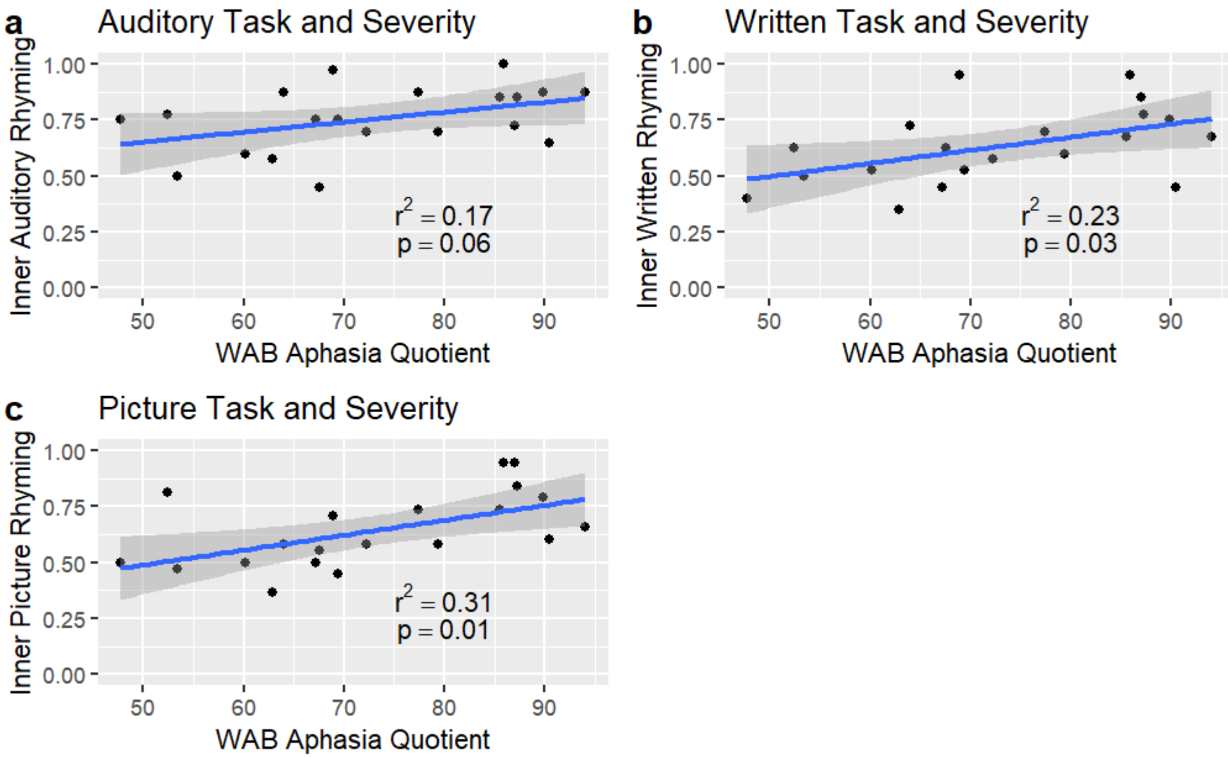

Figure 3: Objective Inner Rhyming Judgement Tasks - a) Auditory: participants hear two words; b) Written: participants see two words; c) Picture: participants see drawings of two objects. x-axis is WAB AQ, where a higher score is milder aphasia severity.

Figure 4

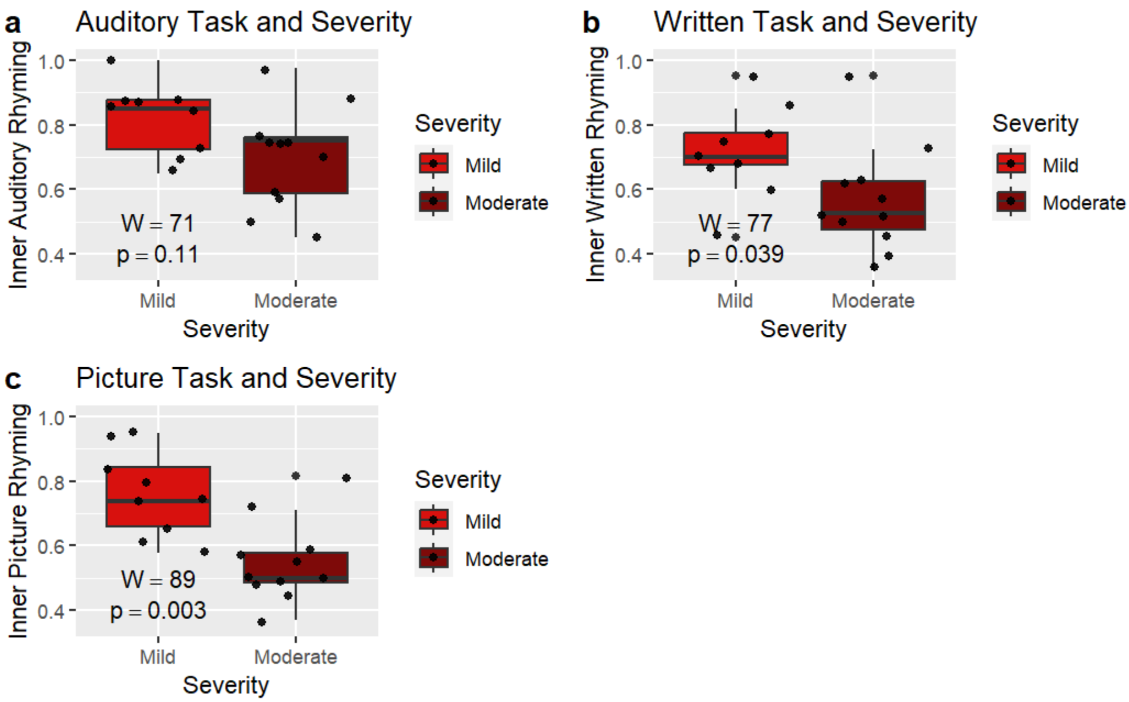

Figure 4: Objective Inner Rhyme Judgement and Categorical Severity - a) Auditory; b) Written; c) Picture. The groups were split into mild severity and moderate severity based on the cut-offs from the WAB. Jittered dots represent individual performance.

Figure 5

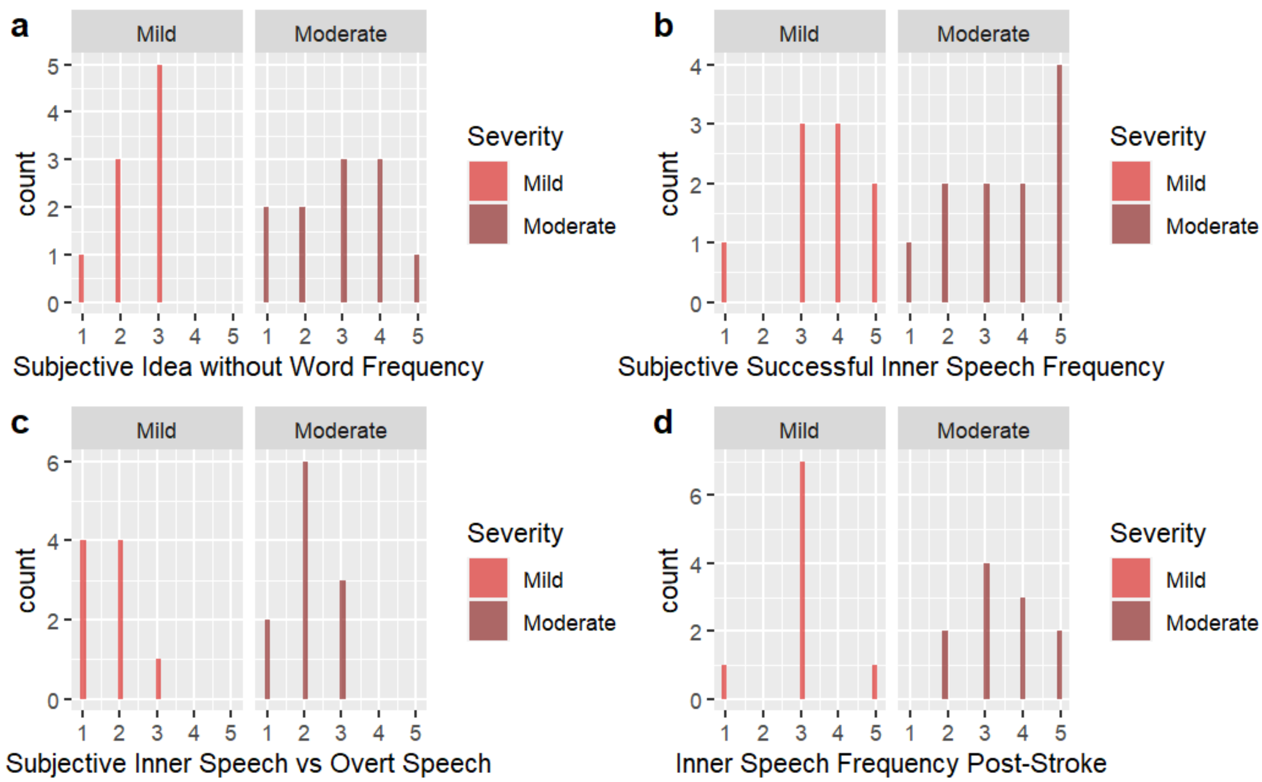

Figure 5: Subjective Ratings, comparisons between mild and moderate aphasia. Idea without Word (IwW), "How often do you see something and know what it is but can't say the word for it out loud?" with 1 being almost never and 5 being almost always; Post-Stroke, "Do you think you talk to yourself in your head less now, more now, or about the same as before your stroke?" with 1 being much less now and 5 being much more now; Successful Inner Speech (SIS), "A few minutes ago we talked about how we sometimes see something and know what it is but can't say the word for it out loud. When that happens to you, how often can you say the right word for it in your head, and have it sound right?" with 1 being almost never and 5 being almost always; Inner vs. Overt Speech, "In general, how do you think your abilities to speak in your head compare to your abilities to speak out loud?" with 1 being much better in head and 5 being much better out loud

Figure 6

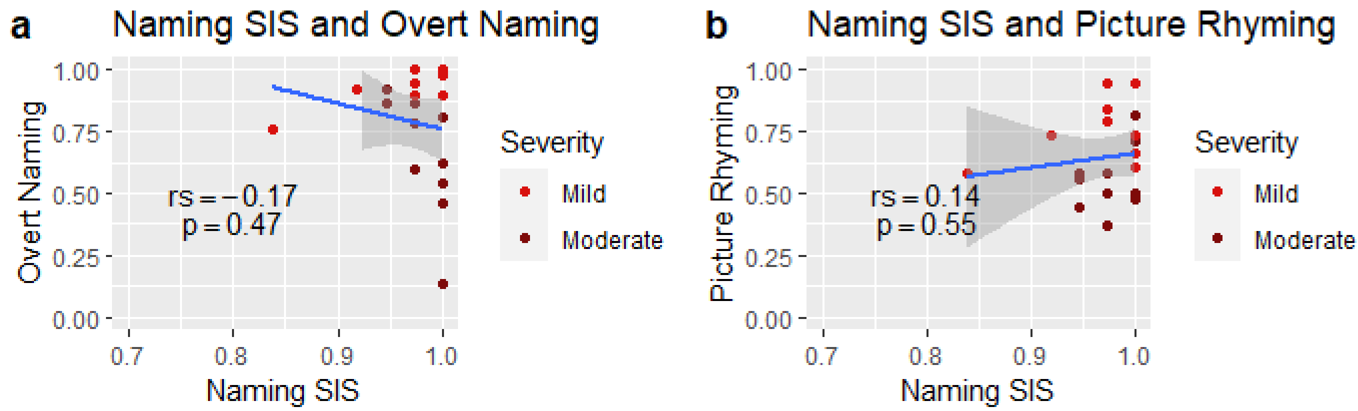

Figure 6: Correlations between reports of successful inner speech (SIS) in-the-moment of silent naming ("Naming SIS") with a) overt (out loud) naming performance and b) performance on the objective picture rhyme judgement task for PWA. Spearman's correlation coefficient (rs) and significance values are provided. Additionally, dots represent each participant, with light red dots representing those with mild aphasia severity, and dark red dots representing those with moderate aphasia severity. The trend line (in blue) includes all participants with aphasia.

Figure 7

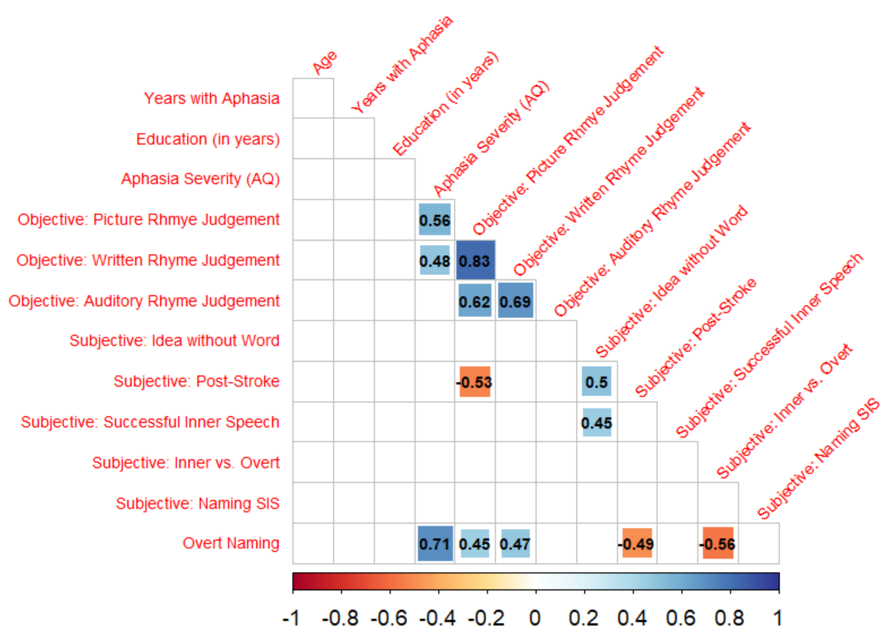

Figure 7: Correlations between objective and subjective measures, as well as demographic

variables. All correlations shown are significant with a 95% confidence interval. Larger squares and darker hues indicate stronger correlations, also shown by the numbers in each square.