Title: Investigating the Effects of Rhythm and Entrainment on Fluency in People With Aphasia

NCT05248295

Document date: 09/23/2024

I. Background and Significance

Aphasia is a language disorder, most commonly as a result of stroke, that can affect speaking, understanding what others say, reading, and writing. It is common, affecting an estimated 2 million people in the U.S. (National Aphasia Association). A common goal in aphasia therapy is to increase meaningful spoken output; especially in people with nonfluent subtypes of aphasia, this may be discussed in terms of improving speech fluency. Apraxia of speech (AOS) is a motor speech planning disorder that frequently co-occurs with nonfluent aphasia, and many treatment tasks simultaneously address both AOS and aphasia. Treatment of comorbid aphasia and AOS often focuses on common expressions and tailored sets of phrases in the hope of maximizing functional gains.

While there are numerous approaches to promoting fluency in people with aphasia (PWA), one common technique is to ask the person to "say it with me." Unison speech, also referred to as choral speech, is explicitly incorporated into several formal treatments for aphasia and/or AOS, including Oral Reading for Language in Aphasia (ORLA; Cherney et al., 1986), Melodic Intonation Therapy (MIT; Albert, et al., 1973; Norton et al., 2009), "integral stimulation" that has long been described in the AOS literature (Rosenbek & Wertz, 1972), and script training approaches such as AphasiaScripts™ (Lee et al., 2009). Notably, unison production has proven effective not just for aphasia and AOS but also for stuttering (e.g., Kalinowski & Saltuklaroglu, 2003; Kiefte & Armson, 2008).

Although unison speaking is a component of many therapy approaches, the conditions under which it works best have not yet been established, and it remains unclear why it works. There is evidence that for PWA, being able to see the speaker's mouth and benefit from audiovisual input, rather than audio input alone, results in improved spoken output, measured as more correctly produced, different words (Fridriksson et al., 2012). Fridriksson et al. (2012) compared audiovisual input (AV), audio-only input (AO), and spontaneous speech (SS) in 13 PWA and found that while output in the AO condition was superior to SS on average, it was not significantly better. In contrast, output in the AV condition was significantly better than in either AO or SS. The benefit conferred by the addition of visual information was attributed to the mirror neuron system, which has been purported to link observation of an action with its execution by the observer (Gallese et al., 1996).

More recent evidence from our group, however, indicates that audio-only input can be quite effective at improving spoken output. Kershenbaum at al. (2017) compared solo repetition to unison repetition (hear a sentence and repeat it, compared to hear a sentence and repeat it in unison with the model) and found significant increases in % syllables correct in the unison condition. There were numerous differences between Fridriksson et al. (2012) and Kershenbaum et al. (2017), including task (reading vs. repetition), stimulus length (connected sentences vs. single sentences or phrases), and output measure (correctly produced words vs. % syllables correct), but one potentially critical difference was that while the first study used sentences with a conversational prosody, the second used sentences spoken metrically. Racette et al. (2006) documented better spoken output (% words correct) under an audio-only unison speech condition compared to solo repetition, using utterances with a typical, conversational prosody (in French), but saw even larger choral gains when the productions were sung, i.e., when they followed a metrical pattern. It seems that whether or not the model utterance is metrical is an important variable impacting the degree

to which PWA benefit from speaking in unison with a model, perhaps particularly when the co-speaker's mouth is not visible (i.e., audio-only conditions).

In considering why and how unison speech improves fluency, an important question is how the participant's production is temporally related to the model. In unison speech, the speaker knows the target utterances ahead of time (e.g., through memorization, repetition, or having the printed text) and can therefore align their productions very closely with the model, with mean between-speaker asynchronies as low as 40-50 ms in some conditions. When two speakers attempt to speak synchronously (as opposed to one speaker aligning with a recording), there is an alternation between who "leads" in time (Cummins, 2003). Taken together, the very short between-speaker asynchronies and lack of a consistent "leader" suggest that speaking in synchrony is an active process, collaborative when it involves two active speakers, and entails prediction of what is coming next in the signal. It is *not* the case that one speaker is closely shadowing the other. Neurotypical speakers are able to generate precise speech timing predictions from their knowledge as fluent speakers of the language in question, a limited amount of data on how the conversation partner speaks, and an understanding of a shared goal, and use those predictions to guide their speaking, a complex motor action. This is *entrainment:* aligning a motor action with an external signal.

None of the existing studies examining unison speaking in PWA have reported asynchronies between the participants' productions and the model. Indeed, it seems that few, if any, of these studies time-locked the recording of participants' productions to the stimulus. It therefore remains unknown whether PWA truly entrain their speech during unison productions, under what conditions they are able to do so, how this varies across individuals with aphasia, and whether true entrainment is needed in order for a unison benefit to be conferred.

Metrical structures are built upon integer divisions of some unit, just as musical rhythms are built from notes of different values; when speech adheres to a metrical framework, syllables occur at a regular interval, or some integer multiple of that interval. In contrast, more typical conversational prosody is based on linguistic knowledge, particularly syntactic and prosodic knowledge. We propose a "two-route" model of speech planning, in which this planning can be accomplished in one of two ways: according to either a metrical or a linguistic framework, with the latter used to produce typical, conversational speech. Further, we suggest that aphasia may differentially affect these frameworks, such that PWA are a well-suited population with which to test the two-route claim. Individuals within the population may show deficits to one route, both routes, or neither. People with nonfluent aphasia are expected to be unable to use syntax and prosody to construct a fluent speech plan, while those with fluent aphasia will often be able to do so, resulting in varied performance across individuals with aphasia when attempting to use a linguistic framework to produce conversational prosodic timing. Concerning the metrical route, PWA frequently show unusual difficulty processing rhythm and timing information (Zipse et al., 2014). Individuals with this deficit should be prone to difficulty using a metrical framework to organize their speech. The proposed study will explore this two-route model by (1) evaluating whether PWA show true entrainment when speaking along to speech that adheres to either a metrical or a linguistic framework, indicating that they can make use of the framework to predict what comes next in the

signal (Aim1), and (2) test whether PWA benefit from a unison model with either a metrical or a linguistic framework (Aim 2).

II. Specific Aims

The specific aims of this study are to:

1. ***Determine how timing alignment of sentences spoken along with an auditory model is affected by metrical vs. conversational prosody, for people with aphasia and for typical speakers.*** Participants will attempt to speak in unison to an audio recording of two types of sentences: (1) sentences with regularly occurring stressed syllables (metrical) and (2) sentences spoken with a typical, conversational timing pattern (conversational).
Prediction: PWA will show varied patterns of performance across these tasks. Successful entrainment with one timing framework will not necessarily be associated with successful entrainment to the other type.

2. ***Determine whether people with aphasia benefit from attempting entrainment when the signal consists of (a) sentences produced with metrical prosody, and (b) sentences produced with conversational prosody.*** In a 2x2 design, sentences with metrical vs. conversational prosody will be produced in repetition vs. in unison with a model. The number of accurate syllables produced will be compared across repetition (imitation) and entrainment (prediction) to evaluate the effect of attempting entrainment.
Prediction: PWAs will vary as to whether they benefit from entrainment to sentences with metrical or conversational prosody. Critically, those who benefit from entrainment with one type of prosody will not necessarily benefit from entrainment with the other type.

III. Subject Selection

One-hundred participants will be recruited for this study, including 40 with aphasia, 40 age-matched controls, and up to 20 controls enrolled for the purposes of piloting and refining stimuli, as well as training study staff on data collection procedures. All participants must be between the ages of 18 and 80 and pass a hearing screening, demonstrating thresholds ≤ 40 dB at 1 kHz, 2 kHz, and 4 kHz in at least one ear, tested in a quiet room. Hearing screens will be conducted by the PI, who is a licensed speech-language pathologist, or by a research assistant trained by the PI, after informed consent has been obtained. Given that the screening thresholds are quite liberal, it is not anticipated that many participants will be excluded from the study based on the hearing criterion.

Control participants must exhibit native-speaker level fluency in American English and must not report any history of speech or language disorders, stroke, or other neurological disease. They must repeat both sentences in the short form of the Sentence Repetition subtest of the Boston Diagnostic Aphasia Examination (BDAE; Goodglass, Kaplan, & Barresi, 2001) without any errors, given a maximum of 2 attempts per sentence. They must also pass the Mini-Cog screening (Borson et al., 2000), a short cognitive screening measure that has been validated in older Americans to have

similar sensitivity and specificity for dementia as the Mini-Mental Status Exam. It consists of memory and recall of a list of three unrelated words after a filled interval of a clock drawing task. To be considered negative for cognitive impairment, participants have to either recall all three words or have a normal clock drawing if they only recall one or two words (Borson et al., 2003).

PWA must be at least 6 months post-onset of aphasia and report pre-morbid native-speaker level fluency in American English, supported by report of other relevant information (e.g., age at which the individual began speaking English and extent they used it in daily life). They must score $\geq$ the 15th percentile, averaged across the 3 auditory comprehension subtests of the BDAE (Word Identification, Following Commands, Complex Ideational Material), and correctly repeat $\geq$50% of words on the Single Word Repetition subtest of the BDAE. In addition, they must score within normal limits on at least 3 of the 5 nonverbal subtests of the Cognitive-Linguistic Quick Test (CLQT; Helm-Estabrooks, 2001).

Participants with aphasia will be recruited from the MGH IHP Aphasia Center, through RSVP for Health, through the Partners Research Portal website, the National Aphasia Association (NAA) website, and via an MGH-IHP IRB-approved Recruitment Database (IRB Protocol 2014P001166). At the MGH IHP Aphasia Center, information about the study will be briefly described and presented in written form as handouts at an aphasia group meeting at the MGH IHP Aphasia Center by the PI or a graduate student research assistant. All those interested in participating will be put in touch with the PI via email and phone. In addition, potential participants will be identified through the Research Patient Data Registry (RPDR) and sent a Research Invitation informing them about the study. Research Invitations will either be sent through the EPIC Patient Navigator system or through USPS. Study staff will check to ensure that patients have not opted out of receiving such invitations. If the research team does not receive a response, a call may be made to the potential participant 2 weeks after the letter has been sent. Control participants will be recruited via RSVP for Health, the Partners Research Portal website, the MGH-IHP IRB-approved Recruitment Database referenced above, flyers posted at the MGH IHP, and announcements seeking participants sent through Broadcast MGH and the IHP Daily News emails. Information about this research study will also be posted on the Institute's website. (Note that ad language for controls includes a version for "healthy adults" and a version for "healthy adults aged 50-80." In our experience, more general ads tend to attract mostly people <30, and people >50 are needed to age-match to the people with aphasia.) Control participants will be pre-screened by phone to confirm that they have native-level fluency in American English and do not have a history of speech or language disorders.

Recruitment materials include an email address for the convenience of participants. Email will be used to establish contact, i.e., send study information in response to an initial contact from a potential participant (see "standard email response" documents), and to send appointment information and directions to the study site. Email will *not* be used to ask screening questions to determine eligibility. The warning language concerning secure vs. unencrypted email, as required by Partners, is included in the standard email response, along with a statement that additional communications will be sent securely unless the person requests otherwise and acknowledges the associated risks. If a potential participant with aphasia responds without acknowledging their acceptance of risks, they will receive a response requesting a phone number to communicate further, since sending future emails with the "send secure" requirements may be difficult and

confusing for a person with aphasia. Email contact is especially helpful to communicate with people with aphasia, many of whom have great difficulty on the phone but may have more success with email due to the reduced time pressure when formulating their thoughts, and the ability to read and re-read a message. Even if a follow-up call is required for security compliance, having the initial email that the person can refer to during the call may be very helpful, especially for scheduling purposes (e.g., choosing dates and times for people who may mix up numbers; seeing the same information printed while they are hearing it can be very helpful).

IV. Subject Enrollment

Informed consent will be obtained at the start of the first (in many cases, only) study visit. This will be done at the MGH IHP, either in a clinic room within the Aphasia Center, or in an assessment room in the research building (for control participants). Informed consent will be obtained from each participant with aphasia by the PI or another study staff member who is a licensed speech-language pathologist with expertise in aphasia. For control participants, informed consent will be obtained by the PI or a research assistant trained by the PI for this purpose. For participants with aphasia, care will be taken to ensure that each person understands the nature of the study and has adequate time to ask questions and consider whether to participate. Study staff obtaining consent will make use of appropriate techniques to ensure good communication with a person with aphasia, including use of gesture, writing down or underlining key words, slow rate of speech, and adequate pauses to allow for processing time. Authorized study staff will verbally summarize the study procedures and the risks and benefits, and will make it clear that participation is voluntary and the participant can choose to discontinue the study at any time. Study staff will offer to read the consent form to the potential participant, in case the participant has difficulty reading due to aphasia. Finally, the authorized study staff member will ask the potential participant if he or she has any questions or concerns, allow time for a response, and clarify whether the person would like to participate before obtaining the participant's signature. Clients with aphasia who meet the study criteria and choose to enroll in the study will participate in the study at the MGH IHP Aphasia Center, likely on the same day as their regular treatment session to minimize extra travel or effort on the part of the participant. Control participants will be asked to come to the MGH IHP at a mutually agreed-upon time. For controls, participation will require 1-2 visits totaling about 3 hours. For PWA, participation will require 2 study visits totaling about 4 hours. For PWA with existing diagnostic assessment data, the total time needed will be about 3 hours, which can be completed in 1 or 2 visits. Controls will be compensated $50 for participation, and PWA will be compensated $75. If any participant does not pass the pre-screening, they will be paid $30.

V. Study Procedures

Tasks Administered to Characterize Participants

The following assessments will be administered to participants with aphasia (PWA) to confirm eligibility for inclusion in this study:

- Hearing screen (pure tones at 1 kHz, 2kHz, 4kHz)
- Long forms of the Single Word Repetition subtest and the 3 auditory comprehension subtests (Word Identification, Following Commands, Complex Ideational Material) of the *BDAE*
- Nonverbal subtests of the *Cognitive-Linguistic Quick Test* (*CLQT*; Helm-Estabrooks, 2001; Symbol Search, Symbol Trails, Design Memory, Mazes, Design Generation)

The following assessments will be administered to control participants to confirm eligibility for inclusion in this study:

- Hearing screen (pure tones at 1 kHz, 2kHz, 4kHz)
- Short form of the Sentence Repetition subtest of the *BDAE*
- Mini-Cog screening (Borson et al., 2000)

In addition to the assessments needed to evaluate whether they meet inclusion criteria, PWA will complete the assessments listed below to characterize aphasia subtype and evaluate fluency, agrammatism, and apraxia of speech (AOS). In cases where Aphasia Center clients have completed these assessments within the past year, scores will be obtained from their clinical file to reduce unnecessary repetition of assessments.

- Sentence Repetition (short form) and Picture Description subtests of the *BDAE*
- One additional picture description, in order to compute Correct Information Units (CIUs) per second as a primary measure of fluency, and syllables/sec and % CIUs as secondary measures (Nicholas & Brookshire, 1993)
- *Boston Naming Test, 2nd edition*, short form (*BNT-2*; Kaplan, Goodglass, & Weintraub, 2000)
- *BDAE* subtests for extended testing of syntactic processing: Touching A with B, Reversible Possessives, Embedded Sentences
- Select subtests of the *Apraxia Battery for Adults, 2nd edition* (*ABA-2*): Diadochokinesis, Increasing Word Length 2A, Limb Apraxia, Oral Apraxia, Repeated Trials

All participants will complete the following tasks in order to evaluate and quantify their nonverbal working memory capacity:

- Corsi blocks: a visuospatial memory span task, requiring participants to observe the investigator point to a sequence of blocks and then repeat the sequence
- Test of Everyday Attention (TEA), Elevator Counting with Distraction: an auditory selective attention task, requiring participants to listen to sequences of tones, counting the lower-pitched tones while ignoring the higher-pitched distractor tones

All participants will complete the following task in order to evaluate and quantify their musical entrainment ability:

- Tapping to music portion of the Beat Alignment Test (BAT; Iversen & Patel, 2008): to evaluate ability to entrain to a musical beat

Experimental Task Design

To address the aims, a 2x2 design will be used to compare the performance of controls and PWA across 4 conditions, which will collectively evaluate performance with metrical vs. linguistic timing frameworks both when repeating (solo production) and when attempting to speak in unison (unison condition) (Fig 1). Each condition will follow a unison repetition task format: the participant hears the stimulus followed by an alerting tone, and then a repetition of the stimulus, to which the participant attempts to speak along.

Prosodic Timing Framework

		Metrical	Conversational
Production Condition	Unison	metrical unison	conversational unison
	Solo	metrical solo	conversational solo

Fig. 1: Four conditions presented

To address Aim 2, two solo repetition conditions, one with metrical sentences and one with conversational sentences, will be included in the task. In these conditions, the participant will hear the stimulus followed by an alerting tone, and then repeat the stimulus. Performance on these conditions will be compared to performance on the corresponding unison repetition conditions.

Additional blocks of trials will allow for follow-up exploratory analyses of the role of self-monitoring one's own auditory feedback, as this has been implicated in speech accuracy in fluency in PWA (Jacks & Haley, 2015; Maas et al., 2015), and self-monitoring ability may be reduced during the unison conditions. In these additional trials, words and sentences will be repeated solo, either with or without masking noise. For the masked trials, masking noise (pink noise, which covers the speech frequencies) will be played through the headphones while the participant speaks, to interfere with their ability to hear their own productions.

Experimental Stimuli

The experimental task is comprised of 4 conditions: metrical sentence unison, conversational sentence unison, metrical sentence solo, and conversational sentence solo. Stimuli for all conditions will be matched for number of syllables and overall stimulus duration. and balanced across timing conditions for syntactic (e.g., wh- questions, compound subjects) and phonemic (numbers of fricatives, affricates, and clusters) complexity.

To enable controlled analysis of entrainment across metrical vs. conversational sentence conditions (Aim 1), these sentences will be developed in pairs. Each pair will have the same number of syllables and the same target words. The same acoustic landmarks within these words will be used for analysis (Fig. 2).

Fig. 2: An example of a 9-syllable sentence pair. Target words are underlined, and stressed syllables are in bold italics. In metrical sentences, stressed syllables occur at regular intervals.

| Metrical: | Re**be**cca com**ple**ted the **con**tract. |
| Conversational: | ***Jane Smith*** com**ple**ted the **big con**tract. |

For the exploratory trials aimed at investigating the role of self-monitoring one's own auditory feedback when speaking, there will be a total of 60 multi-syllabic words across 4 blocks, half with masking noise and half without. There will also be a total of 60 sentences (adapted from the Harvard sentences corpus (IEEE Audio and Electroacoustics Group, 1969) across 6 blocks, including 2 blocks with masking noise, and 4 without (2 prior to any masking noise, and 2 following the masked blocks).

Experimental Task
In order to ensure that task expectations are clear and to minimize the need for verbal cues and reiterating instructions (i.e., to minimize language comprehension demands for the PWA), the 4 conditions will be presented in blocks, with condition-specific instructions and 2 practice items at the start of each block. Within a block, each trial will proceed as follows: the participant will hear the stimulus item, followed by an orienting tone, and then a repetition of the stimulus. In unison conditions, the stimulus will be re-played during the participant's turn, so that the participant produces the stimulus in unison with the recording. In solo conditions, the stimulus will not be re-played, so that the participant repeats the stimulus alone. Gestural cues along with verbal prompts will be used as needed with the practice items to cue participants to "wait their turn" ("stop" gesture) and "speak along" (hand extended to indicate "your turn"). During presentation of experimental items, gestural cues will continue to be used as needed; this approach has been used with success with PWA on a similar task (Kershenbaum et al., 2017).

Data collection
Stimuli will be presented and participant productions recorded using a laptop running Audacity software (https://audacityteam.org/). Stimuli will be played to the participant via adjustable, over-the-ear headphones set to a comfortable listening level. The experimenter will be able to monitor what the participant hears via a single earphone connected to the audio output via a splitter. A headset microphone (Shure WH20XLR Dynamic Headset Microphone) connected to a laptop via an audio interface (Focusrite Scarlett 2i2 multichannel sound card) will be used to obtain all speech recordings. With this arrangement, stimuli and responses are separately recorded on two time-locked channels. At the start of each recording session, line delays will be tested and noted for adjustment during analysis; this entails playing an abrupt stimulus (brief acoustic click) through the headphones into the microphone, and recording the presentation-to-recording delay in Audacity. A directional microphone (Shure SM58S Vocal Microphone) will be used to record tapping to music.

VI. Biostatistical Analysis

Aim 1: Determine how timing alignment of sentences spoken along with an auditory model is affected by metrical vs. conversational prosody, for people with aphasia and for typical speakers.

Version date: 09/23/2024

Audio files of stimuli and responses produced in the unison conditions will be labeled using a partially automated approach to extract the moment of voicing onset for vowels in target syllables (refer to Fig. 2). First, time intervals surrounding each syllable onset will be hand labeled using Praat Software (Boersma & Weenink, 2016). Second, the Praat autoVOT plugin (Keshet et al., 2014) will be used to efficiently extract the timing of voicing onset (fig. 4). The autoVOT plugin has been demonstrated to perform comparably to manual annotation (Buz et al., 2018) and was judged reliable in an initial sample of our own data.
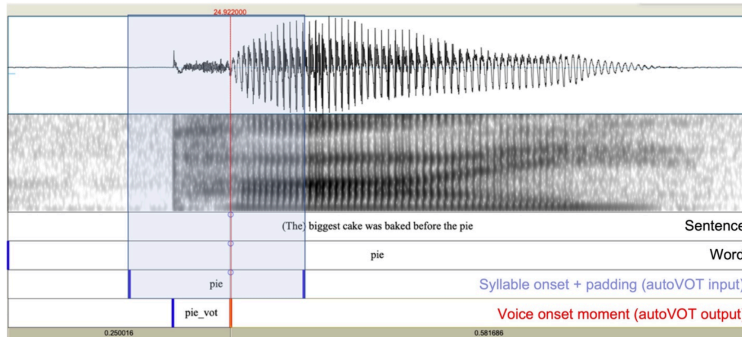

Fig. 4: Example of autoVOT labeling procedure. The blue shaded region indicates the manually labeled search window that is input to the autoVOT program. The red line shows the voice onset moment identified by the autoVOT.

Voice onset moment time points will be used to calculate the timing offset between participant productions and the corresponding time points in the stimuli. The line delay documented at the beginning of each participant's experimental session will subtracted from each extracted timepoint from participant productions, and the stimulus timepoint will then be subtracted from that difference, calculated as

$$(p - d) - s$$

where p represents participant voice onset moment, d represents line delay, and s represents stimulus voice onset moment. The resulting value represents whether and how far productions occur before or after a stimulus.

For statistical analysis, separate mixed-effects linear models will be run for controls and PWAs, with signed timing offset for each target syllable as the outcome.

Aim 2: Determine whether people with aphasia benefit from attempting entrainment when the signal consists of (a) sentences produced with metrical prosody, and (b) sentences produced with conversational prosody.

All sentences used to address Aim 2 will be transcribed and scored for syllable-level accuracy, according to a previously established, comprehensive set of rules (Kershenbaum et al., 2017). Twenty percent of utterances, distributed across PWA, will be independently transcribed and scored by a second rater to evaluate inter-rater reliability.

A 2x2 ANOVA will be used to evaluate the relative benefit of entrainment on % of syllables correct for speech with conversational vs. metrical prosody. Exploratory follow-up correlations will evaluate any possible relationship between a "metrical benefit" and other participant

characteristics including fluency (from the picture description task), AOS severity, and entrainment ability (from the tap to a metronome, entrain a syllable to a metronome, and tap to music tasks).

Power Analysis

Aim 1: The data from this study will be used to establish effect sizes to enable a meaningful power analysis in an external grant application; there are currently no studies of which we are aware comparing time-locked measures of speech entrainment across PWA and controls. To try to ensure that the study is adequately powered, even if some of the PWA are unable to produce all of the target words/syllables or if the target landmarks are not identifiable, we have included a rather large number of target landmarks (4-16 labeled landmarks per stimulus, and 8 stimuli/condition, yielding ~80 data points/condition for each participant).

Aim 2: Based on previous work from our group (Kershenbaum et al., 2017), the effect size of attempted speech entrainment is 0.975 (Cohen's d). This is based on % of accurate syllables produced by PWA, compared across 2 conditions: attempting to entrain to spoken sentences versus repeating them solo. In that study, the spoken stimuli were metrical, with regular, predictable timing. We anticipate that sentences with conversational prosody may result in less of an entrainment gain for many PWA. Therefore, we used 0.7 as an estimated effect size, an alpha level of 0.05, and 0.90 for power. With these constraints, the needed sample size is 19 PWA (sample size analysis conducted with G*Power; Faul et al., 2007). Our proposed sample size of 25 allows for some participant attrition, data loss, or unscorable data (e.g., PWA who cannot produce enough scorable tokens).

VII. Risks and Discomforts

For people with aphasia, there is a risk that the experimental tasks may be difficult and frustrating to complete.

There is also a risk that participant privacy could be compromised in some way. Study data, including audio recordings of participants and calculated results will be maintained on encrypted Partners computers. Any paper records will be kept in a file drawer in a locked office at the MGH Institute.

VIII. Potential Benefits

This study will not benefit participants directly. Rather, there are potential benefits to society, in that information learned in this study may be used to inform the development of more targeted and efficient aphasia therapy approaches.

IX. Monitoring and Quality Assurance

The PI will review the data after each testing session. Scoring may be completed by research assistants. In this case, the PI will initially review their scoring to ensure that it is done correctly. Syllable accuracy will be double-scored by another rater for at least 10% of the recorded samples, randomly selected across participants and conditions. A minimum intra-class correlation of 0.80 will be required for the ratings to be considered reliable.

X. References

Albert, M.L., Sparks, R.W., & Helm, N.A. (1973). Melodic intonation therapy for aphasia. *Archives of Neurology, 29*(2), 130-131.

Boersma, Paul & Weenink, David (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.21, http://www.praat.org/

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., Dokmak, A. (2000). The mini-cog: A cognitive "vital signs" measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry. 15*(11), 1021-1027.

Borson, S., Scanlan, J. M., Chen, P., & Ganguli, M. (2003). The Mini-Cog as a screen for dementia: Validation in a population-based sample. *Journal of the American Geriatrics Society, 51*(10), 1451-1454.

Buz, E., Buchwald, A., Fuchs, T., & Keshet, J. (2018). Assessing automatic VOT annotation using unimpaired and impaired speech. *International Journal of Speech-Language Pathology*, *20*(6), 624–634. https://doi.org/10.1080/17549507.2018.1490817

Cherney, L. R., Merbitz, C. T., & Grip, J. C. (1986). Efficacy of oral reading in aphasia treatment outcome. *Rehabilitation Literature, 47*(5-6), 112-118.

Cummins, F. (2003). Practice and performance in speech produced synchronously. *Journal of Phonetics, 31,* 139-148.

Dabul, (2000), *Apraxia Battery for Adults,* 2nd edition. Pro-Ed.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191.

Fridriksson, J., Hubbard, H. I., Hudspeth, S. G., Holland, A. L., Bonilha, L., Fromm, D., & Rorden, C. (2012). Speech entrainment enables patients with Broca's aphasia to produce fluent speech. *Brain, 135,* 3815–3829.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*(pt. 2), 593-609.

Version date: 09/23/2024

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The assessment of aphasia and related disorders* (3rd ed.). Philadelphia, PA: Lippincott, Williams & Wilkins.

Helm-Estabrooks, N. (2001). Cognitive Linguistic Quick Test (CLQT). San Antonio, TX: The Psychological Corporation, Pearson.

Iversen, J. R., & Patel, A. D. (2008). The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population. *Proceedings of the 10th International Conference on Music Perception and Cognition*, (Icmpc 10), 465–468.

Jacks, A., & Haley, K. L. (2015). Auditory Masking Effects on Speech Fluency in Apraxia of Speech and Aphasia: Comparison to Altered Auditory Feedback. *Journal of Speech, Language, and Hearing Research*, *58*(6), 1670–1686. https://doi.org/10.1044/2015_JSLHR-S-14-0277

Kalinowski, J., Saltuklaroglu, T. (2003). Choral speech: the amelioration of stuttering via imitation and the mirror neuronal system. *Neuroscience and Behavioral Reviews, 27,* 339-347.

Kaplan, E., Goodglas, H., & Weintraub, S. (2000). *Boston Naming Test*, 2nd edition. Pearson.

Kershenbaum, A., Nicholas, M. L., Hunsaker, E., & Zipse. L. (2017). Speak along without the song: What promotes fluency in people with aphasia? *Aphasiology*.

Keshet, J., Sonderegger, M., & Knowles, T. (2014). *AutoVOT: A tool for automatic measurement of voice onset time using discriminative structured prediction.* (Version 0.94) [Computer software]. https://github.com/mlml/autovot/

Kiefte, M., & Armson, J. (2008). Dissecting choral speech: properties of the accompanist critical to stuttering reduction. *Journal of Communication Disorders, 41,* 33-48.

Lee, J. B., Kaye, R. C., & Cherney, L. R. (2009). Conversational script performance in adults with non-fluent aphasia: Treatment intensity and aphasia severity. *Aphasiology, 23*(7-8), 885-897.

Liberman, M. Y., & Streeter, L. A. (1978). Use of nonsense-syllable mimicry in the study of prosodic phenomena. *Journal of the Acoustical Society of America, 63*(1), 231-233.

Maas, E., Mailend, M.-L., & Guenther, F. H. (2015). Feedforward and Feedback Control in Apraxia of Speech: Effects of Noise Masking on Vowel Production. *Journal of Speech, Language, and Hearing Research*, *58*(2), 185–200. https://doi.org/10.1044/2014_JSLHR-S-13-0300

National Aphasia Association. (n.d.). Retrieved Fenrury 26, 2018 from https://www.aphasia.org/aphasia-resources/aphasia-statistics/

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research, 36*(2), 338–350.

Norton, A., Zipse, L., Marchina, S., & Schlaug, G. (2009). Melodic intonation therapy: shared insights on how it is done and why it might help. *Annals of the New York Academy of Sciences, 1169,* 431–436.

Racette, A., Bard, C., & Peretz, I. (2006). Making non-fluent aphasics speak: Sing along! *Brain, 129,* 2571–2584.

Rosenbek, J. C., & Wertz, R. T. (1972). Treatment of apraxia of speech in adults. *Proceedings of the Clinical Aphasiology Conference,* Albuquerque, NM.