Get in Sync: Entrainment Mechanisms and Individual Characteristics Associated with Scripted-Sentence Learning in Aphasia

by

Yina Magally Quique Buitrago

Speech and Language Pathologist, National University of Colombia, 2006

Master in Music Therapy, National University of Colombia, 2012

Submitted to the Graduate Faculty of the

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Yina Magally Quique Buitrago

It was defended on

November 10, 2020

and approved by

Dr. William Evans, Assistant professor, Department of Communication Science and Disorders

Dr. Bharath Chandrasekaran, Vice Chair for Research and Professor, Department of Communication Science and Disorders

Dr. Marta Ortega-Llebaria, Associate professor, Linguistics

Dr. Lauryn Zipse, Associate Professor, MGH Institute of Health Professions

Dissertation Director: Dr. Michael Walsh Dickey, Director of PhD Program and Associate Professor, Department of Communication Science and Disorders

Get in sync: entrainment mechanisms and individual characteristics associated with scripted-sentence learning in aphasia

Yina Magally Quique Buitrago, PhD

University of Pittsburgh, 2020

Script training is a well-established treatment for aphasia, but its mechanisms of action are not well understood, and it remains unclear which people with aphasia (PWA) benefit from it most. Understanding both treatment mechanisms and individual characteristics leading to scripted-sentence learning can improve treatment implementation and personalization. This dissertation investigates mechanisms and individual characteristics associated with scripted-sentence learning via a protocol adapted for Spanish-speaking PWA.

A hypothesized mechanism of action in script training is speech entrainment, the unison production of sentences by patient and clinician. Entrainment relies on integrating rhythmic features, but it is unclear how these features facilitate scripted-sentence learning. Therefore, aim 1 examined the effects of speech entrainment to two types of rhythmic-enhanced sentences on scripted-sentence learning: word-stress aligned beats, which should support lexical retrieval, and metronomic beats, which should enhance memorization via chunking. In addition, identifying individual characteristics associated with scripted sentence learning can improve treatment personalization. Therefore, aim 2 examined behavioral characteristics (language, attention, and rhythmic processing) and cortical tracking (the coupling of neural oscillations and rhythmic speech properties) and their association with scripted-sentence learning.

Fourteen Spanish-speaking PWA participated in a five-session learning paradigm using three conditions: word-stress aligned, metronomic, and control (no beats). Aim 1 analyses showed

significant improvement over time across conditions, demonstrating successful scripted-sentence learning. Rhythmic-enhanced conditions engendered greater learning compared to the control condition, indicating that rhythmic structures of speech entrainment are a key active ingredient for learning scripted sentences. The two rhythmic-enhanced conditions did not differ over time in terms of learning response, suggesting that both rhythmic manipulations may facilitate scripted-sentence learning.

In aim 2 analyses, participants with more severe aphasia showed higher scripted-sentence learning estimates, reflecting that they started at a lower learning intercept and benefited more from the support that speech entrainment provides when learning highly formulaic language. More severely impaired PWA also showed lower cortical tracking, indicating lower perception of rhythmic speech properties. Lastly, attentional deficits and most rhythmic processing measures were not strongly associated with scripted-sentence learning.

Finally, this is the first study to examine scripted-sentence learning in Spanish-speaking PWA, demonstrating cross-linguistic benefits of script training interventions.

Table of Contents

List of Tables

List of Figures

Preface

Science is not a product of the isolated mind that belabors abstractions until reality reveals its treasured secrets. Scientific research is fundamentally a communal enterprise that aims to serve society by discovering ways to improve issues that affect a few of its members. Following this communal pattern, this dissertation is a child raised by an entire village, a child who wants to serve people with aphasia motivated by the vision that *they will find their voice again*. The following lines briefly honor those who, by their contributions of knowledge, patience, and love, brought this project to completion.

During the past years, I have received invaluable opportunities for which I will always be thankful. Above all, I am grateful for those who have helped me grow as a person and as a scientist, for my precious village of advisors, family, and friends who provided unceasing instruction and encouragement. Through their mediation and generosity, I have received training, financial, and professional opportunities that I desperately needed but never imagined could be within my reach. In what follows, I acknowledge those who have built the multiple dimensions of my academic dreams.

Regarding the financial dimension, this dissertation work was supported through generous funding from the American Speech-Language-Hearing Foundation (ASHFoundation) New Century Scholars Doctoral Scholarship; the Council of Academic Programs in Communication Sciences and Disorders (CAPCSD) Ph.D. Scholarship; the University of Pittsburgh's School of Health and Rehabilitation Sciences (SHRS) Ph.D. Student Award; and Fulbright.

Regarding my professional development, the past five years have brought numerous experiences that shaped me as a scientist. An exceptional mentorship team led by Dr. Michael

Walsh Dickey significantly contributed to these experiences. Dr. Dickey, for your genius but especially for your kindness, you have my most profound admiration. Thank you for your candid and enthusiastic view of science. I was also privileged to be mentored by Dr. Will Evans. Will, thank you for empowering me, for always providing a listening ear, and for strongly supporting research in Spanish speakers with aphasia (Pao y Antonio, gracias a ustedes también). *The world needs more mentors like you two!* To the rest of my dissertation committee: Dr. Lauryn Zipse, Dr. Marta Ortega, and Dr. Bharath Chandrasekaran, thank you. You have been instrumental in the culmination of this work.

Regarding my growth in a learning community that understands the challenges of doctoral work, I was fortunate by having shared the Ph.D. journey with such bright and caring friends. Rob Cavanaugh, Haley Dresang, Alex Swiderski, Christina Dastolfo-Hromack, Caryn Herring, Beth Haley, thank you for opening your country, homes, and hearts to me. You are all shining stars in my life, and I look with eager expectation for what the future brings for each one of you. Thank you, Dr. Nike Gnanateja, for your candid instruction and dedicated guidance.

Thank you to those who sustained me emotionally along the way. My deepest gratitude to my husband, the love of my life, for being my adventure partner. Your support has been the engine in this process. Thank you for loving me on both gray and sunny days. Thank you to my siblings Liseth and Sebastian for loving me in the distance. To Paly, for taking care of my little treasures Simon and Mia. Papi, las líneas escritas en este documento son fruto de tu arduo trabajo. Tus manos me enseñaron esfuerzo, tus zapatos me trajeron hasta aquí. Gracias por tu nobleza. A mi madre, por tus esfuerzos cuando era niña, gracias. Me gustaría seguirte viendo crecer en sabiduría y en amor para con los que te rodeamos. Orly, Mechas, y Judy, me siento privilegiada de ser parte de ustedes, les agradezco profundamente su apoyo incondicional. To our adoptive parents in

Pittsburgh, Vicki and Ted, thank you for your hospitality, generosity, and love. You are living proof of a God who provides for our every need!

Finally, to people with aphasia in Colombia (and around the world), thank you for your trust. I will continue working *with* you and *for* you.

1.0 Introduction

Aphasia is a chronic communication disability resulting from damage to brain regions that support language processing (National Institute of Neurological Disorders and Stroke, 2018). Aphasia rehabilitation depends on evidence-based treatments, but the mechanisms of action of some of those treatments remain poorly understood (i.e., how these treatments work). It is also unclear what behavioral and neurological characteristics of *people with aphasia (PWA)* are associated with response to treatment (i.e., for whom these treatments work). Understanding the relationship between treatment mechanisms and individual differences of PWA will improve existing treatments methods and will also allow for individualization of aphasia rehabilitation. Using a protocol adapted for Spanish-speakers with aphasia, the current study investigated mechanisms and individual behavioral and neuro-electrophysiological characteristics associated with script training, or more specifically, with scripted-sentence learning in aphasia.

Script training is a treatment method for PWA that improves the production of everyday sentences through intense rehearsal (Cherney, Halper, & Kaye, 2011; Goldberg, Haley, & Jacks, 2012; Lee, Kaye, & Cherney, 2009). Learning scripts creates opportunities for PWA to participate in conversation and daily situations through rehearsed sentences, helping them to overcome the obstacles to life participation associated with aphasia (Cruice, Worrall, & Hickson, 2006). Several studies have examined script learning in PWA, all of which have reported a successful acquisition of treated scripts and an increase in fluency, rate, and production of script-related content words (e.g., Cherney, Halper, Holland, & Cole, 2008; Lee et al., 2009). However, both the treated scripts and the learning conditions have varied greatly across studies. Hence, there are no guidelines about the optimal conditions for script learning to be successful. The current study examined *scripted* *sentences* as a step forward in understanding the characteristics of script training stimuli that are key to successful learning.

Generally, script training follows a learning process that includes listening, repetition, speech entrainment (also known as choral reading), and independent production. The current study specifically focused on *speech entrainment*, the unison production of sentences by patient and clinician. The benefits of speech entrainment have been studied in PWA (e.g., Fridriksson, Basilakos, Hickok, Bonilha, & Rorden, 2015), but the role of the rhythmic features of speech entrainment on scripted-sentence learning had not been examined to date. Entrainment relies on detection, integration, and production of rhythmic features (Phillips-Silver, Aktipis, & Bryant, 2010), but it is unclear exactly how these features facilitate learning of scripted sentences in aphasia. Two hypotheses regarding this facilitatory effect are: i) unison production of scripted sentences enables PWA to align to word stress, supporting lexical retrieval (e.g., Soto-Faraco, Sebastián-Gallés, & Cutler, 2001), and ii) unison production of scripted sentences enhance sentence memorization via chunking (e.g., Purnell-Webb & Speelman, 2008; Stahl, Kotz, Henseler, Turner, & Geyer, 2011). To test these hypotheses, this study examined the effects of speech entrainment to two rhythm types, stress-aligned vs. metronomic beats, compared to control sentences, to characterize mechanisms underlying scripted-sentence learning.

In addition, identifying individual behavioral and neuro-electrophysiological characteristics associated with scripted-sentence learning will improve both patient selection and treatment customization (i.e., identifying what works best for whom). In terms of *behavioral* characteristics associated with treatment response, PWA vary in their language and cognitive profiles (Halai, Woollams, & Lambon Ralph, 2017), which also generates variability in how PWA respond to any treatment, including script training (Cherney, Halper, et al., 2011; Goldberg et al., 2012; Lee et al., 2009). The current study examined language, attention, and rhythmic processing measures and correlate them with scripted-sentence learning estimates to observe how variability in each measure is associated with scripted-sentence learning performance.

In terms of neuro-electrophysiological characteristics associated with treatment response, *cortical tracking*, the coupling of low-frequency oscillations and the speech envelope (e.g., Ding & Simon, 2014) reflects the successful *alignment of brain responses with rhythmic features* present in the speech signal, thereby enabling efficient processing of linguistic features (Peelle & Davis, 2012; Peelle, Gross, & Davis, 2013). For scripted-sentence learning, cortical tracking may index individual perception of speech-rhythmic properties in connected sentences (Hamilton & Huth, 2018), in turn predicting an individual's ability to learn sentences. In other words, cortical tracking as a measure of perception of speech rhythmic features can provide valuable evidence about whether perceiving speech-rhythmic information facilitates scripted-sentence learning when enhancing speech's rhythmic properties during speech entrainment. Thus, the current study examined the correlation between cortical tracking measures and scripted-sentence learning estimates to observe how variability in this measure was associated with learning performance.

The current study therefore aimed to: i) investigate how speech entrainment to different rhythms affects scripted-sentence learning; and ii) to explore whether individual characteristics in behavioral measures and cortical tracking are associated with scripted-sentence learning. These overall goals help to characterize mechanisms (i.e., how it works) and individual characteristics associated with scripted-sentence learning (i.e., for whom it works), improving aphasia rehabilitation for Spanish-speaking PWA.

2.0 An Overview of the Script Training Literature

This section aims to present an overview of the script training literature that serves as the basis for understanding speech entrainment as a mechanism through which PWA learn scripted sentences. Script training is a treatment method for PWA that targets sentences in either monologues or dialogs for improvement through intense rehearsal (Goldberg et al., 2012). This method relies on pre-determined phrases, sentences, or lager segments of connected speech to improve the encoding, storage and retrieval of memorized language and thereby increase verbal production as a result. Through script training PWA that are severely affected can gain a more complete output of rehearsed sentences and therefore advance in functional communication. The first positive effects of this treatment method were reported in two PWA who were trained in 3 scripts (Youmans, Holland, Muñoz, & Bourgeois, 2005). The scripts were 3-4 sentences long, had a range of 13 to 32 words per sentence, and were reproduced with 97-100% accuracy after 26 sessions, in a single subject multiple baseline design (Youmans et al., 2005).

Script training can be administered through multiple approaches. One of those approaches is a one-to-one format delivered in clinical settings (e.g., Youmans et al., 2005). A *second* approach includes in-person clinical training based on videos containing conversational scripts. Participants use the videos to practice the scripts of real-life scenes (e.g., Bilda, 2011). Given the increasing demand for telehealth, the *third* approach divides a patient's treatment between in-clinic training and training with a clinician via video-conferencing (e.g., Goldberg et al., 2012). The *fourth* approach is based on an at-home computerized treatment with weekly speech and language pathology supervision. In this approach, training can be delivered by a virtual therapist through a software called AphasiaScripts™ (e.g., Cherney & Halper, 2008; Cherney et al., 2008; Kaye &

Cherney, 2016; Manheim, Halper, & Cherney, 2009); through an updated version of AphasiaScripts called AphasiaRx™ (e.g., Van Vuuren & Cherney, 2014); or through a B.A.BAR, a device that uses barcodes to read conversational lines or monologues (e.g., Nobis-Bosch, Springer, Radermacher, & Huber, 2011).

For the *first* approach, one-to-one in the clinic, Youmans et al. (2005) described script training in 2 PWA. Each participant was trained in 3 personalized scripts during three sessions a week (30-45 minutes); participants were expected to practice at home for 15 minutes a day using audiocassettes. Participants were trained in one script at a time; when they accurately produced the target phrases 20 times, then the next script phrase was added. Youmans et al. (2005) measured the percentage of correct script-related words produced (number of script words used divided by the total number of words in the target script), number of errors, and speaking rate. They reported that both participants mastered (97-100%) the three scripts at approximately the same time (26 weeks). In their measure of the number of errors during script production, they reported variability within sessions, with a decrease by the end of the maintenance phase. Finally, for speaking rate, the authors reported variable outcomes from session to session with decreased rate at the end of the maintenance phase. The researchers did not report statistical analysis applied to the measurements mentioned above besides comparing between baseline and maintenance performance.

In the *second* approach, in-person video-based training, Bilda (2011) reported a case series design with 5 PWA. Participants were trained in 20 scripts, three hours a day, for ten days. The scripts initiated, contained, and ended with formulaic language (e.g., hello, thank you, see you). The videos were filmed at realistic locations with actors presenting scenes of daily life. Each participant's performance was measured by the production of a single phrase from the script (second sentence). The target sentence was assessed on a scale from 0-3 based on naming scoring parameters of the Aachen Aphasia Battery (AAT; Huber, Poeck, Weniger, & Willmes, 1983). The results showed a significant improvement in the production of the target sentences for all participants (Friedman test $\chi^2(3,4) = 8.45$, $p<.05$), and an improvement in the use of the untrained scripts (Friedman test $\chi^2(3, 4) = 10.45$, $p< .05$). After treatment, participants also scored higher in the AAT subtests of naming and repetition.

The *third* approach is delivered through video-conferencing/telepractice (e.g., Goldberg et al., 2012; Rhodes & Isaki, 2018). Goldberg et al. (2012) used a single-subject multiple baseline treatment design in two PWA. The subjects were trained in 2 scripts over a three-week period. Goldberg et al. (2012) reported that both participants improved significantly when measuring the percentage of scripted content words with a large effect size for one participant (Cohen's d=4.97) and a smaller effect size for the other (Cohen's d=1.97). Also, across scripts, participants showed improvement in speech rate, defined in this study as the total number of words produced divided by the total amount of time of the participant's turn. The final measure of this study was related to disfluencies, such as repetitions, revisions, and self-interruptions. When expressing the number of disfluencies in proportion to the total number of words, the overall result was a decreasing pattern with high variability in participant performance. The authors probed script generalization with a novel partner in the script's topic, with a novel partner deviating from the target script, and with the clinician deviating from the target script. Their generalization results were highly variable with no consistent trends that could make significant distinctions among the three contexts.

The *fourth* approach is a computerized treatment with weekly supervision. This approach is where most of the script training work has been developed. Before moving on to the studies' specific details, it is important to note that the main argument for providing computerized versions of scripts is to increase the frequency of direct treatment while reducing the costs of face-to-face interventions. According to a recent survey (n= 34 PWA), participants who used computers before aphasia onset report positive experiences when using computers for treatment purposes (Finch & Hill, 2014). From the clinician's side, a survey reported that speech and language pathologists (SLPs; n=107) use software mostly for indirect or supplemental rehabilitation purposes (Davis & Copeland, 2006). Additionally, some computer-delivered treatments, like script training, allow therapists to collect precise data about the amount of time that PWA have practiced to examine treatment-related variables such as dosage, intensity, type of practice (distributed versus massed), and even participant motivation.

Two main software have been used for scripted sentences: AphasiaScripts™ (plus its updated version called AphasiaRx™), and a device known as B.A.BAR. In AphasiaScripts™, a sentence of a dialogue or monologue appears on a computer screen, in a written form, while a virtual therapist reads the sentence aloud (Cherney et al., 2008). Sentences for the script are developed between the patient and SLP. The software includes the possibility of decreasing cues (therapist's voice, written word, movement of the articulators). Participants can click on specific words to repeatedly practice single words. The other software (B.A.BAR), uses barcodes containing linguistic information (words, phrases, sentences, texts) and a barcode scanner. Users get a booklet with codes accompanied by images that can be scanned. The scanner reads the barcode and plays the information related to it. Information can be recorded, stored, and replayed as needed with this device (Nobis-Bosch et al., 2011). The device does not provide feedback, and it's not capable of speech recognition.

Cherney et al. (2008) described a single-subject multiple baseline design using AphasiaScripts™ with 3 PWA. Participants were trained in one monologue and two dialogues.

Comparing pre-post training performance, participants improved in i) the number and percentage of script-related words; ii) the total number of script-related morphemes, nouns, verbs, and modifiers; and iii) the number of script-related words produced per minute. In a similar study, Cherney and Halper (2008) used the same design, number of scripts, treatment intensity, and measurements with three additional PWA with comorbid cognitive deficits. Two out of the three participants showed improvements in all the mentioned measurements. Of note, these studies do not present statistical analyses beyond the examination of the difference between pre- and post-training performance. In a third study, Manheim et al. (2009) trained 20 PWA with the same number of scripts and treatment intensity, but a different outcome measure: the Communication Difficulty subscale of the Burden of Stroke Scale (BOSS; which assesses the physical, cognitive, and psychological burdens of stroke as reported by the patients; Doyle, McNeil, Hula, & Mikolic, 2003; Doyle et al., 2004). There was a significant decrease in the subscale score post-treatment in comparison with the pre-treatment ($\beta = -6.79$, SE = 3.19).

Two subsequent studies, one from Lee et al. (2009) and one from Cherney, Patterson, and Raymer (2011) reported on the script performance of some of the participants from the Manheim et al., (2009) study. The study from Lee et al. (2009) analyzed data from participants with non-fluent aphasia from different sources (n = 14 from Manheim et al., n = 1 from Cherney et al., and n=2 from Cherney and Harper). Lee et al. (2009) described an overall improvement in the production of script-related words, of approximately 45%. In addition, there was a significant correlation between the amount of time in treatment and improvement for both content ($r = 0.67$, $p<.01$) and rate ($r = 0.53$, $p<.05$), which suggests that dosage of intervention is an important factor for script training.

An additional study trained a group of 8 PWA using the AphasiaRx™ software. The results were described in different publications: one reports on script performance using either high or low cues (Cherney, Kaye, & van Vuuren, 2014); one describes the software development (Van Vuuren & Cherney, 2014), and one examines the outcomes of script personalization (Cherney, Kaye, Lee, & Van Vuuren, 2015). The first and third paper results will be discussed below; the second paper will not be considered as it is a software development technical paper. Cherney et al. (2014) reported a crossover study in which participants were trained in 1 script in each condition (high versus low cues). The high-cues condition included multimodality cues with spoken words, visual support from the virtual therapist, and written words. The low-cues condition provided only the written sentence. The results showed that percent accuracy in script production improved from 50% in the baseline to 77.8% (SD=26.4) at the end of training. There were no significant differences between the high-low cue conditions. However, effect sizes for accuracy were larger for the high-cue condition than for the low-cue condition.

One of the advantages of script training is that it can be adaptable to PWA's communicative needs. Cherney et al. (2015) reported the impact of script personalization using scripts that either contained personally relevant items (e.g., in a script about ordering in a restaurant, scripts contained the name of their favorite dish) or generic items. There was a significant increase in the accuracy of script-related words in post-treatment for both personally related items (mean gain of 29.8%, $p<.001$) and for generic items (mean gain of 21.6%, $p<.005$) with no significant differences between them. However, effect sizes were larger for personally related items (Cohen's d=0.9) than for generic items (Cohen's d=0.25). The authors acknowledged a methodological limitation given that the data analyzed was derived from the examination of the use of cues in script training and not from a direct examination of personal relevance.

Finally, Nobis-Bosch et al. (2011) employed a different technological intervention program by using a barcode reader called B.A.BAR. The device reads linguistic information from barcodes that are associated with linguistic output at the levels of words, phrases, sentences, and texts (e.g., after scanning an item, the device reads back the content that has been stored for that particular item). This crossover study involved 18 PWA pair-matched according to gender, age, time post-onset of aphasia, performance on the AAT, aphasia severity, and aphasia type. The members of each pair were assigned to either the B.A.BAR dialogue condition or a non-linguistic condition with visual–cognitive exercises. In the B.A.BAR condition, participants were trained in 48 scripts of three turns each. In the visual-cognitive condition, participants responded to visual matching of a part to the whole, comparing two pictures to find differences, and alike activities, for the same duration as the B.A.BAR condition. After completing training in one condition, there was a washout period of four weeks and then participants crossed to the opposite experimental condition. Nobis-Bosch et al. (2011) reported significant improvements in the targeted sentences; there were no improvements in linguistic or communication abilities with the nonlinguistic visual training.

In summary, several studies have examined script learning in PWA, all of which have reported the acquisition of the targeted scripts. Generalization is highly variable with no evident trends. Most of the evidence of script training comes from learning sentences through a software; the reports include increases in script related content words, in the rate, and in the fluency of production of the scripted sentences. Computerized treatment is a viable form for learning scripts and is growing as a treatment delivery option since it is cost-effective. Studies that compare scripting approaches have not been completed, therefore it is unknown if one approach yields superior results in comparison to the others. Finally, the trained scripts have included monologues and dialogues that vary greatly in length, content, and complexity; hence, there are no guidelines about the optimal conditions for script learning to be successful. Nevertheless, learning scripts opens an opportunity for participating in conversation and daily situations through rehearsed sentences.

2.1 Script Training Open Questions and the Current Study

This section presents open questions of the script training literature and how the current study aims to address them. Specifically, as this study focuses on mechanisms and characteristics associated with scripted-sentence learning response, this section centers on how variables like stimuli complexity and amount/distribution of practice are controlled in the current design. Script training targets sentence-level information, but there is no certainty about how variability in the scripted stimuli has affected treatment and generalization responses. There has been variability in the number of trained and untrained scripts with a typical range between 2 and 20 scripts. There has also been variability in the number of words per script ranging between 13 and 173 in monologues and dialogues. Similarly, variables such as length, content, complexity, and sentence type are yet to be investigated. The lack of research in stimulus-related variables (the number of sentences, words, turns, and overall sentence properties) makes it unclear how changes in stimulus variability have impacted sentence learning. The current study focuses on *scripted sentences* as a minimal learning unit for script training, and it controls the number of sentences, the number of words, and linguistic properties. Concentrating on scripted sentences, instead of monologues or dialogues, is a step forward in controlling the characteristics of the stimuli in which script learning takes place.

Another important open question in the script training literature is the determination of optimal dosage levels in terms of practice time, distribution of practice, and number of sessions. The patterns of practice times reported by the computer logs are highly variable. For example, in the study from Lee et al. (2009), there was a range of practice from 2 hours to more than 16 hours per week, when participants were required to practice 3.5 hours per week. Thus, it is unknown how changes in the amount of training or the practice's distribution influence scripted-sentence learning. The current study does not manipulate different dosage levels; instead, it controls for this variable while examining individual characteristics associated with scripted-sentence learning. Specifically, the methodological design provides sufficient opportunities for practice, based on previous script training studies, while carefully controlling and measuring the amount of practice. All participants received the *same* amount and distribution of practice, five two-hour sessions, distributed over two weeks.

Overall, the current study uses a protocol of scripted-sentence learning adapted to Spanish speakers with aphasia. This adaptation contributes to the evidence-based knowledge for a large and currently understudied portion of the world population, Spanish speakers with aphasia. It's important to note that a recent study reported a successful application of script training on a bilingual Spanish-English speaker with aphasia (Grasso, Cruz, Benavidez, Peña, & Henry, 2019). However, the application of script training in monolingual Spanish speakers is yet to be documented. Beyond contributing to the script training body of evidence, the current study mainly examines the relationship between script training *treatment mechanisms* (i.e., how it works) and *individual characteristics of PWA*, at the behavioral and neuro-electrophysiological levels, that are associated with scripted-sentence learning (i.e., for whom it works).

3.0 The Role of Rhythmic Features of Speech Entrainment in Scripted-Sentence Learning.

3.1 Key components of script training

To begin examining script training mechanisms (i.e., how this treatment works), the following paragraphs describe the process by which PWA have learned scripted sentences. The majority of script training studies have employed the following components as part of the script learning process: listening, repetition, choral reading (also known as speech entrainment), and independent production. A standardized procedure for script training has yet to be developed; therefore, every study has the potential to be based on a set of different but related components. It is worth noting that multiple studies have investigated the listening, repetition, and independent production components of aphasia treatments (e.g., Conroy, Sage, & Lambon Ralph, 2009c). However, few studies have explored the potential importance of speech entrainment as a critical treatment component, as will be further discussed below.

The listening, repetition, and independent production components of script training are rooted in errorless learning principles (for applications of errorless learning in aphasia see Fillingham, Hodgson, Sage, & Lambon Ralph, 2003). The central claim is that by listening first and then correctly repeating a target stimuli multiple times, errors can be avoided (hence the label "errorless"). Errorless learning is based on Hebbian concepts in which a *correct or incorrect* repetition can create an association that increases the likelihood of those responses being repeated in the future (Hebb, 1949, 2005). Therefore, during the learning process, the goal is to strengthen the desired listening-target association by reducing errors until a correct independent production is achieved.

Moreover, the components of *listening* and *repetition* are ubiquitous in aphasia treatments, including script training. Although most of the studies have focused on the word level (e.g., Conroy et al., 2009c; Fillingham et al., 2003; Fillingham, Sage, & Lambon Ralph, 2005a, 2005b, 2006) rather than on the sentence level, the aphasiology field has made considerable progress in understanding how listening and repetition may influence learning in aphasia. For example, a growing body of literature has examined the successful use of errorless repetitive learning, effortful retrieval and its balance with repetition (e.g., Middleton, Schwartz, Rawson, & Garvey, 2015; Schuchard & Middleton, 2018b), and the critical role of semantic and phonological cues in maximizing effort while maintaining successful repetition (e.g., Wambaugh, 2003; Wambaugh, Doyle, Martinez, & Kalinyak-Fliszar, 2002). Given that the listening and repetition components have been more extensively investigated and that the current study examines scripted-sentence learning, the next paragraphs will focus on a generally overlooked component in script training, speech entrainment.

Speech entrainment refers to the unison production of sentences by patient and clinician (Fridriksson et al., 2012). This speech synchronization between patient and clinician may be a key mechanism for learning scripted sentences. Before exploring the specifics of the rhythmic features that allow this synchronization, it is important to note that the word entrainment has been used to describe different phenomena in relation to synchronization between two or more signals. Entrainment can refer to i) *conversational entrainment,* the alignment that takes place between partners engaged in conversation (Borrie, Lubold, & Pon-Barry, 2015); ii) *speech entrainment*, the unison production of sentences by patient and clinician (Fridriksson et al., 2012); iii) coordinated rhythmic movement in response to an external stimulus (e.g., clapping, dancing, walking to music; Phillips-Silver et al., 2010); iv) *cortical entrainment* (a.k.a cortical tracking), the coupling between low-frequency brain responses and the speech rhythm as reflected in the speech envelope (Ding & Simon, 2014); and to other phenomena in the fields of physics and biology.

Although the emphasis of this study is on *speech entrainment* (as a potential mechanism of action for scripted sentence learning) and in *cortical tracking/entrainment* (as an individual characteristic associated with script sentence learning), the following paragraphs briefly describe *conversational entrainment*, an inherent element of communication. As a preview, Aim 1 provides details about speech entrainment using rhythmic enhancements in word stress and metronomic beats (sections 3.4 and 3.5), and Aim 2 describes the specifics of cortical tracking (section 5.2). Figure 1 represents the different types of entrainment mentioned above. The current study focuses on the types highlighted in gray.

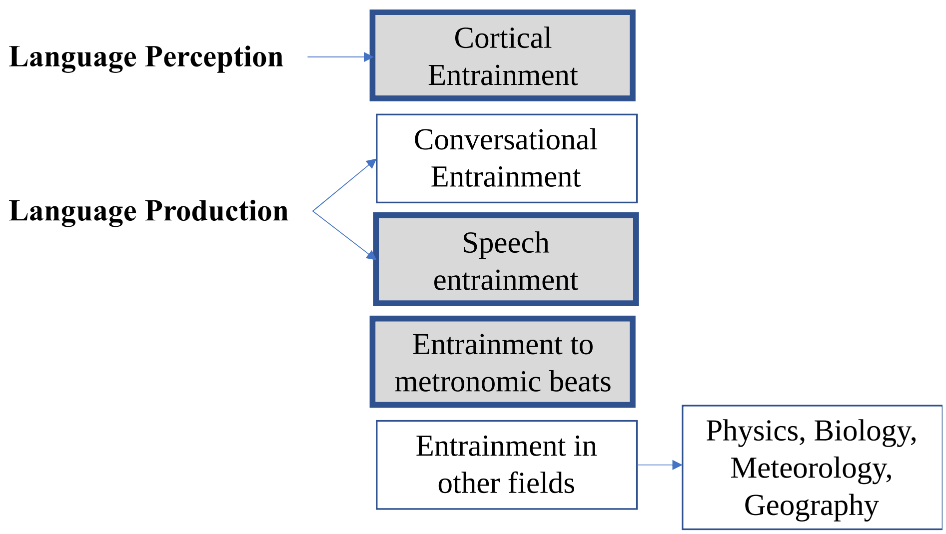

Figure 1 Types of entrainment.The shadowed and outlined boxes will be measured or manipulated in the present study.

3.2 Conversational Entrainment

Conversational entrainment describes processes that occur when two people are engaged in a dialog. Speakers modify their verbal and non-verbal behaviors to align with their communication partners by coordinating syntactic structures (e.g., Branigan, Pickering, & Cleland, 2000), lexical choices (e.g., Brennan & Clark, 1996), imitating phonetics (e.g., Babel, 2012), and also modifying gestures and facial expressions (e.g., Louwerse, Dale, Bard, & Jeuniaux, 2012). Conversational entrainment happens continually in conversations and plays an essential role in social communication by creating affiliation, rapport, and enhancing familiarity with the interlocutor (Gueguen, Jacob, & Martin, 2009). While not the focus of this study, conversational entrainment does affect communication success while interacting with others and therefore is important to be reviewed here.

Conversational entrainment can be used as a treatment strategy to improve PWA's language performance by taking advantage of the alignment between conversational partners at multiple linguistic and paralinguistic levels. Despite its potential influence on communication, few studies have examined conversational entrainment in individuals with communication disorders. It has been reported that adults modify their speech rate and their pitch variation to match people with dysarthria (Borrie & Liss, 2014). Also, dialogues between dyads of a neurotypical individual and a person with dysarthria showed less parallel modulations of prosodic information, less similarity of pitch features at each turn boundary, and an indication of less acoustic entrainment in comparison with a dyad of neurotypicals (Borrie et al., 2015).

Recent efforts have been made toward implementing a computational methodology to capture conversational entrainment, with potential clinical applications (Borrie, Barrett, Willi, &

Berisha, 2019). In its current stage of development, this methodology can analyze conversational entrainment at the rhythmic, articulatory, and phonatory dimensions of speech, and has been validated in neurotypical individuals. Future applications may include additional linguistic and paralinguistic elements of conversational entrainment and may be validated in individuals with communication disorders. These applications could increase our current knowledge of conversational entrainment and shed light on using the alignment between conversational partners to support communication in PWA. The type of alignment seen in conversational entrainment may improve lexical retrieval, syntactic encoding, and communicative success for PWA. Despite its potential importance for communication, conversational entrainment in the aphasia population has yet to be explored.

As mentioned above, conversational entrainment is an inherent element of communication directly related to synchronization between speakers, which is why it was briefly reviewed in this section. However, the emphasis of this study is on speech entrainment (Fridriksson et al., 2012), as a potential mechanism for script sentence learning. The next section will review the current literature about speech entrainment as used in aphasia rehabilitation.

3.3 Speech Entrainment in People with Aphasia

S*peech entrainment*, also called choral reading in the script training literature, refers to the unison production of sentences by a patient and clinician. Speaking in unison can only take place on the basis of rhythm (as detailed below), and this synchronization may provide the necessary support for PWA to learn scripted sentences. Fridriksson et al. (2012) explored speech entrainment in PWA by using audio-visual feedback. In this study, 13 PWA learned ~1-minute scripts using a speech entrainment audio-visual condition; a speech entrainment audio-only condition; and spontaneous speech, a control condition which consisted of talking about a script topic. By comparing the number of different words per condition, Fridriksson et al. (2012) found that the speech entrainment audio-visual condition elicited a larger variety of words (mean=66.04%, SE=4.36%) compared with both the speech entrainment audio-only condition (mean=40.79%, SE=4.22%, $t(12)=3.62$, $p<0.004$) and the spontaneous speech control condition (mean=29.99%, SE=5.00%, $t(12)=4.30$, $p<0.001$). Both entrainment conditions positively influenced the number of script-related words that PWA produced in comparison with the control condition. The same participants (n=13) completed a 6-week computerized program in which they practiced three scripts for two weeks using a speech entrainment audio-visual condition. The authors found an increase in the number of different words at one week after treatment, but not at six weeks. Overall, this study provided evidence for using speech entrainment as a tool to increase sentence-level output in PWA by relying on both auditory and visual information.

In a follow-up study, Fridriksson et al. (2015) aimed to identify cortical damage patterns that predicted fluency outcomes achieved with speech entrainment. The study included 44 stroke survivors, 32 of which were diagnosed with aphasia according to the Western Aphasia Battery (WAB; Kertesz, 1982). Participants performed a picture description task and also learned short scripts using speech entrainment. The main neurological finding was that people with damage in the inferior frontal gyrus (pars opercularis and triangularis) benefited more from the fluency effects provided by speech entrainment than participants with damage in other regions of the left hemisphere. In their behavioral results, 25 participants (23 with aphasia and 2 without aphasia) produced a higher number of different words per minute during the speech entrainment task when compared to a picture description task. Only participants with non-fluent aphasia showed a significantly higher number of different words learned during the speech entrainment relative to the picture description task. Overall, this study provides further evidence of the positive effects of speech entrainment when learning scripted sentences, and it indicates that people with non-fluent aphasia may benefit more from treatment using speech entrainment than other aphasia patients.

In sum, Fridriksson et al. (2012, 2015) provided evidence of the beneficial use of speech entrainment in aphasia. These findings raise two further questions: how does speech entrainment work, and why is it beneficial for scripted-sentence learning in PWA? Based on the Hierarchical State Feedback Control model (HSFC; Hickok, 2014), Fridriksson et al. (2015) suggested two possible mechanisms of action regarding the positive effects of speech entrainment. The first possibility is that speech entrainment works at the somatosensory motor circuit level, a lower-level system that supports the production of individual phonemes by providing multisensorial phonemic support that enhances speech production. In other words, speech entrainment supports the visual and auditory perceptual system associated with phonemic proprioception facilitating speech production. A second possibility is that speech entrainment works at the higher syllable-level system by providing auditory-visual syllabic information that can be mapped onto articulation. These two neurological accounts are detailed in the Fridriksson et al. (2015) and Hickok (2014) publications. *However, neither of these hypothesized mechanisms account for the simultaneous production of speech, which is a critical feature of speech entrainment*. Thus, the unique contribution of entrainment, as opposed to tasks such as repetition with visual and auditory feedback, is still understudied.

In terms of simultaneous speech production, it is important to note that entrainment critically relies on detecting, integrating, and producing rhythmic features (Phillips-Silver et al., 2010). In other words, *simultaneous speech production takes place based on rhythm*, which can come from speech rhythm itself or other types of rhythm, such as metronome beats. Moreover, the rhythmic features that underly speech entrainment have a facilitatory effect on sentence learning in aphasia. Two hypotheses regarding this facilitatory effect are: i) unison production of scripted-sentences enables PWA to align to word stress, supporting lexical retrieval (e.g., Kimelman & McNeil, 1987; Soto-Faraco et al., 2001). ii) unison production of scripted-sentences enhances sentence memorization via chunking (e.g., Purnell-Webb & Speelman, 2008). Evidence in support of these hypotheses will be considered below.

3.4 Word Stress in Language Comprehension and Production in People with Aphasia

As argued in relation to the first hypothesis above, unison production of sentences by patient and clinician —*speech entrainment*— enables an alignment to the rhythm of speech, which in turn can support lexical retrieval. Two main assumptions motivate this claim. The first one is that speech has rhythmic properties over which speech entrainment can take place. The second assumption is that these rhythmic properties are related to lexical access. To support these assumptions, the next paragraphs will briefly describe the broad notion of speech rhythm, its taxonomy, and the use of rhythm —evident in word stress— for lexical activation, particularly in the Spanish language. These descriptions will help to explain how speech entrainment may support lexical access in aphasia.

3.4.1 A Note on Speech Rhythm

Speech rhythm refers to the "hierarchical organization of temporally coordinated prosodic units" (Cummins & Port, 1998, p. 145). The rhythm present in speech helps to i) bind events into organized sequences under a superordinate prosodic cycle; and to ii) coordinate multiple movements, timings, and processes necessary for speech production (Cummins & Port, 1998). Speech rhythm is conceptualized in terms of duration, spacing between elements, and their relative intensity/prominence (e.g., alternation of stressed and unstressed syllables). A quest for a definition of rhythm in speech has been tied to developing a taxonomy that can help to catalog languages based on rhythm. Although this taxonomy has been investigated for decades, there is still disagreement about language-specific rhythmic classifications (see Nolan & Jeon, 2014).

Multiple models have been proposed for how to characterize the rhythmic structure of the world's languages. In early studies, Pike (1945) classified speech rhythm into two types, syllable-timed (e.g., Spanish) and stress-timed (e.g., English; Pike, 1945). A third category, mora-timed, was added some decades later (Abercrombie, 1982) to include languages like Japanese. Two main measures are still in use to classify rhythm in speech, the duration of vocalic/consonantal intervals (%V, %C, and their standard deviation ΔV, ΔC; Ramus, Nespor, & Mehler, 1999); and the Pairwise Variability Index of successive acoustic-phonetic intervals (PVI; Grabe & Low, 2002). Using these measures, prototypical languages are classified as expected (e.g., Spanish is classified as syllable-timed). However, these measures have yielded conflicting classifications for the same language, particularly for non-prototypical languages (Arvaniti, 2012; Grabe & Low, 2002) with a classification success around 33% (Arvaniti, 2009). In addition to the weak reliability of the measures, the elicitation method and the selection of testing materials significantly affect rhythmic classifications. Measures that capture rhythm in its multidimensionality are being developed, including enveloped-based rhythmic analysis (Tilsen & Arvaniti, 2013), which will be discussed in the cortical tracking section (section 5.2).

On a parallel view to the classification of rhythm in speech in discrete categories, Dasher and Bolinger (1982) argued that the rhythm in speech could be placed on a continuum. This continuum has a gradient scale that ranges from syllable–, stress– and mora-timed languages (Cummins, 2002), and it may be the result of language-specific phonetic and phonological properties (Ramus et al., 1999). For example, vowel duration and intensity carry accentual information that can help identify a syllable as stressed or unstressed, and these properties vary widely across languages.

Returning to the first assumption, that speech has rhythmic properties over which speech entrainment can take place, the literature briefly described above reveals that speech indeed has prominent rhythmic properties, which have been investigated for decades. These properties seem to be language-specific, and efforts to taxonomize them have moved from discrete categories to a multidimensional view of rhythm placed on a continuum. The current study does not contribute to the definition of rhythm in speech, nor does it clarify its taxonomy. Instead, it focuses on the enhancing contrastive rhythm provided by word stress and its role in lexical retrieval in aphasia.

In summary, in the current study, speech rhythm is understood from the perspective of contrastive rhythm, the alternation of strong and weak elements (White, Mattys, & Wiget, 2012). This view of rhythm is suitable for the Spanish language in which the alternation between strong (stressed) and weak (unstressed) syllables is highly marked. In other words, for the purpose of this study, rhythm in speech is viewed from the perspective of word stress, which can be defined in terms of the prominence that a syllable receives in comparison with other syllables within the same word. This prominence is generally related to longer durations and intensities than their unstressed counterparts (Hualde, 2012).

3.4.2 The role of word stress in lexical access in Spanish

This section begins with a brief introduction to the lexicostatistical differences between English and Spanish. Next, it describes word stress assignment in Spanish, followed by the acoustic characteristics (pitch, duration, and intensity) that lead to the perception of word stress in this language when contrasted with English. Finally, it examines the role that word stress plays in lexical access in Spanish.

Regarding lexicostatistics, a study from Cutler, Mister, Norris, and Sebastian-Gallés (2004) analyzed data in both Spanish and English using real (naturally-occurring) speech and linguistic databases. In terms of *number of phonemes*, they reported that Spanish has a twenty-five-phoneme inventory while English has a forty-four-phoneme inventory. A language with fewer phonemes can use a couple of alternatives to cover the same number of lexical concepts as a language with more phonemes. One option is homophony, in which the same word form can have multiple meanings. Another option is to have longer words, which seems to be the case in Spanish. For example, Cutler et al. (2004) reported that the mean *word-length* (measured in number of phonemes) is 8.3 in Spanish and 6.94 in English. Their estimates of real speech revealed a mean word-length of 4.62 in Spanish and 3.54 in English. In terms of *number of syllables*, Spanish has a mean of 3.48 syllables per word, while English has a mean of 2.72 syllables per word. In their real speech estimates, Spanish had 2.02 syllables per word and English 1.42 syllables. Overall, Cutler (2012) concluded that English has a large phonetic inventory to build an extensive set of short words; in contrast, Spanish has compensated for a comparatively small phonetic inventory using longer words.

Having longer words in Spanish is relevant for word stress assignment. Hualde (2012) reported that, in Spanish, stress is placed on one of the last three syllables on lexical words. Words with stress on the last syllable are called oxytonics (e.g., ala**crán** 'scorpion'); words with stress on the second to last syllable are called paroxytonics (e.g., cucu**ra**cha 'cockroach'), and words with stress on the third to last syllable are called proparoxytonics (e.g., li**bé**lula 'dragonfly'). Overall, there is no word stress assignment further back from this "three-syllable window" (Hualde, 2012). In terms of the frequency of these stress assignments, Hualde (2012) reports that most words ending in a consonant are oxytonic, and most words ending in a vowel are paroxytonic. In fact, approximately 95% of nouns and adjectives follow these stress rules. Adverbs also follow this rule, although some frequent adverbs ending in vowel are oxytonics (e.g., a**quí**, a**llá**). Verbs follow different rules depending on the tense: some verbs in the present tense are paroxytonics, and some verbs in the future tense are oxytonics. In sum, word stress in Spanish is assigned in a "three-syllable window" forming oxytonic, paroxytonic, and proparoxytonic words that vary in frequency according to their word types (e.g., nouns, verbs, adverbs).

Beyond lexicostatistics and word stress assignment, acoustic characteristics (pitch, duration, and intensity) lead to the perception of syllabic prominence —word stress— in Spanish. Among these, *pitch* has been shown to be a reliable marker of stressed syllables in declarative sentences (Ortega-Llebaria, 2006). *Duration* has also been shown to be a consistent correlate of syllabic stress in Spanish. For instance, speakers tend to produce longer durations in stressed syllables than unstressed syllables (Llisterri, Machuca, de la Mota, Riera, & Ríos, 2003; Llisterri, Machuca, de la Mota, Riera, & Ríos, 2005; Mora, 1998; Ortega-Llebaria & Prieto, 2011). In terms of *intensity*, some studies have shown its acoustic correlation with the perception of syllabic stress in Spanish while others have shown a lesser, or inconsistent, effect across subjects (Delattre, 1969; Nadeu, 2014; Ortega-Llebaria, 2006; Ortega-Llebaria & Prieto, 2011).

Using acoustic characteristics (pitch, duration, and intensity), Ortega-Llebaria, Gu, and Fan (2013) compared stress perception of Spanish sentences in a group of native English-speakers with extensive experience speaking Spanish and a control group of native Spanish-speakers. Declaratives and reporting clauses were used as intonational contexts in which target words were embedded. The target words were digitally manipulated to create variations in stress acoustic characteristics (pitch, duration, and intensity). A native Spanish-speaker recorded the sentences, including the target words. After each trial, participants were asked to indicate whenever a target word was perceived as an oxytone (i.e., having lexical stress in the last syllable). It was found that frequency and duration had a significant effect on subjects' stress perception. Moreover, interactions between frequency and duration with group indicated that stress is not perceived identically between English and Spanish speakers. When perceiving stressed syllables, *duration* has a stronger effect on English speakers than for Spanish speakers. In contrast, *pitch* has a stronger effect on Spanish speakers than on English speakers. *Intensity* did not show an effect on subjects' stress perception nor an interaction with group. Overall, when listening to the same stimuli, English and Spanish speakers have different perceptions of stress depending on acoustic characteristics (Ortega-Llebaria et al., 2013; Romanelli & Menegotto, 2015).

The cross-language differences in stress perception could be attributed to speakers using stress cues available in their native language to identify stress patterns in another language. Thus, English speakers would show difficulties perceiving stress in Spanish stimuli because they use characteristics relevant to the perception of stress in their native language (Ortega-Llebaria et al., 2013; Romanelli & Menegotto, 2015). For example, in English, segmental cues (e.g., a full vs. a reduced vowel) are more critical for word recognition than suprasegmental cues (e.g., stressed vs. unstressed syllables). In support of this claim, Fear, Cutler, and Butterfield (1995) conducted a study presenting words to English speakers that had their initial vowel sounds switched. For example, the initial vowels were switched in the words: *audience* (initial vowel receives primary stress), *auditorium* (initial vowel receives secondary stress), *audition* (initial vowel is unstressed but not reduced), *addition* (initial vowel is unstressed and reduced). In this study, Fear et al. (1995) asked participants to rate how natural the modified words sounded compared to the original word. A change in stress was rated as not different from the original word; however, a change in segmental information (e.g., in the *addition* example) was rated as significantly less natural. Overall, English speakers found that a reduced vowel (i.e., segmental information) was more important for rating a word as "natural" than changes in word stress (i.e., suprasegmental information).

In Spanish, however, unstressed vowels are never reduced (i.e., stressed syllables only change suprasegmentally and not segmentally). Therefore, Spanish allows the examination of suprasegmental information independent of segmental information. For example, the words **ca**so [*case*] and ca**só** [*married*] have the same segmental structure, with no phonetic variation in the consonants or vowels, and only differ in the suprasegmental stress-related information (Cutler, 2012). Therefore, Spanish-speakers make optimal use of suprasegmental cues to distinguish between words, but as seen above, English-speakers rely more on segmental cues for word recognition. Overall, *stress cues for Spanish-speakers are crucial for reducing word candidates and accessing lexical information*; these cues are less critical for English-speakers (Cutler, 2012; Ortega-Llebaria, Olson, & Tuninetti, 2019).

Soto-Faraco et al. (2001) argued that current speech processing models have failed to include the role of suprasegmental information (e.g., word stress) in lexical activation. The failure to include this information is because most models are based on English speakers' data, where it is likely that lexical stress does not play a strong role in lexical retrieval (Cutler & Carter, 1987). To begin to understand the role of suprasegmental information in lexical access, Soto-Faraco et al. (2001) conducted priming experiments in Spanish. As seen above, words in this language have one syllable marked for primary stress, and stress/unstressed syllables do not differ in their segmental information. For example, sa**ba**na (savannah) vs. **sa**bana (sheet), or **ca**so (case) vs. ca**so** (third-person singular past tense of the verb "to marry") vary in their stress patterns, but not in their segmental information. As mentioned above, having identical segmental information and variations in the suprasegmental information makes Spanish a suitable language to explore the independent role of suprasegmental information in lexical access.

In their three priming experiments, Soto-Faraco et al. (2001) had participants listening to sentences ending with word fragments (e.g., **pr**inci-) after which they had to make lexical decisions (e.g., **pri**ncipe "prince", prin**ci**pio "beginning"). Different groups of subjects participated in each experiment. The first experiment was designed such that word fragments either i) matched in their stress patterns with the targets (e.g., **prin**ci and **prin**cipe); ii) mismatched in their stress patterns with the targets (e.g., prin**ci** from the word prin**ci**pio and **prin**cipe); or iii) were control items (e.g., mos from the word mos**qui**to and **prin**cipe). A significant effect of prime type was found, such that speed and accuracy were higher for matching primes than mismatched and control primes. This result was interpreted as demonstrating that Spanish listeners use stress to access lexical information. When items are segmentally identical, word stress provides a lexical retrieval advantage to a word candidate that matches the target in stress information.

In their second experiment, stress patterns were held constant while segmental information (a single vowel sound) was manipulated. For example, the word fragment either i) matched segmentally (e.g., aban- from the word aban**do**no “abandonment” for the target aban**do**no); ii) mismatched segmentally (e.g., abun- from the word abun**dan**cia “abundance” for the target aban**do**no); or iii) were control items (e.g., e- from the word "e**las**tico" for the target aban**do**no). Again, a significant effect of prime type was found such that fragment primes that matched the onset of the targets speeded response times. Speed and errors indicated inhibition for targets preceded by vowel-mismatching prime fragments. Overall, the results were interpreted as showing that Spanish listeners use segmental information similarly to stress information for lexical activation.

The third experiment used consonant matches/mismatches rather than vowel matches/mismatches. The consonants varied in either i) a single feature (e.g., pa**ti-** from the word pa**ti**lla "watermelon" vs. pa**pi-** from the word pa**pi**lla "pap" for the target pa**pi**lla); in ii) multiple features (e.g., bofe- from the word bofe**ton** "slap" vs. bole- from the word bole**tin** "bulletin" for the target bofe**ton**); or iii) were control items (e.g., ce- from the word ce**ne**fa "valance" for the target pa**pi**lla vs. gav- from the target gavi**lan** "hawk" for the target bofe**ton**). A significant effect of prime type was found such that fully matched primes facilitated lexical decisions whereas mismatches generated inhibitory responses. This result was interpreted as showing that Spanish listeners use segmental information from consonants to access lexical information irrespective of whether the mismatch involves a single phoneme or multiple phonemes.

In sum, Soto-Faraco et al. (2001) showed that matching primes facilitated lexical decision responses in both suprasegmental and segmental manipulations. This supports the claim that Spanish speakers can use suprasegmental (i.e., word stress) cues to access lexical information.

Comparable results to the Soto-Faraco et al. (2001) study were found in Dutch speakers (Cooper, Cutler, & Wales, 2002) such that Dutch speakers use stress to reduce the set of lexical alternatives. As seen above, listeners use all available cues in their language to access lexical information (Ortega-Llebaria et al., 2013). In the case of Spanish, pairs of words can be segmentally identical but suprasegmentally different; therefore, Spanish speakers use this information to distinguish between words and access lexical information (Baqué, 2017). Moreover, the contribution of segmental and suprasegmental information for lexical access seem to be equivalent in Spanish. However, in English segmental information is more relevant. Given the cross-language differences described above, and the fact that research on the facilitatory effects of rhythm and word stress to PWA has been studied mostly in English, it is critical to assess the effects of rhythm and word stress in Spanish speaking PWA.

Unfortunately, studies related to the effects of word stress on language processing *in PWA* are non-existent in Spanish. Consequently, the next section examines the role of word stress in English speakers with aphasia. These reports are not optimal given the language-specific evidence provided above, but they provide confirmation of the use of speech rhythm, evidenced in word stress, in aphasia treatment. Motivated by the evidence described above demonstrating that word stress plays an important role in lexical access in Spanish, the current study uses a protocol adapted for Spanish-speakers with aphasia to explore how the unison production of scripted-sentences enables PWA to align to word stress, supporting lexical retrieval (e.g., Soto-Faraco et al., 2001).

3.4.3 Word Stress in Language Comprehension in Aphasia

Prior to describing the limited evidence to date regarding the role of word stress (in particular, emphatic stress) in facilitating language performance in aphasia, it is first necessary to describe in further detail how word stress impacts language processing in English, the language in which these aphasia studies have been conducted. Cutler (1989) identifies two primary levels of stress: *Lexical prosody*, which refers to the stress patterns at the word level; and *metrical prosody*, which refers to the rhythmic patterns of strong and weak syllables observed on longer chains of speech, such as in sentences. Importantly, metrical prosody is built on lexical prosody. The current study examines beats highlighting the existing *metrical prosody* of sentences. These beats may also align with and thereby accentuate the existing *lexical prosody* of individual words within those sentences. This alignment between metrical and lexical prosody is important to understanding how stress may impact comprehension in English, in particular in connected speech (sentence) contexts.

A crucial role of metrical and lexical prosody is the segmentation of continuous speech into meaningful linguistic units, a critical step before lexical access can occur. Lexical access is thought to operate in discrete units; therefore, segmenting continuous speech provides individual lexical units from which PWA can more easily retrieve meaning. However, it is important to note that *speech segmentation strategies are language-specific, as those depend on syntactic rules, phonetic rules, and the distributional rhythmic patterns of each language* (Cutler, 2012). Based on rhythmic patterns, Cutler (1990) posited the Metrical Segmentation Strategy (MSS), later known as the Rhythmic Segmentation Hypothesis. This hypothesis's main idea is that the speech rhythm that characterizes a given language will be used for speech segmentation. In other words, speakers will use the particularities of their language's rhythm to mark word boundaries and then access lexical information. Evidence regarding speech segmentation through word stress in Spanish was reviewed in section 3.4.2 above. See also Peñaloza et al. (2015) for an example of speech segmentation in aphasia using an artificial language.

Cutler (1989) indicated that segmentation strategies are built on the concept of metrical prosody. As an example, in English and other Germanic languages, most words tend to have a trochee strong-weak stress pattern (i.e., **tab**le, **win**dow, **so**fa, **pa**yer, **men**tor) rather than an iambic one (Chi**nese**, pa**yee**, men**tee**). This frequent strong-weak stress pattern, or trochaic bias, guides the segmentation of a speech chain by setting word onsets where stressed syllables are found. Evidence in this regard comes from misheard sentences where the assumption that a stressed syllable is part of a new word onset, leads to the perception of a new word (e.g., the target spoken sentence "it was il**le**gal" is perceived as "it was an **ea**gle") (Cutler & Norris, 1988). Generally, mis-segmentations of speech in English are more likely to involve stressed syllables being erroneously interpreted as word-initial stress. These findings suggest that listeners often treat metrical prosody, the presence of strong syllables, as being aligned with lexical prosody, and then exploit this alignment to identify word onsets, thereby facilitating word recognition and retrieval.

Overall, the trochaic bias plays a role in speech segmentation in English and other Germanic languages. That is, English speakers take advantage of the high-frequency strong-weak pattern to identify word boundaries in continuous speech. This strategy facilitates subsequent lexical access. Thus, even though word stress is less important for lexical access in English than it is in Spanish (e.g., Soto-Faraco et al., 2001), it can nonetheless be of benefit for identifying words and accessing lexical representations, particularly in sentence contexts. This reasoning suggests that cues which highlight stressed syllables might be beneficial for people with aphasia and in aphasia rehabilitation. However, the evidence regarding the use of stress in aphasia rehabilitation is non-existent in Spanish. In English, there is some evidence in favor of the benefit of cues that highlight strong syllables. This evidence comes from research on the effects of emphatic stress on language performance in English-speaking PWA.

The following paragraphs briefly review those studies and includes the main limitations of this literature. Emphatic stress refers to the strategy in which a word, in a sentence context, is presented as more salient than other words by overemphasizing its stressed syllable (e.g., what did you **do** yesterday? show me the **red** pencil). This strategy, which should facilitate lexical access, is frequently used in clinical practice to improve language processing in PWA. However, i) studies about the use of emphatic stress in language comprehension of PWA are relatively scarce; ii) the studies that have addressed this topic have limitations; and as mentioned above, iii) the evidence of the use of this strategy has come solely from English and might not be applicable to other languages. Specific to Spanish, the emphatic stress strategy does not translate well because in many contexts it is considered ungrammatical (Gutierrez-Bravo, 2002; Zubizarreta, 1998). Therefore, the current study uses beats to make word stress more salient, rather than emphasizing the words from the speech signal itself.

Early aphasiology studies reported that when listening to paragraphs and answering yes/no questions, PWA (n=9) performed better when the paragraphs had emphatic stress relative to natural speech (Kimelman & McNeil, 1987). In contrast, emphatic stress was not beneficial when emphasizing a single target word within the context of paragraphs. PWA (n=10) did not show an advantage in answering yes/no questions regarding an emphatically-stressed target word (Kimelman, 1991) embedded in a paragraph. Although the authors concluded that stressing target words alone does not significantly contribute to auditory comprehension, the context surrounding the target words was digitally manipulated to be prosodically neutral. Generally, prosodic cues accompany (or lead to) a specific emphasized word; by removing these contextual prosodic cues, the beneficial role of salient stress for PWA was also reduced.

Kimelman (1999) examined whether aphasia severity and linguistic complexity were associated with the benefit of emphatic stress on auditory comprehension in PWA. Eight stories were divided into two complexity levels (fifth- to sixth-grade and eighth- to ninth-grade reading levels). Each sentence in the story had a designated target word that was read with either emphatic stress or natural prosody (e.g., Edward Bok was born in a village in the **Holland** countryside. When he was a **child,** he moved to America. In time, he became a famous **author** and made a fortune…). To assess auditory comprehension, each target word had two related yes/no questions (e.g., did he move to America when he was an adolescent?). A target was considered correct if both questions were answered accurately. PWA were divided into groups according to their scores on the Auditory Comprehension subtests of the Boston Diagnostic Aphasia Examination (BDAE; Kaplan, Goodglass, & Weintraub, 2001; Kaplan, Goodglass, Weintraub, & Goodglass, 1983) into mild (n=9), moderate (n=8), and severe (n=9), and an additional control group (n=10). Participants heard the stories in both complexity levels in different sessions.

Overall, auditory comprehension was better in the emphatic stressed condition than natural prosody for all groups across complexity levels. There were no significant differences in auditory comprehension between the high and low levels across participants in terms of complexity. However, a three-way interaction effect (severity, stress, and complexity) showed that the most severe aphasia group benefited more from emphatic stress, in the high level of complexity in comparison with the low. Cutler claims that stress cues, the strong-weak pattern, derived from metrical stress, support segmentation. These cues make it easier to segment words in sentences and access the corresponding lexical representations during comprehension. If Cutler's hypothesis is correct, stress cues, or emphatic stress that makes those stress cues more salient, will be most beneficial when people need to segment potentially hard-to-identify words from running speech, or when people need to use as many available cues as possible (most severe PWA) to process connected speech. The three-way interaction that Kimelman (1999) found is consistent with this view.

In sum, PWA benefited from emphatic stress when answering yes/no questions about information presented in paragraphs (Kimelman & McNeil, 1987). Also, adding emphatic stress on words throughout a passage facilitated comprehension (Kimelman, 1991), supporting that contextual prosodic cues, leading to a target emphasized word, can benefit auditory comprehension in aphasia. Lastly, people with more severe aphasia benefited more from emphatic stress cues when increasing stimuli complexity (Kimelman, 1999). It could be the case that emphatic stress is most favorable in connected speech (when the segmentation process is more difficult); and in cases of more severe aphasia, when all available cues are required to access the lexical representations. Overall, these results are in agreement with Cutler's claims, in that stress cues support the segmentation of connected speech, which facilitates later lexical access that results in better comprehension.

Using a different methodology from the studies described above, an additional study examined lexical access in PWA (n=12) and controls (n=12) performing a dual-task of a yes/no semantic judgment task and a yes/no lexical decision task (Slansky & McNeil, 1997). The stimuli consisted of 160 phrases with the following order: "we saw the X and the Y today". X were nouns presented either with or without emphatic stress, and Y were nouns presented without emphatic stress. Participants had to perform either a semantic judgment on X or a lexical decision task on Y. In the semantic judgment task, PWA were both significantly more accurate and faster when an emphatically stressed X was used, in comparison to when unstressed nouns were used. This finding provides some evidence that emphatic stress supports access to lexical representations in aphasia.

This result also highlights the facilitatory effect of stressed words on task performance in response time and accuracy for PWA. It is also consistent with the yes/no comprehension question results from Kimelman and McNeil (1987).

In the *lexical decision task*, however, PWA were not significantly faster or more accurate when emphatic stress was applied to the preceding semantic judgment target. In contrast, controls were significantly slower and significantly less accurate when emphatic stress was applied to the preceding semantic judgment target (indicating interference effects). The authors' explanation for PWA not being differentially affected by the presence of emphatic stress in the lexical decision task was that PWA have difficulties deactivating a salient stimulus (emphatic stress on X) while having few resources for the (Y second task) lexical decision task, or that the dual-task demand exceeded the processing capacity for PWA. Besides the overload that the dual-task required from PWA, the lexical decision results reported in this study are not informative for the role of emphatic stress on language comprehension. A missing manipulation from the Slansky and McNeil (1997) study was an examination of the role of empathic stressed/unstressed lexical decisions on X (the target word), as they did for semantic judgments. In sum, in a semantic judgment task, PWA were significantly more accurate and faster when the targets were provided with emphatic stress (Slansky & McNeil, 1997). Even though lexical decisions in this study were not examined for words with emphatic stress, the semantic judgment task provides an additional piece of evidence that emphatic stress can support lexical access, including PWA.

An important clarification about the four studies reviewed above is that word stress (syllabic prominence and weak-strong alternations) is present in words and sentences regardless of emphatic stress. Thus, the strategy of emphasizing stress is related to an amplification of existent acoustic characteristics of words and sentences. This amplification reflects changes in speech intensity, frequency, and duration, making acoustic cues more salient and enhancing segmentation and later lexical retrieval in aphasia. However, a critical limitation of these studies is that those acoustic characteristics were not reported or controlled for, making it difficult to understand their role in auditory comprehension in aphasia. Another limitation is that the evidence reported on these four studies comes from a narrow team of researchers, leaving open the need for replication. These studies are also outdated, and new evidence about this topic has not been reported in more than two decades. Lastly, as mentioned above, the rhythmic properties of speech are language-specific, which leaves a need to adapt and test these findings in other languages different than English.

3.4.4 Word Stress in Language Production in Aphasia

Evidence supporting the benefits of emphatic stress for PWA (exaggerated prosodic cues that emphasize metrical and lexical prosody) are not limited to comprehension. A parallel set of evidence indicates that lexical stress benefits *language production* in aphasia, particularly in repetition and naming tasks. In terms of *repetition*, Howard and Nickels (1999) found that most PWA (n=13) had a lower probability of correctly producing a word with an unstressed initial syllable than a word with a stressed initial syllable. It is worth remembering that often English words start with stressed syllables – described above as the trochaic bias. The most common errors reported in the Howard and Nickels (1999) study were either omission of the first unstressed syllable or reduplication of the stressed syllable. These findings were replicated and extended to confrontation *naming* by Howard and Smith (2002), who reported that difficulties repeating words with an initial weak syllable were even more marked in picture naming tasks than in repetition.

Howard and Smith (2002) explored the production of word stress in PWA (n=12) using words and nonwords (to remove lexical representations). They used a disyllabic nonword repetition task with stress either in the first or the second syllable. They found that nonword repetition was worse for disyllables that initiated with a weak syllable compared to a strong one, particularly when there was a "schwa" in the weak syllable. This finding implies that difficulties producing words with weak-strong stress patterns are not the result of an impaired access to lexical representations; however, the nonword repetition task had floor effects for many of the participants; thus the implication is inconclusive.

Howard and Smith (2002) included in their study a naming and a repetition task for three-syllable words. They found that for both tasks, PWA more accurately repeated words that had stress on the first syllable when compared to words with stress on the second syllable. In fact, words with stress on the second syllable yielded the lowest repetition accuracy. The authors also examined how the repetition of disyllabic words was affected by a preceding monosyllabic word or a preceding disyllabic word with a strong-weak stress pattern. Again, accuracy was lower when repeating a word with a weak-strong pattern preceded by a strong-weak disyllable. The experiments by Howard and Smith (2002) showed that words with initial unstressed syllables are likely to be produced in error; however, as described in the discussion of the trochaic bias in Germanic languages in section 3.4.3, nouns with initial unstressed syllables are less frequent in English, creating a parsing bias that makes difficult to access or produce these type of words. The authors established that stress-related difficulties resulted from the *phonological assembly* as defined by the Levelt, Roelofs, and Meyer (1999) model. Overall, these findings suggest that word stress has a relevant role in retrieving lexical-phonological representations for both naming and repetition tasks.

An additional study from de Bree, Janse, and van de Zande (2007) explored stress assignment in PWA through word reading and non-word repetition tasks. In the reading task, Dutch PWA showed preferences for reading with the expected Dutch regular stress compared to irregular stress. During the non-word repetition task, a tendency towards the regularization of the stress pattern was found for both PWA (n=6) and matched controls (n=6). Also, PWA showed a disproportional difficulty repeating non-words with highly irregular stress patterns compared to controls. Although de Bree et al. (2007) reported that their participants did not show signs of motor speech disorders, it is important to note that some studies have hinted that difficulties in producing word stress in PWA is in part due to the presence of apraxia of speech (e.g., Vergis et al., 2014).

An essential diagnostic criterion for apraxia of speech is an impairment in the production of prosody, with a slow and hesitant articulation (Haley & Jacks, 2019). People with apraxia of speech differ in speech duration metrics (e.g., Pairwise Variability Index and Lexical Stress Ratio) compared to PWA without apraxia, and matched controls (Haley & Jacks, 2019). Also, in the production of three-syllable words (in isolation and within a sentence context), people with apraxia of speech (n=9) have lower duration of vowels in stressed words with weak–strong patterns, when compared to PWA (n=8) and controls (n=8) (Vergis et al., 2014). This lower duration was found only in the sentence context and not in the naming in isolation task. There were no significant differences in intensity among the three groups. However, a shorter duration of vowels could have had a negative impact on the lexical stress contrast between weak–strong syllables, consistent with the perceptual prosodic deficits observed in the diagnosis of apraxia of speech.

A review of the particulars of apraxia of speech is out of the scope of the current study (for systematic reviews see Ballard et al., 2015; Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006). However, because of the potential impact of apraxia of speech on prosodic production, participants in the current study will be tested for the presence of motor speech disorders.

In sum, evidence regarding word stress in language production in aphasia shows that PWA have difficulties in the repetition of disyllabic words that initiate with an unstressed syllable (Howard and Nickels, 1999); which are less frequent in the English language. These same difficulties were observed in naming tasks, repetition of nonwords, and naming/repetition of trisyllabic words that initiated with an unstressed syllable (Howard and Smith, 2002). PWA have also been shown to regularize stress patterns when asked to repeat words and nonwords with irregular stress patterns (de Bree et al., 2007). Notably, motor speech disorders may negatively affect PWA's production of stress patterns (Vergis et al., 2014). Overall, the evidence of word stress production in PWA suggests that stress is used to i) retrieve lexical-phonological representations from memory for both naming and repetition tasks; and ii) encode nonword representations in memory, enabling performance in repetition tasks.

Recalling the main findings about empathic stress supporting lexical retrieval in PWA, it was observed that emphatic stress increased auditory comprehension when PWA responded to questions related to a paragraph (Kimelman & McNeil, 1987; Kimelman, 1991). Also, more severe PWA benefited from empathic stress, particularly when increasing stimulus complexity (Kimelman, 1999). Additionally, PWA performed significantly faster and more accurately in semantic judgment tasks when the target words were provided with emphatic stress (Slansky & McNeil, 1997). Overall, the evidence suggests that emphasized stress in the English language supports connected speech segmentation facilitating language comprehension. Connecting these findings with Cutler’s (2005) claim that stress information has an important role in word segmentation, it can be seen that listeners (both neurotypicals and PWA) attend to stressed syllables and use them to segment words from the speech stream, facilitating later access to lexical representations.

Taken together, results from language comprehension and production in PWA indicate that word stress supports lexical access, retrieval of lexical/phonological representations, and encoding of sub-lexical phonological representations for nonwords. Moreover, emphatic stress can amplify these roles supporting language comprehension and production of PWA. Motivated by these findings, the current study added percussive beats to the sentences to highlight words' primary syllabic stress within a scripted sentence context. It is expected that by highlighting these stress features during speech entrainment, PWA will learn scripted sentences via an improvement in lexical access.

3.5 Metronomic beats as a rhythmic support for chunking and memorization

As mentioned before, speech entrainment, the unison production of sentences by patient and clinician, depends on rhythmic features. The literature reviewed on word stress showed how the rhythmic features inherent to the speech signal facilitate lexical retrieval and, in turn, scripted-sentence learning. One additional possibility by which rhythmic features can facilitate scripted-sentence learning is that the unison production of scripted sentences enhances sentence memorization via chunking. In other words, producing speech in unison can result in regular rhythmic patterns that can facilitate memorization of sentences via information grouping/chunking. This section examines how predictable rhythmic elements, such as metronomic beats that are not related to lexical retrieval (like word stress), can facilitate scripted-sentence learning. Thus, the purpose of examining metronomic beats in language processing is to explore how these beats, in connection with speech entrainment, facilitate scripted-sentence learning in aphasia. The current study added percussive beats to the sentences to highlight metronomic patterns within a scripted-sentence context. It is expected that by highlighting these regular rhythmic features during speech entrainment, PWA can memorize scripted sentences via chunking.

3.5.1 Using metronomic beats to memorize sentences via information chunking, evidence from neurotypicals

Before examining metronomic beats in association with memory, the first step is to establish that entrainment to these beats is possible. In linguistics, repeating phrases in time with metronomic beats is known as the Speech Cycling Paradigm (Cummins & Port, 1998). This paradigm represents a case of speech rhythmic organization in which the only possibility is to produce a phrase that fits with a given rhythmic structure (Cummins, 2002). In other words, speech entrainment to metronomic beats is possible given that these beats define a rhythmic structure (regularly spaced rhythmic intervals) over which entrainment can occur. In this paradigm, words or sentences are produced with regular periodic patterns established by the onset of stressed syllables (Port, 2003). Additionally, when using this paradigm, the cognitive load for language production is assumed to be low due to i) multiple sentence repetitions that do not necessarily require lexical access; and ii) the efficiency that rhythm provides for coordinating structures in motor sequences.

Using the Speech Cycling Paradigm, Cummins and Port (1998) examined English speakers repeating short phrases (e.g., **big** for a **duck**) in time with a two-tone metronome. Participants saw the short phrase, then listened to the pair of tones, and then were asked to repeat, multiple times, the phrase in time with the tones (i.e., speech entrainment to a metronome). Their findings showed that speakers could entrain to a metronome, although they were not precise in producing syllables strictly in time with the tones (Cummins & Port, 1998). Participants produced three reliably rhythmical patterns. In the first one, "big" and "duck" were aligned on successive beats, leaving the next beat silent. In the second pattern, "big" and "duck" were produced in alternating beats. In the last pattern, "big," "for a," and "duck" were aligned in three successive beats.

A critical finding of Cummins and Port (1998) is that the cycle duration for phrase repetition affects the timing of the second stressed word in phrases like “big for a duck.” Speakers showed a preference for producing “duck” at a delay that was a whole-integer ratio of the cycle duration. This finding is important because it argues against precise alignment to beats (at least ones that are perceived to be temporally out of synchrony with a larger timing structure), and provides evidence about the existence of an internal timing “attractor.” Overall, Cummins and Port (1998) provided evidence of the use of the Speech Cycling Paradigm in neurotypicals, supporting the use of entrainment to metronomic beats and revealing a tendency to align beats with stressed syllables. Similar results have been reported in other stress-timed (e.g., Arabic and English) and mora-timed (e.g., Japanese) languages (Tajima, Zawaydeh, Kitahara, & Port, 2000).

Given that speech entrainment to metronomic beats is possible, the next step is to identify its utility for memorizing information via chunking. It is important to recall that producing speech in unison can result in regular rhythmic patterns that could facilitate memorization of scripted-sentences via chunking. Extended observations in psychological research have elaborated on the role of chunking for memorization (for a review, see Gobet et al., 2001). For example, it has been observed that chunking/grouping elements with pauses in between positively influence immediate memory recall for verbal items presented in a list (Hitch, Burgess, Towse, & Culpin, 1996). Chunking has also been shown to facilitate short-term memory recall (with items such as digits and letters) in tasks like serial recall, backward recall, and item recognition (Murdock, 1995).

Furthermore, chunking has been used as a mechanism for learning sentence grammar in artificial languages (Servan-Schreiber & Anderson, 1990). A full review of the memory and chunking literature is out of the scope of the current study; instead, it will focus on chunking through regular rhythmic patterns (e.g., metronomic beats) and its potential association with sentence memorization/recall.

Wallace (1994) examined the effects of rhythm and melody on text (lyrics) memorization. One type of rhythm studied in these experiments was metronomic beats. In the first experiment, neurotypical participants (n=64) heard three verses in either a spoken or sung condition. Participants were presented with the three verses, five times, and were asked to recall by writing after the first, second, and fifth repetitions. When analyzing the percentage of words verbatim recalled, Wallace (1994) found a significantly greater recall in the sung condition in comparison with the spoken condition. In relation to chunking, it was reported that the sung condition generated better line accuracy performance, and the errors that participants made preserved the rhythmic organization of the sentence (e.g., "Oh if I had a sailing ship" was recalled as "I wish I had a sailing ship"). In other words, in the sung condition, participants showed more awareness of structural elements such as rhythm and line breaks.

In an additional experiment, a different group of participants (n=21) heard the same three verses spoken rhythmically to a metronomic beat. Similar to experiment 1, participants were asked to recall after the first, second, and fifth presentations. The metronomic beat condition results were compared to the results of the sung condition in experiment 1. It was found that recall was significantly better for the sung condition than the metronomic beat condition. However, the metronomic beat condition was not compared to the spoken condition, which likely would have been significantly different. Additionally, the figures from their study show that the differences in recall were larger for the sung vs. spoken conditions in comparison to the sung vs. metronomic beat condition. Comparing these three conditions (not conducted in the study) would have clarified Wallace's results. Notably, the difference between the number of lines recalled (a measure of chunking) in the sung vs. metronomic beat conditions was not significantly different. The advantage found for both sung and metronomic beat conditions, compared to the control condition, is consistent with the idea of rhythm as a chunking mechanism to learn sentences.

In an additional set of experiments, Purnell-Webb and Speelman (2008) investigated the effects of *familiarity, rhythm, and chunking* on text recall in healthy participants (n=100). It was hypothesized that *familiarity* of melody or rhythm would facilitate text recall as the associative link between the melody/rhythm and the text is strengthened. In terms of *rhythm*, it was hypothesized that the integration between rhyme, metrical prosody, and beats would facilitate text recall. It was also hypothesized that melody and rhythm would provide a natural verse framework that facilitates text recall through *chunking*. Notably, these experiments had a broader use of rhythm, not restricted to metronomic beats, as shown below.

Purnell-Webb and Speelman (2008) compared the effects of using a familiar/unfamiliar melody, a familiar/unfamiliar rhythm, and a spoken input on written text recall. The familiar melody condition was a well-known melody, and the unfamiliar melody condition was a new melodic composition. The familiar rhythm condition consisted of informing participants that the drum sound was tapping to the rhythm of the familiar melody. In contrast, in the unfamiliar rhythm condition, participants heard the same rhythm but were not informed about the rhythm belonging to a well-known melody. Participants were presented with four verses; after each verse, they were asked to recall by writing as much of the verse's text as possible. Subsequently, participants were presented with the four verses (three times) and ask to recall again. Finally, participants were asked to recall with a delay of 15 minutes. When text was presented by verse, text recall was measured via verbatim words recalled after each verse. When the four verses were presented together, text recall was measured via the percent of lines with the correct number of syllables produced.

In terms of *familiarity*, significantly more text was recalled in the familiar conditions in comparison to the unfamiliar and spoken conditions. There was no evidence to support the claim that as the unfamiliar melody became more familiar (e.g., with more trials), it produced an increase in text recall; therefore, increasing melodic familiarity did not confer additional benefit to recall. In terms of *rhythm*, it was found that in both the known and unknown rhythm conditions, participants recalled text significantly better than in the unfamiliar melody or spoken conditions. In addition, participants recalled text similarly in the two rhythmic conditions in comparison with the familiar melody condition. In other words, when providing rhythmic information, recall is facilitated to the same extent as when a familiar melody is used. In terms of *chunking*, the study examined the mean percent of verse lines with the correct number of syllables recalled. There were no significant differences in lines recalled between the familiar melody and the rhythm conditions. However, a significant difference between the rhythm conditions and the unfamiliar condition indicated that participants recalled more lines in the rhythmic conditions than the unfamiliar melody condition.

Overall, Purnell-Webb and Speelman (2008) found that rhythmic and melodic patterns can facilitate text recall. Melodic familiarity generated better recall when compared to unfamiliar melodies and spoken input, but increasing melodic familiarity did not improve recall. Rhythm showed to be a mnemonic and chunking device as effective as a familiar melody. As will be discussed in the following section, this finding is similar to what was reported by Stahl et al. (2011), where PWA learned sentences in a comparable way using metronomic and melodic conditions. Also, performance in the rhythmic conditions was better than in the unfamiliar melody and spoken conditions. Overall, with or without a melodic accompaniment, rhythmic patterns supported sentence recall and chunking in the Purnell-Webb & Speelman (2008) study. Previous studies have reported that musical accompaniment aids sentence-level recall, as measured by the number of syllables and chunks recalled, compared to when the same information is presented in a spoken form (McElhinney & Annett, 1996).

In summary, evidence from the Speech Cycling Paradigm exemplifies that it is possible to entrain to metronomic beats (Port, 2003). This paradigm has been successfully used in English (Cummins & Port, 1998), Arabic, and Japanese (Tajima et al., 2000). Beyond the ability to entrain to metronomic beats, as reported by the Speech Cycling Paradigm, this type of beats has been useful for memorization via chunking. In terms of chunking, it is known that grouping information facilitates learning and memorization of verbal stimuli (e.g., Gobet et al., 2001; Hitch et al., 1996). Wallace (1994) suggested that melodies help in text *encoding*, by chunking lines of text which facilitates learning; and in *retrieving*, by providing a melodic framework over which structural information can be recalled. As seen above, this is also the case for rhythm (Purnell-Webb & Speelman, 2008) and for a specific type of rhythm, metronomic beats (Wallace, 1994). Overall, regular rhythmic features can facilitate sentence memorization via chunking. This regular rhythmic structure exists during speech entrainment aiding scripted-sentence learning. Thus, metronomic beats can provide an enhanced regular structure that facilitates the memorization of scripted sentences via chunking in people with aphasia.

As a reminder, the current study investigates speech entrainment as a key mechanism for scripted-sentence learning. Entrainment depends on the detection, integration, and production of rhythmic features (Phillips-Silver et al., 2010), but it is unclear how these features facilitate learning of scripted sentences in aphasia. One hypothesis is that speech entrainment enhances sentence memorization via chunking (e.g., Purnell-Webb & Speelman, 2008). In other words, producing speech in unison can result in regular rhythmic patterns that can facilitate memorization of sentences via information grouping/chunking. In this study, we applied metronomic beats to sentences, enhancing this regular rhythmic pattern, to facilitate the learning of scripted sentences in PWA. The next section will review evidence regarding the use of rhythm, in general, and metronomic beats, in specific, for sentence learning as it occurs in aphasia.

3.5.2 Benefits of rhythm for language learning in PWA

Before examining the evidence of the use of metronomic beats in PWA, it is crucial to briefly characterize rhythmic processing (perception and production) in aphasia. Zipse, Worek, Guarino, and Shattuck-Hufnagel (2014) studied the discrimination and production of rhythmic patterns in post-stroke aphasia (17 PWA and 21 matched controls). To test *discrimination*, participants judged pairs of rhythms as the same or different. To test *production*, the study included tapping and continuation activities. In the tapping activity, participants listened to a rhythmic pattern of non-isochronous rhythms (uneven beats) repeated six times and were asked to tap along eight times (entrainment task). In the continuation activity, participants were asked to tap the rhythm six more times after the rhythmic pattern had stopped playing.

On average, PWA performed significantly worse than the control group on five out of six measures: i) PWA had more difficulty accurately identifying similar and different rhythmic patterns; ii) they showed more variability when tapping along to a rhythm; iii) they had poorer accuracy; iv) they had greater variability when tapping a rhythm from memory; and, v) they had less abilities in reproducing the relative durations of the timing intervals (Zipse et al., 2014).

Interestingly, the only measure in which most PWA with aphasia did not differ from controls was in the tapping-along task, entrainment. This last finding suggests the potential use of entrainment to rhythmic beats for PWA. The relative preservation of entrainment to rhythm found by Zipse and colleagues may also help explain why speech entrainment benefitted learning of scripted sentences in the studies of Fridriksson and colleagues (2012, 2015).

Adding text to metronomic beats, for example adding phrases or scripted sentences, and entraining to that rhythmic information has shown to improve learning and memorization of sentences in PWA. In a recent study, 12 PWA and 10 controls learned 2-phrase excerpts taken from unknown children's songs (Kershenbaum, Nicholas, Hunsaker, & Zipse, 2017). Participants completed three excerpts in four experimental conditions: rhythmically spoken unison (speech entrainment), rhythmically spoken solo (speech alone, no entrainment), sung unison (melodic entrainment), sung solo (sing alone, no entrainment). The authors measured the percentage of correct syllables in each condition in a group-level analysis. They found better performance in the entrainment conditions than in solo conditions ($F(1,20)=7.16$, $MSE=484.71$, $p=.02$). There was also better performance in the speech entrainment condition than in the singing condition ($F(1,20)=4.54$, $MSE=28.60$, $p=.046$). The interaction between group and unison versus solo was also significant, such that PWA had better results in entrainment compared to the solo condition relative to controls ($F(1,20)=5.52$, $MSE=484.71$, $p=.03$). In short, the participants' ability to learn target excerpts benefited from the use of speech entrainment, singing/speaking along, in comparison to solo production.

A study by Stahl and colleagues examined how PWA (n=17) learned three types of lyrics: *original lyrics,* which corresponded to a German nursery rhyme; *formulaic lyrics*, which contained stereotyped every-day phrases; and *non-formulaic lyrics,* which were syntactically correct but unlikely to occur phrases, such as the text of a poem (Stahl et al., 2011). This study had three conditions: melodic intoning, metronomic beat (called rhythmic speech in the study), and a control condition that consisted of a metronomic beat shifted by an eighth note (called spoken arrhythmic in the study). In the melodic intoning condition patients sang to a playback, and in the metronomic beat conditions (rhythmic speech and control) participants spoke along (entrained) to a playback. An analysis of the percentage of correct syllables produced in response to the different lyric types revealed no effects of the melodic condition when contrasted to the metronomic conditions. On the other hand, the metronomic beat condition showed an effect on the correct number of syllables compared to the spoken arrhythmic condition. Therefore, entrainment to metronomic beats positively affected the production of correct syllables across lyric types. In addition, the degree of benefit for learning was reported to be comparable for metronomic and melodic conditions.

In a follow-up longitudinal study, Stahl, Henseler, Turner, Geyer, and Kotz (2013) treated 15 PWA in either singing therapy, rhythmic therapy, or standard therapy. In the singing therapy condition, participants learned formulaic phrases set to a well-known melody. In the rhythmic therapy condition, participants were trained with the same lyrics, but the lyrics were rhythmically spoken (entraining to a metronomic beat). Standard therapy did not include singing or rhythm but was based on SLP-standard treatment using non-formulaic phrases. Generalization was tested using non-formulaic phrases. Participants in *both* singing and rhythmic therapy showed a significant improvement in the percentage of correct syllables for formulaic phrases; however, rhythmic therapy resulted in a higher percentage of correct syllables than singing therapy (this percentage was not reported as statistically significant). Participants in the standard therapy showed improvement in non-formulaic phrases and were the only group to show generalization to unknown phrases. A caveat of this study is that while the singing therapy and rhythmic therapy groups received training in formulaic sentences, the standard therapy group did not; this change in the stimuli does not allow for a direct comparison between the standard therapy and the singing or rhythmic groups.

In summary, PWA can entrain to rhythmic beats such as speaking in time to a metronome, similar to the Speech Cycling Paradigm. Although some PWA have shown difficulties in rhythm perception and production, some others have shown good performance on entrainment to beats (Zipse et al., 2014). When rhythm is used in rehabilitation, including entrainment, a positive effect has been shown on PWAs' ability to learn sentences (Kershenbaum et al., 2017). Stahl et al. (2011) showed that entraining to metronomic beats positively affected the number of correct syllables PWA produced when learning lyrics. In a follow-up longitudinal study, PWA improved the percentage of correct syllables produced for formulaic phrases under both sung and metronomic beat conditions, but showed greater improvements for metronomic beats (Stahl et al., 2013). Overall, converging evidence presented in this section suggest that metronomic beats are beneficial for sentence memorization in PWA.

3.5.3 A note on metronomic beats as facilitators of phonetic encoding in motor-speech disorders

In addition to sentence learning in PWA (Stahl et al., 2013; Stahl et al., 2011), metronomic beats have also been used in the treatment of motor speech disorders. An early single-subject design successfully used metronomic beats in the improvement of oral motor control tasks, motion rate tasks, isolated word repetition (1 syllable per beat), and sentence repetition (1 syllable per beat) in a participant with severe apraxia of speech and mild anomia (Dworkin, Abkarian, & Johns, 1988). The improvements achieved were sustained in the absence of metronomic beats during the maintenance phase. Another single-subject multiple baseline study reported a participant with mild to moderate apraxia of speech and Broca's aphasia, as evidenced by the WAB, who received a combined treatment using both metronomic beats and hand tapping (Wambaugh & Martinez, 2000). The participant was trained to produce one syllable per beat while tapping his hand to the metronome. This participant showed a significant improvement in the percentage of consonants produced correctly and in word accuracy for both treated and untreated three-syllable words.

In addition, Metrical Pacing Therapy (MPT; Brendel & Ziegler, 2008) uses rhythmic cues to facilitate the organization of speech movements with the aim of improving articulation. A study with a cross-over design compared MPT with a control segment-based treatment. Nine participants with apraxia of speech and mild to moderate aphasia were presented with various rhythmic tones matching the number of syllables and stress patterns of a sentence (metrical templates). The participants were asked to articulate sentences in unison with the template. When a participant failed to entrain to the provided rhythmical cue correctly, a clinician facilitated sentence production by emphasizing word onsets, tapping the pacing rhythm on the patient's hand, or modeling the task fading out gradually. Sentence complexity (sentence length, syllabic complexity, and production rate) were adapted to apraxia severity. Across participants, MPT resulted in an improvement in the acoustic measures of apraxia of speech; for example, there was a significant reduction of sentence duration and a significant decrease of disfluencies relative to the control condition. In other words, participants were more fluent, took less time, and were more accurate in producing a set of sentences (Brendel & Ziegler, 2008).

Overall, metronomic beats have been used to treat motor speech disorders causing significant improvements in acoustic measures for apraxia of speech. For example, studies have reported that participants achieved better control in motor tasks and motion rate (Dworkin et al., 1988); improvements in word-level articulation accuracy (Wambaugh & Martinez, 2000); and an increase in sentence-level fluency (decreased sentence duration and fewer disfluencies) (Brendel & Ziegler, 2008). Although the idea of metronomic beats for memorization has been entertained in the field of motor speech disorders, this type of rhythm has only been examined at a lower motoric and acoustic level and not at a sentence memorization level. The current study will assess motor speech disorders since they commonly co-occur in aphasia (Duffy, Strand, & Josephs, 2014); however, the focus will be on a higher language encoding level of scripted-sentence memorization.

3.5.4 Summary: Evidence and hypotheses regarding speech entrainment as a mechanism for scripted-sentence learning in aphasia.

Major elements can be observed from the evidence regarding speech entrainment as a mechanism through which PWA learn scripted sentences. Entrainment depends on the detection, integration, and production of rhythmic features (Phillips-Silver et al., 2010). In other words, simultaneous production of speech takes place on the basis of rhythm that comes from speech rhythm (e.g., word stress), or from other types of regular rhythm (e.g., metronomic beats). Based on these types of rhythm, two hypotheses emerge about how entrainment assists PWA in learning scripted sentences.

The first hypothesis is that speech entrainment to scripted sentences enables PWA to align to word stress supporting lexical retrieval. Evidence in favor of this hypothesis includes i) metrical and lexical prosody playing an important role in the segmentation of continuous speech, supporting subsequent lexical access in neurotypical English speakers (Cutler 1989; Cutler & Norris, 1988; Cutler, 2005); ii) word stress contributing directly to lexical access in neurotypical Spanish speakers (Soto-Faraco et al., 2001); iii) emphatic stress speeding semantic judgments tasks in PWA (Slansky & McNeil, 1997); iv) emphatic stress aiding sentence comprehension in PWA (Kimelman & McNeil, 1987), in particular when there is an increase in stimulus complexity (Kimelman, 1999); v) PWA showing better repetition and naming performance for words and nonwords with the most common stress patterns in both English (Howard and Smith, 2002) and Dutch (de Bree et al., 2007), suggesting that alignment with regular stress patterns facilitates encoding and recall of word form information. Overall, evidence from neurotypicals and PWA suggests that emphatic stress facilitates lexical retrieval, including in sentence contexts, and that regular stress patterns facilitate word and word form production.

The second hypothesis is that speech entrainment enhances memorization of scripted sentences via chunking. The production of sentences using speech entrainment results in regular rhythmic patterns that can facilitate sentence memorization. Regular rhythmic patterns may be enhanced using metronomic beats, improving memorization via chunking. Evidence in favor of this hypothesis comes from i) the possibility of entrainment when repeating phrases in time with metronomic beats, as seen in the Speech Cycling Paradigm (Cummins & Port, 1998); ii) the successful use of metronomic beats in PWA for sentence learning, particularly for formulaic language (Stahl et al. 2011, 2013); iii) rhythm, in general, as a mnemonic and chunking device (e.g., number of syllables and words, and stress patterns) that facilitates text recall in neurotypicals as effectively as recalling text from singing a familiar melody (Purnell-Webb et al., 2008); and, iv) rhythm enhancing information chunking when neurotypicals recall lyrics (Wallace, 1994). Overall, the evidence suggests that metronomic beats can be used to enhance sentence memorization via chunking in neurotypicals and PWA.

Given these two hypotheses, the current study aims to examine the effects of speech entrainment to two rhythm types, word stress-aligned and metronomic beats, to characterize mechanisms underlying scripted-sentence learning. The stress-aligned manipulation was similar to previous emphatic stress findings in that it highlighted the primary syllabic stress of words; however, it was different in that the stress-related emphases were added, using beats, to all words of the sentences rather than highlighting only one word in comparison to the rest of the sentence. The metronomic beats condition was similar to previous studies using regular, predictable beats in congruency with the sentence. The use of beats in both stress-aligned and metronomic manipulations made these conditions comparable. Additionally, the current study aims to determine whether individual characteristics in behavioral and electrophysiological measures (cortical tracking) are associated with scripted-sentence learning. These aims help characterize mechanisms and individual characteristics in scripted-sentence learning, improving aphasia rehabilitation for Spanish-speaking PWA.

4.0 Aim 1: Identify the effects of speech entrainment to two rhythm types on learning scripted sentences.

As described above, script training improves the production of everyday sentences through intense rehearsal (Cherney, Halper, et al., 2011; Goldberg et al., 2012; Lee et al., 2009). PWA learn scripted-sentences via *speech entrainmen*t (Fridriksson et al., 2015; Fridriksson et al., 2012), the unison production of sentences by patient and clinician. Entrainment relies on the detection, integration, and production of rhythmic features (Phillips-Silver et al., 2010), but it is unclear how these features facilitate learning. Two hypotheses regarding this facilitatory effect are: i) unison production of scripted-sentences enables PWA to align to word stress, supporting lexical retrieval (e.g., Cutler, 2005; Soto-Faraco et al., 2001), and ii) unison production of scripted-sentences using metronomic beats enhance sentence memorization via chunking (e.g., Purnell-Webb & Speelman, 2008). To test these hypotheses, this study examined the effects of entrainment to two rhythm types, stress-aligned vs. metronomic beats, characterizing the mechanisms underlying scripted-sentence learning.

4.1 Methods for Aim 1

4.1.1 Participant recruitment and eligibility criteria

Fourteen participants were recruited in Colombia from a university-sponsored aphasia support group (located in the city of Bogota) and from an outpatient neurology institute (located in the city of Medellin). The 14 participants with aphasia were between 31 and 78 years old and native Spanish speakers. Participants had aphasia secondary to a single left hemisphere cerebrovascular accident. Participants also completed an audiometric screening at 40 dB at .5, 1, 2, and 4 kHz. This study was reviewed and approved by the University of Pittsburgh's Institutional Review Board (IRB), by the ethics committee at the outpatient neurology institute (Neuromedica) in Medellin, and by the university-sponsored aphasia support group (EFUNA) in Bogota.

Before enrolling participants to be assessed with standardized cognitive and language assessments, potential participants were screened using the Language Screening Test (LAST; Flamand-Roze et al., 2011). They needed to score between 4 and 13 on the LAST to be eligible for the study. The LAST was previously used in Colombian PWA to screen for language impairments secondary to stroke (Quique, Dresang, & Dickey, 2018). In addition, potential participants were screened for apraxia of speech using a Spanish adaptation of the Duffy (2005) protocol. Participants were excluded if they presented with severe apraxia of speech. Additional exclusionary criteria included multiple strokes, progressive neurological disorders, or medical conditions that severely affected cognitive function (e.g., dementia) and difficulties tolerating 2+ hour testing sessions.

4.1.2 Short-term learning paradigm

Planned involvement for each participant was eight sessions lasting $\leq$ 2.5 hours each. The first two sessions were dedicated to language, attention, and rhythmic processing assessment (see detailed information about these assessments in section 5.1.1). The third session was dedicated to the cortical tracking measure. This measure was only collected with enrolled participants in Medellín (n=9) because, during the data collection period, the EEG equipment was not accessible in the city of Bogota. The final five sessions were dedicated to the short-term learning paradigm in which participants learned the scripted sentences (see the 4.1.3 stimuli section, below).

The short-term learning paradigm was used to identify the effects of speech entrainment to two rhythm types on learning scripted sentences. In other words, this paradigm allowed the investigation of speech entrainment as a key mechanism through which PWA learn scripted sentences. This paradigm was selected given that at the start of language treatment, there are consistent improvements in communication abilities, which appear as positive slope changes in accuracy for treated stimuli (e.g., Quique, Swiderski, Hula, & Dickey, 2018; Quique, Evans, & Dickey, 2019). A short-term learning paradigm can capture these initial changes in slope that result from initial response to treatment (i.e., treatment effects). In this case, after five sessions of learning scripted sentences, the paradigm was sensitive to initial slope changes.

The short-term learning paradigm differs from a treatment study mainly in the duration of training (five sessions), in the lack of a baseline administration, and in that probes were collected at the end of each session rather than at the beginning of it. Post-session probes allowed us to collect five data points per participant, rather than four if we had collected probes at the beginning of each session which would have limited probe collection to the start of the second training day. This type of probes also captured short-term learning within a training session instead of longer-term changes in learning.

In addition, this learning paradigm allowed a careful control of the amount of practice (dosage) that each participant received. Controlling for dosage is a step forward in answering how changes in the amount of training or the practice's distribution influence scripted-sentence learning. In the current study, we provided sufficient opportunities for practice while controlling the amount of practice for each participant (see Figure 3, below). Overall, the short-term learning paradigm was designed to control for dosage such that participants received the same amount and distribution of practice, five two-hour sessions, distributed over two weeks.

Finally, this learning paradigm is similar to what is known in the pediatric language treatment literature as dynamic assessment (Poehner & Lantolf, 2005) in which a strategy is explored for a brief amount of time to examine its effectiveness and to adjust it to individual needs. Overall, this short-term learning paradigm offers a model to test initial response to a given treatment, in this case, response to scripted-sentence learning.

4.1.3 Stimuli

Participants learned 30 sentences, consisting of 150 words in total, over five 2-hour sessions. The total number of sentences was divided into 10 sets of 3. Each sentence set was matched for number of words, number of syllables, syntactic structure, and roughly matched for phonetic complexity. Additionally, using a lexical frequency Spanish database (Cuetos, Glez-Nosti, Barbón, & Brysbaert, 2011), lexical frequency was compared and matched as closely as possible across the set members (the complete set of sentences can be found in Appendix 1). All participants learned the same 30 target sentences, but randomly assigned to one of the three conditions for each participant. The sentences were generated in consultation with an expert in Spanish linguistics at the University of Pittsburgh. To avoid ceiling effects for participants with milder aphasia, the sentence sets varied in complexity level based on each set's linguistic properties.

4.1.4 Conditions

Each sentence was recorded by a Spanish native speaker using clear speech. For this study's purposes, clear speech is understood as articulating phonemes precisely and accurately, slowing speech rate, increasing pause durations, and increasing intensity (Uchanski, 2005). Using clear speech was important for the current study for multiple reasons: it benefits intelligibility when listening to speech-in-noise, it has shown to improve memory for sentences (Van Engen, Chandrasekaran, & Smiljanic, 2012), and it allowed the addition of beats for the manipulated conditions. The recording made using clear speech served as the control condition for this study.

Each sentence appeared in one of three conditions. *Control Condition*: in this condition, the recorded sentence was presented without the addition of any beats. *Stress-aligned Condition*: in this condition, digitally added beats aligned with word stress. To ensure clarity and following the methods reported by Stahl et al. (2011, 2013), the added beats' intensity was decreased by ~10 dB compared to the recorded sentence. Beats aligned with syllable onsets, one beat per syllable, such that when syllabic onsets were voiced, the beat aligned with the beginning of the syllable, but when the syllabic onset was voiceless, the beat aligned with the onset of the vowel. This resulted in a more natural alignment where each beat corresponded to a syllable. Additionally, stressed- vs. unstressed-syllable beats differed in volume and pitch. Accentuated beats (aligned with stressed syllables) had a higher percussive pitch, and an intensity level decreased by ~8 dB in comparison to the sentence; all remaining beats had a lower percussive pitch and an intensity level decreased by ~10 dB. *Metronomic Beat Condition:* as in the stress-aligned condition, the beats were digitally added at syllable onset. Similar to a metronome, the beats were equally spaced across the sentence. Beats in this condition had a percussive pitch similar to the stressed-syllable beats in the stressed-aligned condition. A schematic example of the three conditions is presented in Figure 2.

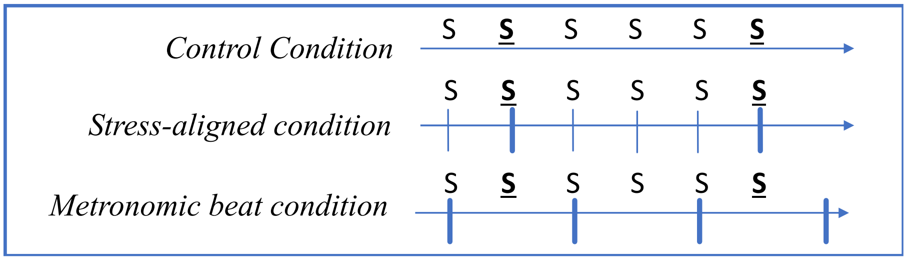

Figure 2 Schematic example of the three conditions. S = syllables, S = Stressed syllables.

Procedure

All experimental procedures – instructions, practice, experimental stimuli, and probe stimuli – were presented using PsychoPy version 3.2.4 (Peirce et al., 2019). Before beginning to learn the targeted scripted sentences, participants were familiarized with the computer program and the entrainment task by practicing with three sentences, one in each condition (see practice sentences in Table 1). During practice, participants were provided with instructions about the sequential steps of listening, entrainment, and independent production. These instructions were aphasia friendly consisting of a related image and the words: “listen” (“escucha”), “let’s say it together” (“digámoslo juntos”), and “you” (“dilo tú”). During practice trials, they also received feedback about their performance. These trials were not scored or included in the analysis.

Table 1 Practice sentences.

Phrase	Number of words	Number of syllables	Syllabic stress
Vaso de leche *[Glass of milk]* Jugo de fresa *[Strawberry smoothie]* Jarra de vino *[Jar of wine]*	3	5	**S** S S **S** S

Similar to the practice trials, in the experimental trials, participants saw instructions on a computer screen signaling when to listen, when to speak in synchrony with the sentences (speech entrainment), and when to attempt independent sentence productions. These steps, part of the short-term learning paradigm, are shown in Figure 3. The number of repetitions within these steps follow errorless approaches reported for nouns and verbs (Abel, Schultz, Radermacher, Willmes, & Huber, 2005; Conroy et al., 2009c). In this case, PWA repeated the sentence in synchrony with the recording a total of 4 times (*speech entraining*) before producing it independently. Participants heard the sentences via headphones and saw an image related to each sentence. This image was later used as a cue for probe administration. Participant productions were audio-recorded and scored offline.

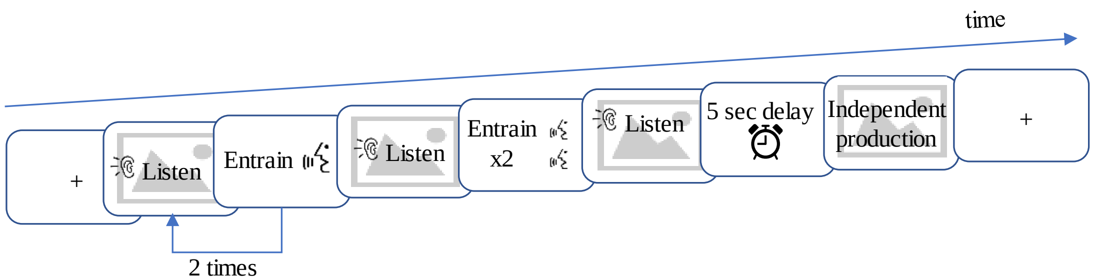

Figure 3 Steps of the sentence-learning paradigm.

Sentences were blocked by condition and randomized across participants. Participants saw the complete set of 30 sentences in each session: ten sentences in each condition with breaks between conditions (2-minute pause) to avoid fatigue. In addition, the complete set was presented twice in each session, with breaks between each time (5-minute pause). This procedure was repeated on each of the five days of the short-term learning paradigm, over two weeks.

Learning was measured via *post-session probes* in which participants produced the learned sentences in response to a cue (the related image unique to each sentence). See Figure 4 for a schematic example of a learning session. Although probes are typically administered before each treatment session, in the current study, post-session probes were selected to obtain a total of five measurements, given that the short-term learning paradigm included five sessions in total. Of note, the administration of post-session probes reflected the initial response to the proposed short-term learning paradigm within a session instead of longer-term changes in learning.

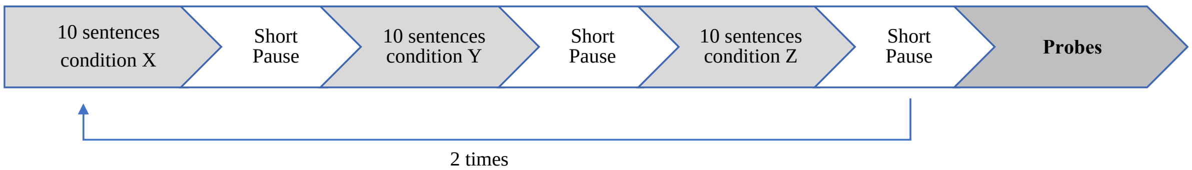

Figure 4 Schematic example of a single session

4.1.6 Dependent variable and scoring procedures

In the current study, learning was analyzed at a group-level using probe accuracy at the syllable-level. Syllable-level accuracy was chosen as a dependent variable for multiple reasons. First, both the metronomic and stress-aligned conditions emphasize syllables and consequently, the sentence's metrical structure. Second, previous studies that have used rhythmic interventions in PWA (Stahl et al., 2011, 2013) have used accuracy at the syllable level as a dependent variable. Third, studies in neurotypicals that have examined the effects of rhythm on sentence recall have used syllable-level information (Purnell-Webb & Speelman, 2008). Lastly, for a learning paradigm that focuses on syllabic and segmental information, syllable-level accuracy can provide a fine-grained measure of language production in PWA.

Regarding scoring procedures, each of the five probes gathered at the end of each session was scored offline. To score accuracy, we adapted the Philadelphia Naming Test (PNT) scoring rules (Roach, Schwartz, Martin, Grewal, & Brecher, 1996). From the sentence produced in response to the image (i.e., probes), we identified the words that were *complete attempts* (i.e., not self-interrupted, with clear downward or upward/questioning intonation, and a noticeable pause between words). After identifying complete attempts, we counted the number of syllables that belonged to the target sentence (*number of target syllables*). We allowed for the addition and deletion of plurals, and participants with a motor speech disorder (e.g., apraxia) were also allowed one sound omission, addition, or substitution per word. Syllable additions due to Spanish re-syllabification were also counted as correct target syllables. For example, the sentence el gato-esta dormido [the cat is sleeping], in which "to-es" become a single syllable, was counted as having the correct number of syllables. Similarly, the sentence el gato esta dormido, in which "to es" are produced as two different syllables, was also counted as having the correct number of syllables. To track progress per session, we also scored the independent production (see Figure 3) of each sentence during the short-term learning paradigm following the same scoring procedures.

Data processing was performed mainly with the tidyverse package, version 1.3.0 (Wickham et al., 2019) and the dplyr package, version 0.8.5 (Wickham, Romain, Lionel, & Müller, 2020). Figures were made mainly using the ggplot2 package, version 3.3.0 (Wickham, 2010) and the sjPlot package, version 2.8.3 (Lüdecke, 2020).

4.1.7 Analysis

Logistic binomial models, using a binomial distribution, can be applied to count data where each data point represents the number of successes in a given number of trials (Gelman & Hill, 2007). In the current study, this translates into the proportion of syllables that PWA got correct out of the total number of syllables of a given target sentence. Therefore, the proportion of syllables correct out of the total number of syllables per sentence (probe accuracy at the syllable-level) was modeled as count data using generalized mixed-effects logistic regression models (Breslow &

Clayton, 1993) with random intercepts and slopes for subjects and items (Baayen, Davidson, & Bates, 2008).

It is suggested that a maximal model structure should include random intercepts by-items and by-individuals, as well as random slopes by-items and by-individuals for each of the fixed effects (Barr, Levy, Scheepers, & Tily, 2013). Therefore, the maximal model planned a priori included: i) fixed effects of condition and time (session); and ii) a random effects structure with random intercepts and slopes, for the fixed effects of interest, nested within-participants and within-items[1]. However, data from treatment studies with small sample sizes do not always allow for a maximal model (Wiley & Rapp, 2019).

In this case, the planned maximal structure had convergence issues for which the maximal model, as permitted by the data, was selected. The selected model included fixed effects of condition and time (sessions) and random effects for session within-participant, for condition within-items (sentences), and for observations[2]. When using a binomial distribution, with count data, overdispersion is a common issue where the data has more variation than is explained by the model (Gelman & Hill, 2007). Overdispersion causes underestimation of the error variance; correcting for it widens the confidence intervals and make p-values appropriately larger. In this study, to correct for overdispersion, a trial-level random effect (1 | observations) was included. By including this term in the model, all the estimates are appropriately adjusted for the overdispersion. The overall random effects structure was determined based on model convergence and model fit.

---

[1] *Planned maximal model*: glmer(cbind(successes,failures) ~ Condition * Session + (1 + Condition * Session | Subject) + (1 + Condition * Session | Sentence).

[2] *Maximal model used*: glmer(cbind(successes,failures) ~ Condition * Session + (1 + Session | Subject) + (1 + Condition | Sentence) + (1 | observations)

As planned, condition was coded using two orthogonal contrasts: one that compared the two rhythmic conditions against the control condition, and one that compared the two rhythmic conditions with each other. Session was centered around its mean. Models were fitted using the lme4 package, version 1.1-21 (Bates, Maechler, Bolker, & Walker, 2015) and the lmerTest package, version 3.1-1 (Kuznetsova, Brockhoff, & Christensen, 2017). The analyses were conducted in the R statistical software version 3.6.3 (R Core Team, 2020) using RStudio version 1.1.456 (RStudio Team, 2016).

4.1.8 Hypotheses and predicted results

Hypothesis 1a: it was hypothesized that entrainment to *both* rhythmic-enhanced conditions would increase the number of syllables produced correctly relative to the control condition. This would indicate that rhythm, regardless of type, had beneficial effects for scripted-sentence learning in PWA (Fridriksson et al., 2015; Stahl et al., 2011). In addition, this finding would provide evidence indicating that speech entrainment mechanisms are key for the learning of scripted sentences in aphasia.

*Hypothesis 1b***:** if speech entrainment to the stress-aligned condition resulted in better performance than the metronomic beat condition, it would suggest that word-stress-aligned production improved lexical retrieval (Cutler, 2005; Soto-Faraco et al., 2001). Specifically, higher syllable accuracy in the stress-aligned condition, in comparison with the metronomic beat condition, would indicate that lexical retrieval access improved as a result of the segmentation and lexical retrieval provided by word stress.

*Hypothesis 1c***:** if speech entrainment to the metronomic beat condition resulted in better performance than the stress-aligned condition, it would suggest that sentence learning is facilitated by sentence memorization via chunking (Stahl et al., 2013; Stahl et al., 2011). Specifically, higher syllable accuracy in the metronomic beat condition, in comparison with the stress aligned condition, would indicate that PWA are memorizing sentence chunks using regular rhythmic structures (e.g., metronomic beats).

4.2 Results for Aim 1

Before reporting group-level results, it is important to note that participant COL13 was not included in the short-term learning paradigm. After working on two different sessions, he did not consistently follow the instructions regarding when to listen, when to entrain, and when to have an independent production. In these two sessions, COL13 received continuous feedback to follow the computer instructions, but his responses were highly variable. In agreement with his family, it was decided not to continue his participation in the study. The rest of the participants with aphasia (n = 13) completed the five sessions of the short-term learning paradigm, and their results are reported below. See Table 5 for demographic information.

The results are divided into three main segments. The first segment reports performance during the short-term learning paradigm. The second segment reports on probe accuracy. The third segment reports model results for the group-level analysis.

4.2.1 Performance during the short-term learning paradigm

Performance during the short-term learning paradigm refers to the proportion of correct syllables produced during the independent production step of training (see Figure 3). As depicted in Figure 5, there was high variability during this independent production step. For example, some participants displayed a positive slope suggestive of improvement over time during training (e.g., COL1 and COL9); other participants performed almost at ceiling during training (e.g., COL5 and COL14); and some others showed a shallower slope suggesting similar performance across sessions without ceiling effects (e.g., COL4 and COL12). See Table 2 for the mean proportion of syllables produced correctly during training collapsing across participants and sessions.

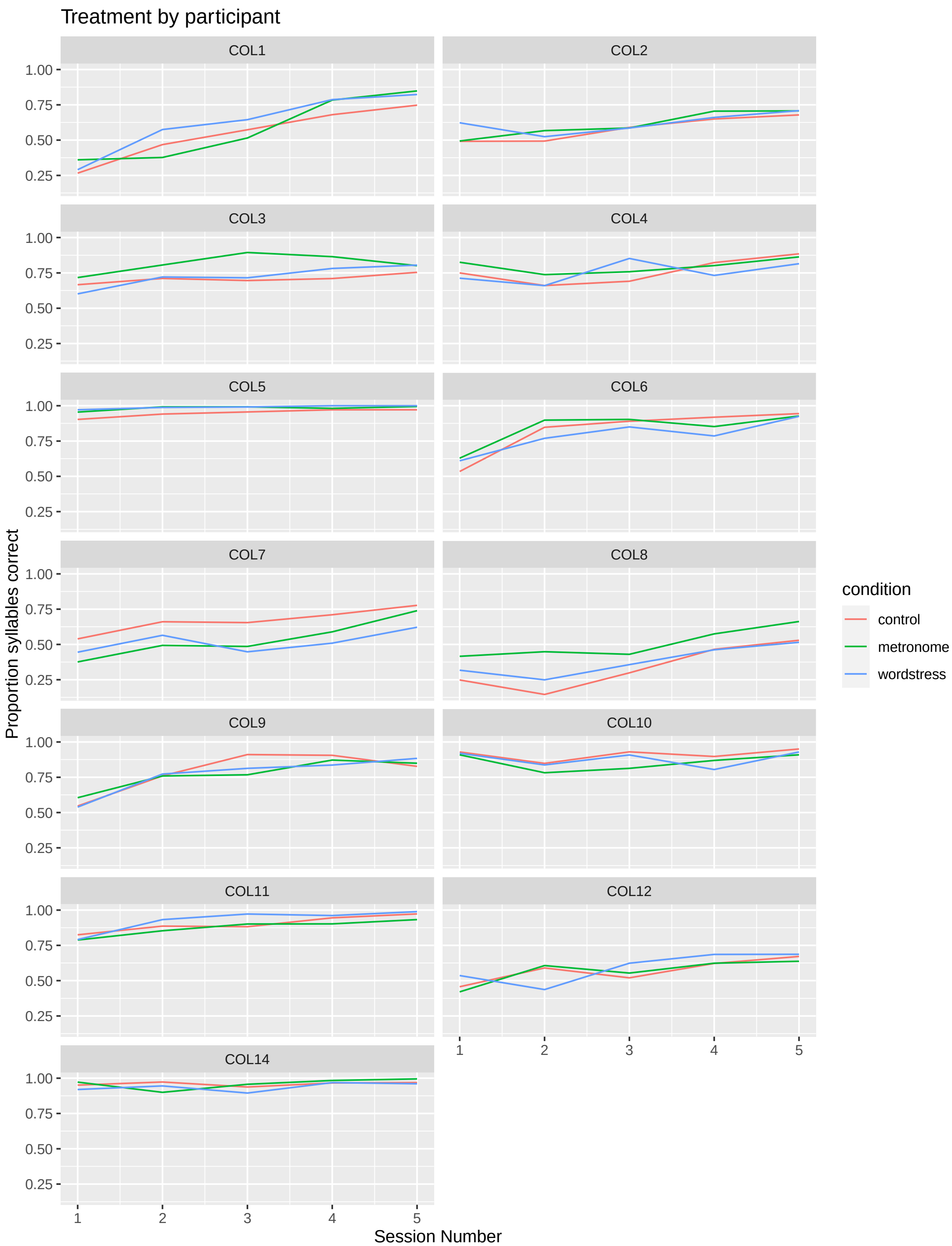

Figure 5. Participant performance during the short-term learning paradigm

Table 2. Syllable-level performance during the short-term learning paradigm across participants and sessions

Condition	Mean proportion syllables correct	SD proportion syllables correct
control	0.73	0.30
metronome	0.75	0.28
stress-aligned	0.74	0.29

Notes: SD = standard deviation

4.2.2 Probe accuracy

Probe accuracy at the syllable level, *the dependent variable*, refers to the proportion of correct syllables produced during probe collection. As a reminder, probe accuracy at the syllable-level was chosen as a dependent variable following studies that have used rhythmic interventions in PWA (Stahl et al., 2011, 2013) and studies that have used rhythm to enhance sentence memorization in neurotypicals (Purnell-Webb and Speelman, 2008). Also, as noted above, syllable-level accuracy provides a fine-grained measure of language improvement for PWA that do not produce an entirely correct scripted sentence (i.e., word by word), but do have a production that is close enough to match the intended meaning of the sentence. As depicted in Figure 6, there was high variability in probe performance. Most participants showed a positive slope suggesting improvement over time (e.g., COL1, COL3, COL12), while others showed a shallower slope (e.g., COL8 and COL11) with little or no improvement. Learning over time will be examined in the group-level analyses below, but see Table 3 for the mean proportion of syllables produced correctly during probes collapsing across participants and sessions.

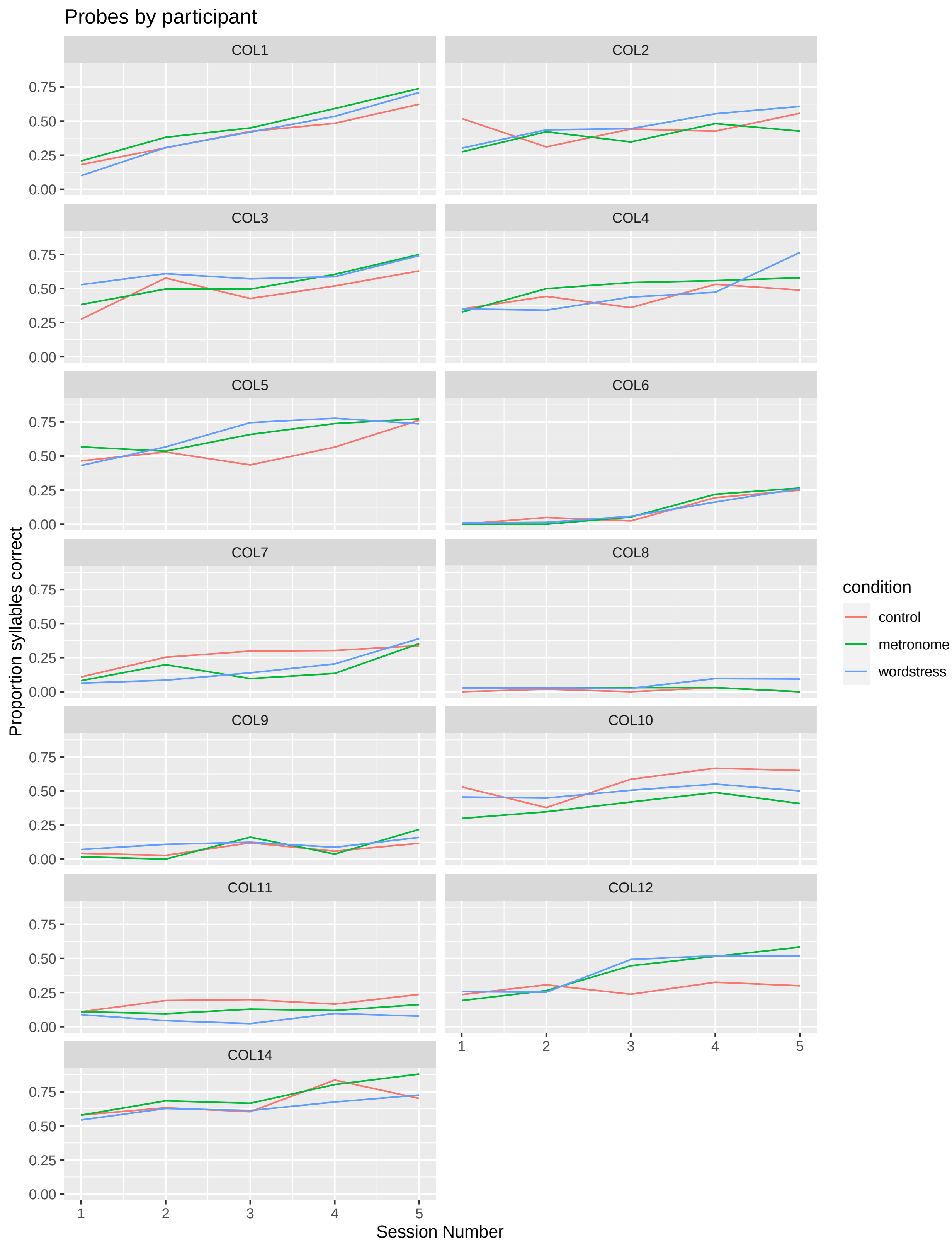

Figure 6. Participant probes performance

Table 3. Probe accuracy at the syllable level across participants and sessions

Condition	Mean proportion syllables correct	SD proportion syllables correct
control	0.35	0.31
metronome	0.35	0.33
stress-aligned	0.36	0.32

Notes: SD = standard deviation.

4.2.3 Group-level results

Mixed-effects logistic regression models (Baayen et al., 2008; Breslow & Clayton, 1993) were used to examine probe accuracy at the syllable level. Fixed effects of condition and session, and random effects for subjects and items were included[3] (Baayen et al., 2008). Condition was coded using two orthogonal contrasts: one that compared the two rhythmic conditions against the control condition (contrast 1) and one that compared the two rhythmic-enhanced conditions with each other (contrast 2).

Collapsing across conditions, the odds of producing a correct syllable were 1.5 times greater with each additional session (ß = 0.41, SE = 0.06, z = 6.18, $p < 0.001$). In other words, there were significant improvements over time across all conditions demonstrating successful sentence learning and providing further evidence of script training as an efficacious treatment for aphasia. Importantly, this is the first study to examine scripted-sentence learning in Spanish speakers with aphasia, demonstrating cross-linguistic benefits of script-training interventions.

[3] glmer(cbind(successes,failures) ~ Condition * Session + (1 + Session | Subject) + (1 + Condition | Sentence) + (1 | observations)

In relation to the differences between the rhythmic-enhanced conditions and the control conditions, no significant differences were found in the odds of producing a correct syllable when comparing the control condition with the rhythmic-enhanced conditions (contrast 1), collapsing across sessions ($ß = 0.03$, $SE = 0.08$, $z = 0.35$, $p = 0.72$). In other words, PWA produced a similar number of correct syllables in the rhythmic-enhanced and control conditions averaging across sessions.

In relation to the differences between the rhythmic-enhanced conditions, no significant differences were found in the odds of producing a correct syllable when comparing the metronomic condition with the stress-aligned condition (contrast 2), collapsing across sessions ($ß = 0.04$, $SE = 0.15$, $z = 0.28$, $p = 0.78$). In other words, PWA produced a similar number of correct syllables in each of the rhythmic-enhanced conditions averaging across sessions.

Although these main effects are of interest, the crucial results are based on the interaction effects between condition and session. These interaction effects provide evidence regarding *differences between conditions in learning effects over time*. There was a significant interaction effect between contrast 1 (rhythmic-enhanced conditions vs control condition) and session indicating the difference between the rhythm-enhanced and control conditions increased, as session increased ($ß = 0.12$, $SE = 0.05$, $z = 2.47$, $p = 0.014$). In other words, PWA produced a significantly higher number of syllables in the rhythmic-enhanced conditions when compared to the control condition as the number of sessions increased. This finding is consistent with the claim that rhythm inherent in speech entrainment (either from word stress or from regular rhythmic structures) is a facilitative mechanism for scripted-sentence learning. See Figure 7 for a depiction of this interaction effect.

Additionally, there was no interaction effect between contrast 2 (metronomic vs. stress-aligned) and session indicating that the difference between the metronomic and stress-aligned conditions did not vary as session increased (ß = -0.01 , SE = 0.05, z = -0.18, p = 0.86). In other words, the number of correct syllables produced did not differ in the two rhythmic-enhanced conditions as the number of sessions increased. This finding indicates that both types of rhythms, stress-aligned (aiding lexical retrieval) and metronomic beats (improving memorization via chunking) benefited scripted-sentence learning.

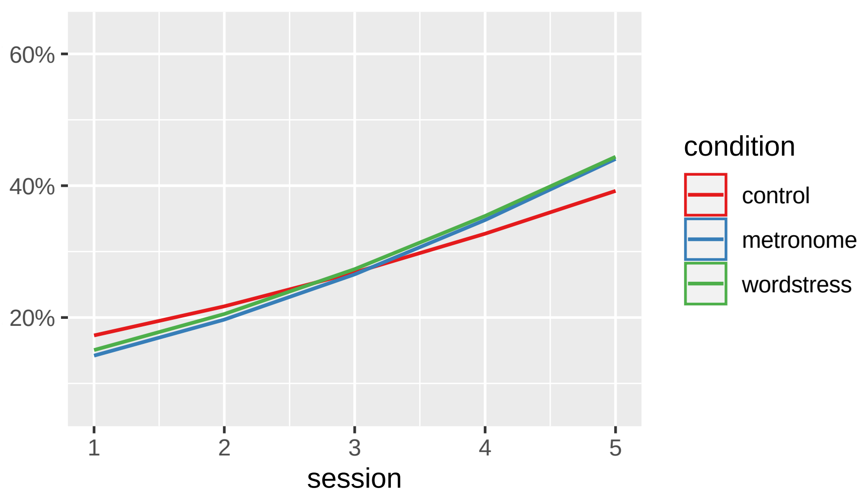

Figure 7. Effects of increasing the number of sessions by condition on the proportion of correct syllables produced.

***Note*: The y-axis representes persentage of syllables correct. The x-axes represents the number of sessions.**

It is important to mention that the primary statistical analysis was planned with a set of contrasts comparing the two rhythmic conditions against the control condition, and to each other. In other words, the second contrast coding was designed to examine the differential contributions of speech entrainment to rhythmically enhanced word stress (supporting lexical retrieval; Soto-

Faraco et al., 2001) and to metronomic beats (supporting memorization via chunking; Purnell-Webb & Speelman, 2008; Stahl et al., 2011). To further test each of the two rhythmic-enhanced conditions, against the control condition, we ran a post-hoc analysis using treatment coding with the control condition as a reference value (i.e., dummy coding). As expected, it was found that as session increased, the difference between the metronomic condition and control condition increased (ß = 0.12 , SE = 0.05, z = 2.21, p = 0.02); similarly, as session increased, the difference between the stress-aligned condition and control condition also increased (ß = 0.11 , SE = 0.05, z = 2.04, p = 0.04). These beta coefficients can be interpreted as a measure of the effect of one additional session on probe performance in each of the rhythmic conditions when compared to the control condition.

4.3 Discussion for Aim 1:

At the group level, the significant improvements seen over time across conditions provide critical evidence for three conclusions regarding script training. First, it indicates successful sentence learning and provides further evidence for script training as an efficacious treatment for Spanish speakers with aphasia. This finding is consistent with the breadth of evidence about the efficacy of script training in English speakers with aphasia (Cherney et al., 2008; Cherney, Halper, et al., 2011; Cherney et al., 2015; Cherney et al., 2014; Kaye & Cherney, 2016; Lee et al., 2009; Manheim et al., 2009). Second, it strengthens the evidence for the use of speech entrainment for scripted-sentence learning (Fridriksson et al., 2015; Fridriksson et al., 2012). As speech entrainment relies on detecting and integrating rhythmic features, highlighting these features (e.g., word stress or regular rhythmic structures) facilitated scripted-sentence learning. Last, evidence to date regarding script training's benefits and mechanisms has come almost exclusively from monolingual English speakers. A recent study, however, reported a successful application of script training with a bilingual Spanish-English speaker with post-stroke aphasia (Grasso et al., 2019). The current study is the first study to examine scripted-sentence learning in Spanish speakers with aphasia, demonstrating the cross-linguistic benefits of script training interventions.

4.3.1 Training and probe performance

As shown in Figure 5, Figure 6, Table 2, and Table 3, there was high variability in performance during both the independent production step of the short-term learning paradigm and during probes. This variability in response to language interventions is commonly seen in multiple aphasia interventions, including in script training (e.g., Cherney, Halper, et al., 2011; Goldberg et al., 2012; Lee et al., 2009). Importantly, performance was considerably higher during the training phase of the short-term learning paradigm (where the mean proportion correct was ~0.74) in comparison to performance during probes (where the means proportion correct was ~0.35). In other words, as a group, participants performed twice as well in the independent production during training (immediate recall) than in probes (delayed recall). Although there were no statistical analyses to directly examine the relationship between training and probe performance, it may be the case that speech entrainment in the context of training is necessary but not sufficient for durable learning of scripted sentences as assessed during probes. Future research can examine the relationship between probe and training performance with careful and theoretically-motivated models. Recent work has begun to examine the relationship between performance during training and longer-term aphasia treatment outcomes in aphasia (e.g., Simic et al., 2020).

4.3.2 Differences between the rhythmic conditions

It was hypothesized that speech entrainment to both rhythmic-enhanced conditions would increase the number of syllables produced correctly relative to the control condition. Evidence in favor of this hypothesis came from an interaction effect between session and the rhythmic-enhanced conditions. This interaction effect indicated that rhythm, regardless of the type, facilitated scripted-sentence learning when compared to the control condition. Of note, this finding does not provide definitive evidence that speech entrainment is the *only* mechanism underlying the benefits of script training, but it indicates that the rhythmic features critical to speech entrainment are a key active ingredient for scripted-sentence learning (Stahl et al., 2013; Stahl et al., 2011), because enhancing rhythmic features was found to have additional benefits for learning this type of sentences.

It was also hypothesized that the rhythmic-enhanced conditions could be beneficial in one of two ways. First, word-stress-aligned language production may enhance lexical retrieval facilitating learning of scripted sentences (Cutler, 2005; Soto-Faraco et al., 2001); second, metronomic beats which support information memorization via chunking may facilitate sentence learning (Stahl et al., 2013; Stahl et al., 2011). Higher performance in either the stress-aligned condition or the metronomic condition, when compared to each other, would have provided evidence about how rhythm supported scripted-sentence learning. However, neither of these hypotheses was ruled out because, the conditions (stressed-aligned vs. metronomic) did not differ from each other at the group level. This result suggests that both rhythmic-enhanced conditions similarly benefited scripted-sentence learning for this group of participants as a whole.

The lack of a difference between rhythmic conditions opens multiple possible interpretations. First, it could be that both learning mechanisms (lexical retrieval and memorization via chunking) are at play and are similarly beneficial for scripted-sentence learning. Second, it could be that neither specific learning mechanism is behind the effects and that *any* rhythmic enhancement applied to speech entrainment could help in scripted-sentence learning. Third, it could be that both learning mechanisms are at play, but some people benefit from one, and other people benefit from the other. The following paragraphs will expand on these three possible interpretations.

Regarding the first interpretation, the current study provides initial group-level evidence of both of these learning mechanisms being favorable for scripted-sentence learning. Specifically, PWA learned sentences in a similar way in the two rhythmic-enhanced conditions compared to the control condition. The similar effect in sentence learning in these conditions suggests that through speech entrainment PWA aligned to word stress, which supported lexical retrieval (e.g., Soto-Faraco et al., 2001) and that speech entrainment to metronomic beats helped PWA to memorize sentences via chunking (e.g., Purnell-Webb & Speelman, 2008; Stahl et al., 2011). Both of these rhythmic-enhanced conditions had facilitative effects of a similar magnitude. This first interpretation of these findings encourages clinical implementation of rhythmic cues (stress-aligned or metronomic) to facilitate learning of formulaic language. Practicing clinicians can easily take advantage of using rhythm to enhance learning in aphasia as previously considered in treatment approaches such as melodic intonation therapy (e.g., Zumbansen, Peretz, & Hébert, 2014a; Zumbansen, Peretz, & Hébert, 2014b).

The second interpretation, that *any* rhythmic enhancement (i.e., not necessarily tied to improving lexical retrieval or memorization via chunking) could help in scripted-sentence learning is motivated by the fact that speech entrainment depends on rhythmic features (Phillips-Silver et al., 2010). Thus, it could be the case that enhancing the rhythmic features of speech entrainment in different ways could be beneficial for scripted-sentence learning. A possibility to further explore this interpretation is to examine the effects of other rhythmic cues (e.g., syncopation, metronomes with different accentuations) on scripted-sentence learning. These rhythmic cues would have to be less directly mappable to lexical retrieval or memorization via chunking, but still preserving a regular and expected rhythmic structure for speech entrainment. If any of these rhythmic enhancements improve learning, this would provide evidence that the underlying mechanism is not lexical retrieval or memorization via chunking but rather the inherent rhythmic features of speech entrainment.

Regarding the third interpretation, that some people benefit more from one mechanism than the other, the current study will provide preliminary evidence in the second aim of this study (identifying individual characteristics associated with scripted-sentence learning response). Notably, some participants benefited slightly more from the metronome condition during training (e.g., COL8) while others from the stress-aligned condition (e.g., COL1; see Figure 5). In contrast, during probes, some participants benefited slightly more from the metronome conditions (e.g., COL1) while others from the stress-aligned condition (e.g., COL2; see Figure 6). However, it is important to note that none of the participants showed an unequivocal response to one of the rhythmic-enhanced conditions in comparison to the other.

In summary, the evidence emerging from Aim 1 analyses provides group-level evidence of the value of rhythmic features and the importance of speech entrainment as a key active mechanism for scripted-sentence learning in aphasia. In addition to group-level evidence, it is critical to examine individual characteristics associated with scripted sentence learning using these rhythmic features. The reason for this examination is to identify the best candidates for a specific intervention; in this case, scripted-sentence learning that takes advantage of the speech entrainment and the rhythmic features that support it. Identifying individual characteristics associated with greater benefit from one set of rhythmic cues or another can improve treatment personalization; that is, matching characteristics (e.g., language impairment measures) with a specific treatment approach. The identification of individual characteristics associated with scripted-sentence learning will be the focus of Aim 2 in the following chapters.

5.0 Aim 2: Identify individual characteristics in behavioral measures, and cortical tracking, associated with scripted-sentence learning response

5.1 Behavioral measures associated with scripted-sentence learning response

People with aphasia vary in their *behavioral* (Halai et al., 2017) and *neurological* profiles (Butler, Lambon Ralph, & Woollams, 2014). Due to profile heterogeneity, there is also variability in how PWA respond to any treatment, including script training (Cherney, Halper, et al., 2011; Goldberg et al., 2012; Lee et al., 2009). Even when PWA are equated in variables such as language impairment and months post-onset, there is still significant variation in response to treatment (e.g., Conroy, Sage, & Lambon Ralph, 2009a; Conroy, Sage, & Lambon Ralph, 2009b). In other words, in aphasia rehabilitation "one size does not fit all"; therefore, an understanding of individual characteristics, and their association with treatment active ingredients and mechanisms, will improve aphasia rehabilitation and allow for treatment customization. The current study examines individual characteristics, at the behavioral and neuro-electrophysiological levels, that may be associated with scripted-sentence learning response.

In terms of *behavioral* predictors, one of the most recurrent and strong factors associated with language recovery is language impairment, as measured by overall aphasia severity (e.g., Liang et al., 2001; Quique et al., 2019). Consistent with this association, naming impairment has also been associated with gains in therapy (e.g., Fillingham et al., 2005a, 2005b; Quique et al., 2019). For example, a relatively large naming study (n=33, *word-level* intervention) found correlations between pre-treatment language measures and treatment outcomes (naming $r=0.68$, $p<.05$, semantic memory $r=0.61$, $p<.05$, word reading $r=0.71$, $p<.05$; Lambon Ralph, Snell, Fillingham, Conroy, & Sage, 2010). Additionally, divided attention $r=0.48$, $p<.05$, and visuospatial memory $r=0.41$, $p<.1$ were correlated with treatment outcomes. Comparable studies have highlighted the importance of language impairments and cognitive performance, particularly attention, in treatment outcomes (e.g., Gilmore, Meier, Johnson, & Kiran, 2019).

To measure treatment outcomes on five clinically-relevant dimensions (language, neurological, neuropsychological, psychosocial, and socio-economic), van de Sandt-Koenderman et al. (2008) developed the Multi-Axial Aphasia System (MAAS). In this study, fifty-eight participants received semantic and phonological treatment and received the MAAS pre- and post-intervention. Multiple regression analysis revealed that the overall MAAS rating predicted treatment outcomes measured by the Amsterdam Nijmegen Everyday Language Test (ANELT, a measure of language performance; Blomert, Kean, Koster, & Schokker, 1994). In addition, neuropsychological factors (attention, concentration, memory, and executive functioning) contributed independently to treatment outcomes. In the cognitive domain, executive function and visual short-term memory have also been shown to predict the magnitude of anomia treatment response (Gilmore et al., 2019). Overall, language performance and cognitive factors, such as attention, seem to be strongly associated with aphasia treatment outcomes at the word-level.

At the *sentence-level*, the relationship among language, cognitive measures, and treatment outcomes in aphasia is less clear. Some of the main sentence-level aphasia treatments are Verb Network Strengthening Treatment (VNeST; Edmonds, Mammino, & Ojeda, 2014; Edmonds, Nadeau, & Kiran, 2009), Treatment of Underlying Forms (Thompson & Shapiro, 2005), Verb Argument Structure Treatment (VAST; Thompson, Riley, den Ouden, Meltzer-Asscher, & Lukic, 2013), and Mapping Therapy (Rochon, Laird, Bose, & Scofield, 2005; Schwartz, Saffran, Fink, Myers, & Martin, 1994). One common topic in the reported results of these treatment methods is the *heterogeneity in treatment response*. PWA have differed in their responses to treatment, both in their improvement for treated sentences and generalization to untreated sentences. Frequently, more severe language impairments lead to poorer treatment outcomes (e.g., Schwartz et al., 1994), and higher pretreatment sentence construction abilities predict better treatment outcomes (e.g., Edmonds, Obermeyer, & Kernan, 2015). Other person-level variables such as time post-onset of aphasia and degree of sentence-comprehension impairment have not been shown to be predictive of sentence-level treatment response (Quique, Swiderski, et al., 2018).

The *neurological* predictors of aphasia treatment response have been examined relatively widely; it is known that factors such as lesion characteristics, perfusion, and grey and white matter density influence language recovery (Kiran & Thompson, 2019). Larger scale initiatives such as PLORAS (Predicting Language Outcome and Recovery After Stroke) are also being developed to observe group and individual treatment outcomes at the neurological level (Price, Seghier, & Leff, 2010; Seghier et al., 2016). The discussion of these neurological factors and supporting evidence is out of the scope of this study. Please see Crosson et al. (2019) and Kiran and Thompson (2019) for recent and complete reviews. As mentioned above, the current study examines individual characteristics at the behavioral and neuro-electrophysiological levels associated with scripted-sentence learning response.

5.1.1 Methods: behavioral measures associated with scripted-sentence learning response

Participants received language, attention, rhythmic assessments, and general demographic questions (e.g., age and education). To characterize language performance, PWA were given the *Western Aphasia Battery* (WAB; Kertesz, 1982) in its Spanish adaptation (Kertesz & Pascual-Leone, 2000). The WAB assesses spontaneous speech, auditory comprehension, repetition, and naming. From the WAB, we obtained the Aphasia Quotient (WAB-AQ), a measure of severity in which a higher score indicates a milder aphasia.

To characterize sustained and selective auditory attention, PWA were given the Elevator Counting and the Elevator Counting with Distraction subtests of the *Test of Everyday Attention* (TEA; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994). In the *Elevator Counting subtest,* participants pretend they are in an elevator where the floor indicator is broken; they are asked to count recorded "elevator tones" to find the number of their current floor. This subtest has a total of 7 items with a maximum score of 7. The score indicating probable impairment is < 7, and the score indicative of a deficit is ≤ 5. *The Elevator Counting with Distraction* subtest follows the same idea as the Elevator Counting subtest, but participants are required to focus only on low-pitch tones while ignoring high pitch tones. This subtest has a total of 10 items with the age-scaled scores ranging between 0-10.

The TEA has been used in people with traumatic brain injury (Bate, Mathias, & Crawford, 2001; Chan, 2000), and has shown good test–retest reliability in patients with stroke (Chen, Koh, Hsieh, & Hsueh, 2013). This test has also been recommended for assessing attention in PWA (Murray, 2002). Assessing attention was relevant for the current study for two main reasons: first, previous findings have reported attention as a predictor of treatment outcomes in aphasia (e.g., Blomert et al., 1994; Lambon Ralph et al., 2010); and second, it is likely that the beats (tones added to the sentence) used in the current study can help PWA in focusing attention on stressed syllables (Attentional Bounce Hypothesis; Pitt & Samuel, 1990) and metronomic beats (Dynamic Attending Hypothesis; Large & Jones, 1999).

To characterize rhythmic processing, PWA were given the *rhythm subtest* of the *Montreal Battery of Evaluation of Amusia* (MBEA; Peretz, Champod, & Hyde, 2003). In this subtest, participants listen to pairs of short musical excerpts and decide whether they are the same or different. The test is designed such that the differences between the two excerpts are achieved by changing the duration values of two adjacent tones (i.e., changing the rhythmic grouping while retaining the meter and total number of sounds between the two excerpts). This test's highest possible score is 30 points, where higher scores denote better performance; the cut-off score established in neurotypicals is 23 points (representing two standard deviations below the mean; Peretz et al., 2003). The MBEA has shown to be sensitive to musical disorders in perception and memory disorders and has also demonstrated good reliability (Peretz et al., 2003). Additionally, the MBEA has been used successfully in stroke survivors with aphasia (Sihvonen et al., 2016).

To examine beat perception, PWA were given the synchronization with the beat of music task from the *Beat Alignment Test* (BAT; Iversen & Patel, 2008), and the perceptual judgment of the beat in its adapted computerized version (CA-BAT; Harrison & Müllensiefen, 2018b). In the *synchronization with the beat of music* task, participants were asked to listen and then tap to the beat of 12 musical excerpts with a variety of musical styles. In this task, participants listened to each of the excerpts twice via headphones and used a drumstick in their left hand to tap near a recording microphone.

Taps were recorded using Audacity® software, version 2.3.3 (Audacity Team, 2019). Offline scoring for the synchronization task consisted of the following steps: i) identifying the second presentation of each musical excerpt; ii) using the sound finder audacity tool to label each tap of the second presentation; iii) manually checking audacity automatized labels; iv) exporting labels. A custom script was made to omit both extra-long intervals, defined as >1.4 times the raw mean, and unusually short intervals, defined as < 0.6 times the raw mean. This script was also designed to analyze inter-tap intervals and correlate them to the closest excerpt tactus level, as established by Iversen and Patel (2008). Tactus levels, or tactus inter-onset intervals, are the most likely to occur tapping tempos to which participants synchronize (main, half-time, double-time).

The outcomes of the synchronization with the beat of music task involves three main measures suggested by the test's authors (Iversen & Patel, 2008). These measures were computed separately for each participant. The first two measures are the average tapping interval (inter-tap mean) and the variability of tapping (inter-tap standard deviation) for each of the 12 excerpts. The last measure is the match between tapping rate and the tempo of the musical excerpt, adjusting for tactus (correlation between inter-tap mean and tactus level). This last measure is taken as an ability estimate for synchronizing with the beat of music. Participants with a good ability to synchronize with the beat of music would have a higher correlation coefficient value. As this coefficient estimates synchronization ability, it was explored in its association with learning estimates.

In the *perceptual judgment of the beat task* (CA-BAT), participants listen to superimposed beats on musical excerpts. Each musical excerpt is presented twice, once with the superimposed beats arranged in synchrony with the excerpt, and once with the beats either too fast or too slow in relation to the excerpt. Participants are asked to indicate in which of the two presentations the beats are aligned "on the beat" with the excerpt. In the CA-BAT, each trial is adjusted according to complexity depending on the participant's performance in the previous trial (Harrison & Müllensiefen, 2018a, 2018b). Based on item response theory, the CA-BAT provides two measures: i) an ability estimate that ranges from -4 to +4, in which a higher score indicates better performance; and ii) a standard error of measurement for the ability estimate. This study is the first to implement the CA-BAT in participants with aphasia. We used the psychTestR implementation of this task, version 1.5.1 (Harrison & Müllensiefen, 2018a) with the packages cabat version 0.8.0 (Harrison, 2019a) and psychTestR version 2.9.0 (Harrison, 2019b).

5.1.1.1 Analysis

The primary analysis examined correlations between individual learning estimates and performance in the language, attention, and rhythmic processing measures. Individual learning estimates for each condition were necessary to identify which PWA benefited from the experimental (metronome and stress-aligned) in comparison to the control condition. However, a maximal model that included a condition by session interaction term, nested under participant, did not converge. Therefore, two separate models were used to obtain learning estimates: one for rhythmic-enhanced items and one from the control items[4].

The decision to use two models (rhythmic-enhanced and control) rather than three (metronome, stress-aligned, and control) was motivated by the lack of variability at the group level between the two rhythmic conditions. Therefore, we collapsed across the metronomic and stress-aligned conditions to obtain a learning estimate for the rhythmic-enhanced conditions. This simplification allowed fitting models with the same structure as the model reported in aim 1, targeting two subsets of the data (rhythmic-enhanced and control). All group-level fixed effects estimates were consistent with the model reported above. From these subset models, the random slope for session was used as the by-participant learning estimate for the rhythmic-enhanced and control items. In other words, these models provided estimates of an individual's learning ability

[4]
```
model_rhythm<- glmer(cbind(successes,failures) ~ session +
    (1+session|participant) + (1+condition|sentence) + (1 | observations), family = binomial, data = subset_rhythm)
model_control<- glmer(cbind(successes,failures) ~ session +
    (1+session|participant) + (1|sentence) + (1 | observations), family = binomial, data = subset_control)
```

(rate of improved probe performance over time) separately in the rhythmic-enhanced and control conditions.

Correlations examined how variability in each domain (language, attention, and rhythmic processing) were associated with individual learning estimates for the rhythmic-enhanced and control items. Below, whether the correlations are statistically significant before or after Bonferroni correction is reported. However, due to the nature of these exploratory correlations, special attention was placed on the correlation coefficients and their direction rather than only examining p-values (see Ioannidis, 2005, 2019).

5.1.1.2 Hypotheses and predicted results

The following set of exploratory hypotheses pertain to the relationship between individual *learning estimates* and PWA's *performance in the measures of language, attention, and rhythmic processing*. Some of the associations between learning estimates and performance were expected to affect specific experimental conditions (rhythmic-enhanced and control conditions) in a distinctive way; conversely, some associations between learning estimates and performance did not have an a-priori prediction for specific conditions. Given that these associations have not been systematically tested, the aim of examining individual predictors of script learning in PWA was *exploratory*. The goal of examining them was to identify individual characteristics in PWA that could inform treatment customization when training scripted sentences. These analyses may also reveal preliminary evidence regarding the cognitive mechanisms that support response to scripted sentence learning by identifying abilities or capacities correlated with better or worst learning response.

Regarding *language performance*, it was hypothesized that WAB-AQ would be positively associated with learning performance. This would indicate that PWA with more preserved language ability (higher WAB-AQ scores) would also have higher learning estimates across conditions (rhythmic-enhanced and control conditions). Previous studies show that overall language-impairment scores positively correlate with treatment response, with higher scores (indicating milder aphasia severity) being associated with better recovery (e.g., Liang et al., 2001; Quique et al., 2019). Conversely, a script training study by Lee et al. (2009) found a small negative correlation between scores on the Western Aphasia Battery (WAB-AQ; Kertesz, 1982) and the production of script-related words following treatment. There were no specific predictions about how overall aphasia severity would influence learning estimates for specific conditions.

Regarding *attention*, it was hypothesized that attention performance in the TEA would be negatively associated with learning performance in the rhythmic-enhanced conditions. This would indicate that people with poorer attention performance would benefit more from the beats provided in the rhythmic conditions. The rationale for this hypothesis is that the beats could direct attentional resources towards the stressed syllables and metronomic pulses (Large & Jones, 1999; Pitt & Samuel, 1990), helping to overcome the concomitant attentional deficits that negatively affect PWA's language production (Murray, 1999, 2012). In contrast, PWA with better attention performance, who may not need the same rhythmic support, may perform similarly across the experimental conditions.

Lastly, regarding *rhythmic processing* abilities, it was hypothesized that people with better rhythmic performance would also have higher learning estimates of scripted sentences in the rhythmic-enhanced conditions relative to the control condition. This would indicate PWA use their abilities to detect, integrate, and produce rhythmic features (Phillips-Silver et al., 2010) during speech entrainment, resulting in turn in increased script sentence learning.

5.1.2 Results: behavioral measures associated with scripted-sentence learning response

The learning estimates for the control items and the learning estimates for the rhythmic-enhanced items correlated significantly ($r = 0.58$, $p = 0.03$). This correlation was expected, given that both item sets belong to the same scripted-sentence learning paradigm, involving the same learning procedures for both control and rhythmic-enhanced conditions.

5.1.2.1 Demographics and scripted-sentence learning response

None of the demographic variables correlated significantly with either of the learning estimates (rhythmic-enhanced and control items). See Table 4 for demographic information and Figure 8 for demographic correlations with the learning estimates.

Table 4. Demographic information and language performance.

ID	Gen	Edu	MPO	Age	WAB_{sp}	WAB_{au}	WAB_{rep}	WAB_{n}	WAB_{AQ}	AoS
COL1	F	8	12	47	11	7.45	2.6	3.5	49.1	N
COL2	F	15	108	56	18	8.45	2.4	7	71.7	N
COL3	F	7	15	55	13	8	1.4	5.1	55	Y
COL4	F	17	37	65	13	7.6	5.2	7	65.6	N
COL5	M	20	211	53	14	8.05	8.4	7.8	76.5	Y
COL6	M	16	17	53	3	4.65	5.3	0.6	27.1	Y
COL7	F	12.5	24	56	10	4.25	5.2	2.4	43.7	N
COL8	M	13	24	40	8	7.2	3	4.3	45	N
COL9	M	2	27	58	1	4.8	3.4	1	20.4	Y
COL10	F	14	24	42	11	8	6.4	6.7	64.2	N
COL11	F	10	240	55	9	6	6.4	4.3	51.4	Y
COL12	M	11	42	31	7	6.15	3.5	2.2	37.7	Y
COL13	M	12	14	78	8	7.15	2	2.6	39.5	N
COL14	F	8	372	51	14	7.6	6.5	6.7	69.9	N
Mean		**11.8**	**83.4**	**52.9**	**10**	**6.8**	**4.4**	**4.4**	**51.2**	
SD		**4.6**	**111.4**	**11.2**	**4.5**	**1.4**	**2.1**	**2.4**	**17.0**	

Notes: WAB = Western Aphasia Battery; ID = participant identification; Edu = years of education; Gen = gender; F = female; M = male; Age = age in years; MPO = months post-onset of aphasia; WAB_{sp} = WAB spontaneous Speech Score (ranges between 0-20); WAB_{au} = WAB Auditory Verbal Comprehension Score (ranges between 0-10); WAB_{rep} = WAB Repetition Score (ranges between 0-10); WAB_{n} = WAB Naming and Word Finding Score (ranges between 0-10); WAB_{AQ} = WAB Aphasia Quotient (ranges between 0-100). AoS = Apraxia of speech (Y = yes, N = no); SD = standard deviation.

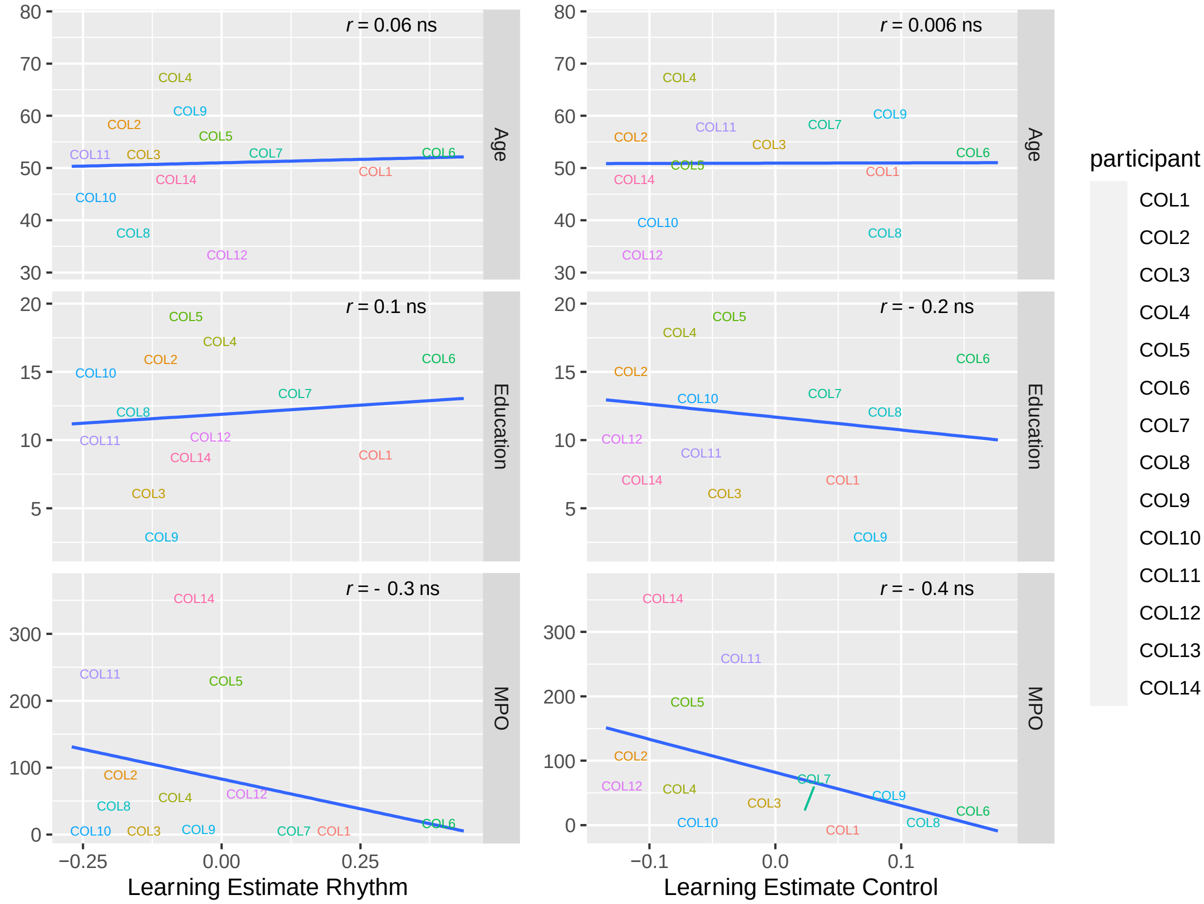

Figure 8 Correlations between demographic information and learning estimates.

Age and Education are measured in years; MPO = months-post onset of aphasia; ns = not significant either before or after Bonferroni correction.

Note that the x-axis has different ranges.

5.1.2.2 Language performance and scripted-sentence learning response

Regarding *language performance*, all participants (n=14) were assessed with the Western Aphasia Battery (WAB; Kertesz, 1982) in its Spanish adaptation (Kertesz & Pascual-Leone, 2000). Aphasia Quotient (WAB-AQ) varied from 20.4 to 76.5, indicating severity ranges from mild to very severe aphasia. Detailed WAB scores, including spontaneous speech, auditory comprehension, repetition, and naming scores per participant are presented in Table 4.

All WAB scores, except for the WAB repetition subscale, correlated significantly with the control-condition learning estimates but not the rhythmic-enhanced learning estimates. However, after Bonferroni correction, only WAB-AQ correlated significantly with the control learning estimate. See Figure 9 for language impairment correlations. The repetition score did not have strong correlations with the learning estimates in either the control or the rhythm-enhanced conditions. It is important to note that the WAB-AQ and comprehension, naming, and spontaneous speech scores share strong correlations ($r = 0.83$, $r= 0.97$, $r = 0.93$, respectively). Interpreting differences between measures that share high variance (collinearity), such as these language measures requires major caution. Therefore, the correlations between these measures and learning estimates are reported separately in Figure 9 but interpreted together in the discussion section.

The trends seen in the slopes (Figure 9) suggest that PWA with more severe language impairments (i.e., lower WAB scores) also had higher learning estimates. One potential cause of this finding could have been that people with more severe aphasia had lower initial performance, and therefore more room to improvement. To test whether PWA with higher aphasia severity started at a lower learning intercept, we examined the correlations between WAB-AQ and the random intercept. This random intercept was obtained from the random effects structure from which the learning estimates were taken (see analysis section 5.1.1.1 above). This uncentered random intercept gives an estimate of learning at the beginning of the short-term learning paradigm. The correlation between WAB-AQ and the random intercept was 0.8 for the rhythmic-enhanced conditions and 0.78 for the control condition. These correlations indicate that PWA with more severe language impairments (i.e., lower WAB-AQ scores) also had a lower intercept across conditions. In other words, PWA with more severe language impairments showed lower performance at the beginning of the short-term learning paradigm, which gave them more opportunity to improve.

The repetition WAB score merits examination separately from the other WAB scores for two main reasons. First, it does not correlate significantly with other WAB scores. In fact, the correlation coefficients between repetition, WAB-AQ, and naming are below 0.36, and below 0.1 for comprehension and spontaneous scores. Second, the repetition WAB score did not strongly correlate with either the rhythmic ($r = -0.07$) or control condition ($r = -0.2$) learning estimates. The low correlation coefficients suggest no meaningful association between repetition tasks and scripted-sentence learning performance through speech entrainment.

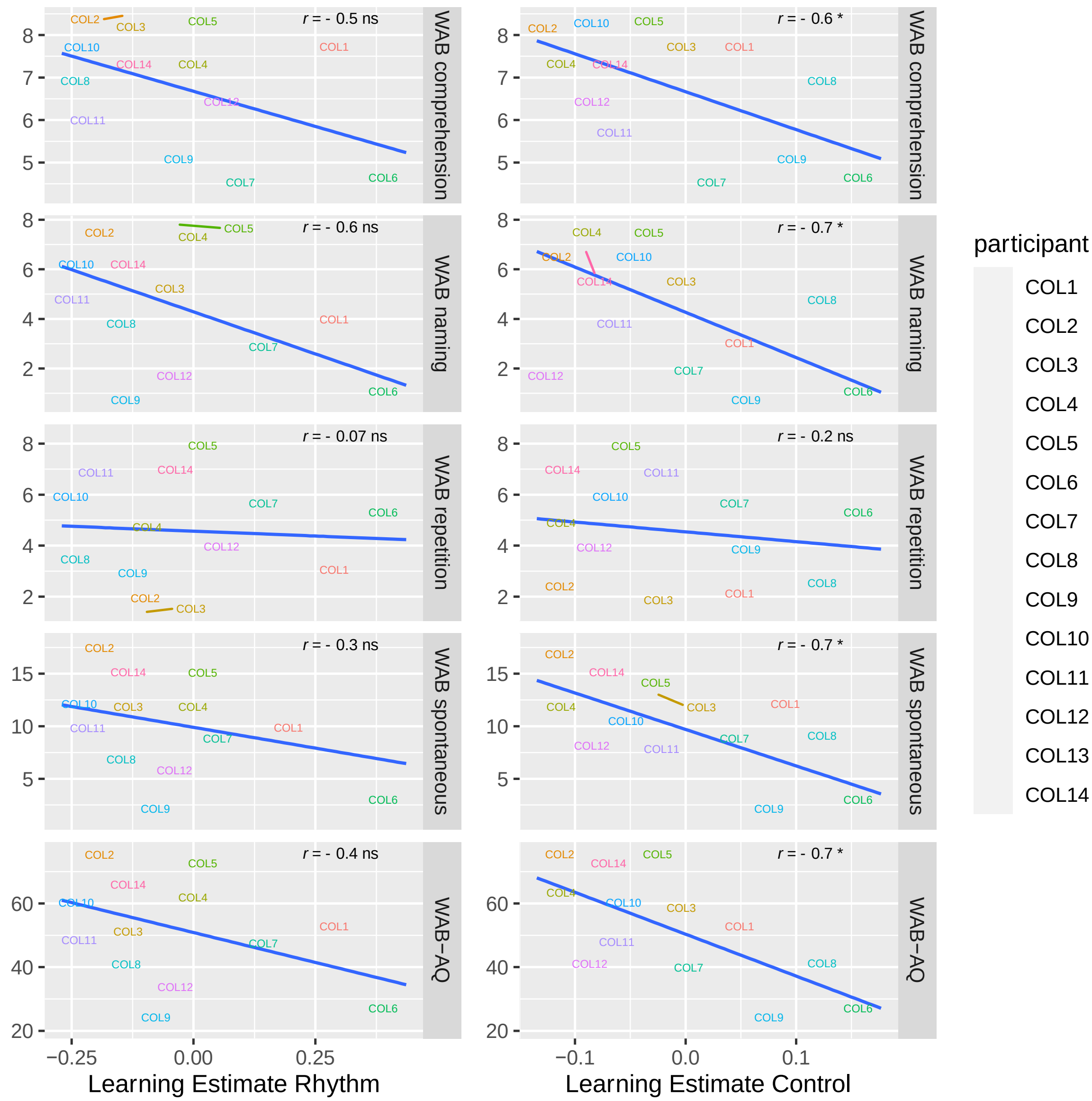

Figure 9 Correlations between language impairments and learning estimates.

WAB = Western Aphasia Battery; WAB Comprehension = Auditory Verbal Comprehension Score (ranges between 0-10); WAB Naming = Naming and Word Finding Score (ranges between 0-10); WAB Repetition = Repetition Score (ranges between 0-10); WAB spontaneous = spontaneous Speech Score (ranges between 0-20); WAB-AQ = WAB Aphasia Quotient (ranges between 0-100); ns = not significant either before or after Bonferroni correction; * = significant before Bonferroni correction. Only the correlation between WAB-AQ and the control learning estimate reached significance after Bonferroni correction. Note that x-axis has different ranges.

5.1.2.3 Attentional performance and scripted-sentence learning response

Regarding *attention*, participants were assessed with the Elevator Counting and the Elevator Counting with Distraction subtests of the Test of Everyday Attention (TEA; Robertson et al., 1994). In the elevator counting subtest, half of the participants performed within the probable impairment range (< 7), and three participants performed within the deficit score range (≤ 5). The group mean was = 6.2, and the standard deviation was = 1.2, indicating low performance in sustained auditory attention. In the elevator counting with distraction, the average age-scaled score for the current group of participants was 6.6 (SD = 2.1), indicating low performance in selective auditory attention. See Table 5 for attention performance scores per participant.

Table 5. Attention performance.

ID	Counting	Distraction	Distraction scaled
COL1	7	5	7
COL2	7	5	10
COL3	7	4	5
COL4	7	4	6
COL5	7	9	10
COL6	7	5	6
COL7	5	2	4
COL8	7	5	9
COL9	7	0	4
COL10	3	1	5
COL11	5	2	7
COL12	6	5	6
COL13	6	1	5
COL14	6	5	9
Mean	**6.2**	**3.8**	**6.6**
SD	**1.2**	**2.4**	**2.1**

Notes: ID = participant identification; Counting = elevator counting subtest (ranges between 0-7); Distraction = elevator counting with distraction subtest (ranges between 0-10); Distraction scaled = age scaled-score equivalents for the elevator counting with distraction subtest (ranges between 0-13); SD = standard deviation.

None of the attention performance variables had high correlation coefficients with either of the learning estimates (rhythmic-enhanced and control). See Figure 10 for attention correlations.

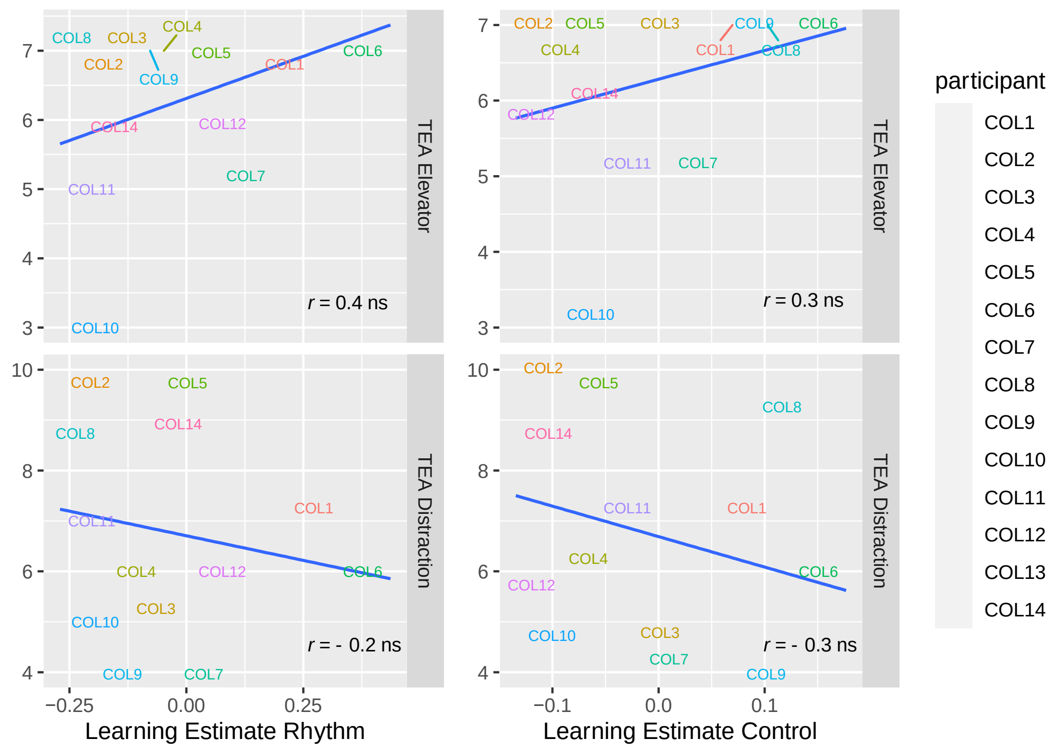

Figure 10 Correlations between attentional performance and learning estimates. TEA Elevator = elevator counting subtest (ranges between 0-7); TEA Distraction = age scaled-score for the elevator counting with distraction subtest (ranges between 0-10). ns = not significant either before or after Bonferroni correction. Note that x-axis has different ranges.

5.1.2.4 Rhythmic processing performance and scripted-sentence learning response

Rhythmic processing performance was measured using multiple tests. The first one was the rhythm subtest of the Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al., 2003). In this test, 78.5% (n=11) of the participants scored below the cut-off of 23 points (Peretz et al., 2003), indicating low abilities in discriminating whether two given rhythms were the same or different; only three participants performed above the cut-off (group mean = 20.4, SD = 3.9). For individual performance scores on this test, see Table 6.

The second test to assess rhythmic processing was the synchronization with the beat of music task from the Beat Alignment Test (BAT; Iversen & Patel, 2008). In this test, participants were asked to tap to the beat of 12 short musical excerpts. As a group, participants displayed continuous and steady tapping to the excerpts (group inter-tap mean = 0.49 seconds, group inter-tap standard deviation = 0.11 seconds), which provides evidence of both an acceptable beat perception and a motor expression of this perceived beat. See Table 6 for inter-tap mean (BAT_{mean}) and inter-tap standard deviation (BAT_{sd}) per participant.

In terms of the correlation between inter-tap mean and tactus level (the critical measure of synchronization with the beat of music), most participants (n=13) had correlation coefficients that ranged between 0.79 and 0.9. Notably, in the normative data from Iversen & Patel, most participants had correlations above 0.9, but older neurotypicals have shown correlations closer to 0.85 and above. Overall, 10 participants in the current sample were above 0.85, indicating good synchronization with the beat of music. See Figure 11 for correlations per participant. An additional participant (COL9) was close to this value (r = 0.84). Lastly, two participants had an acceptable synchronization with the beat (COL7 and COL8), and one participant showed poor performance in this task (COL12, r = 0.64). Without further testing, it is difficult to identify why COL12 had difficulties synchronizing to the beat of music; prior evidence suggests potential disruptions in auditory-motor mapping (Sowiński & Dalla-Bella, 2013). Overall, the correlation coefficient across all participants was 0.95, indicating good performance in synchronization with the beat of music across participants.

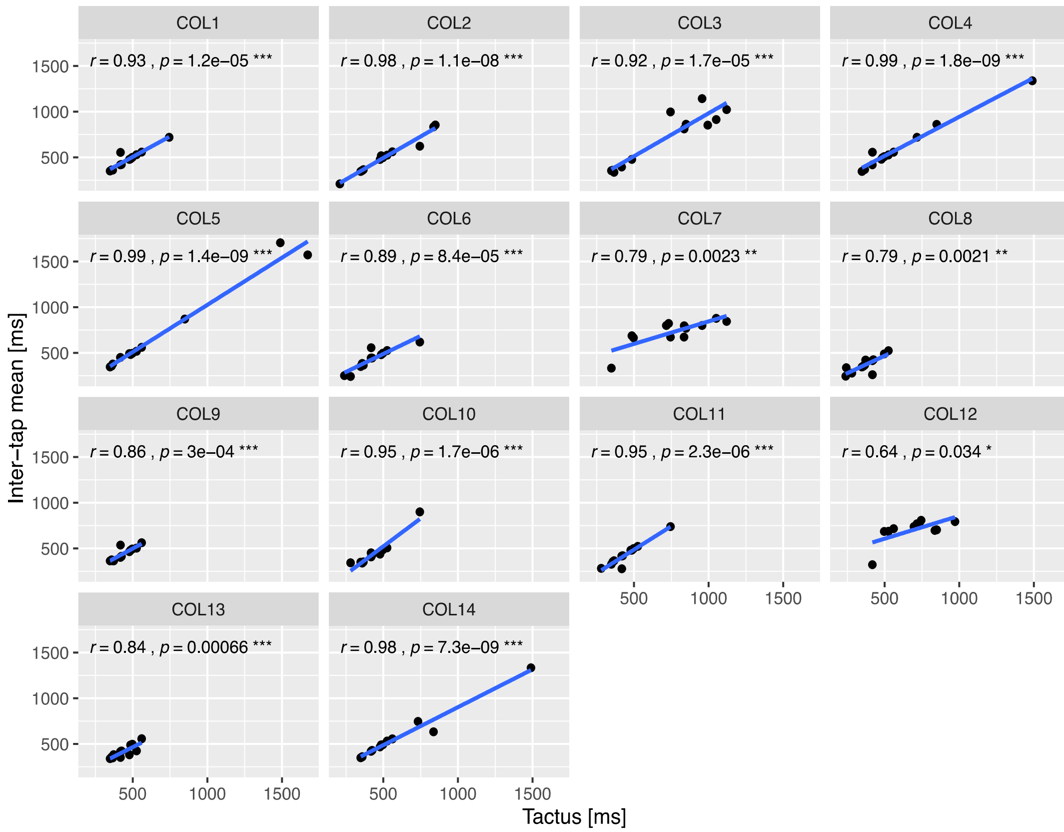

Figure 11 Correlation between inter-tap mean and adjusted tactus per participant. Each data point represents performance for one excerpt (12 excepts in total). Clusters of points indicate that participants repetitively chose to synchronize with a given tactus level. r = correlation coefficient; * significance at < 0.001; ** significance at 0.01; * significance at < 0.05.**

The last test used to assess rhythmic processing was the adaptive version of the perceptual judgment of the beat task (CA-BAT; Harrison & Müllensiefen, 2018). The group mean for the ability estimate was low (mean = -2.376, standard error = 1.081), indicating poor beat perception performance. See Table 6 for ability estimate mean ($BAT_{ability}$) and its standard error of measurement (BAT_{sem}) per participant. To the best of our knowledge, this is the first implementation of the CA-BAT in aphasia; therefore, more research would be needed to enable performance comparisons between groups of participants.

Table 6. Rhythmic processing performance.

ID	MBEA	$BAT_{ability}$	BAT_{sem}	BAT_{mean}	BAT_{sd}	BAT_{corr}
COL1	20	-2.567	1.114	0.477	0.052	0.930
COL2	21	-2.803	0.881	0.465	0.024	0.983
COL3	11	-0.608	0.655	0.629	0.092	0.924
COL4	24	-3.871	2.151	0.544	0.061	0.988
COL5	24	-2.064	0.987	0.553	0.026	0.989
COL6	20	-1.464	0.890	0.407	0.056	0.895
COL7	14	-1.963	1.038	0.69	0.083	0.789
COL8	22	-0.741	0.666	0.354	0.038	0.792
COL9	20	-2.047	0.983	0.438	0.041	0.863
COL10	26	-4.000	2.205	0.426	0.062	0.953
COL11	22	-3.442	1.608	0.407	0.032	0.950
COL12	22	-3.502	1.867	0.671	0.095	0.639
COL13	21	-2.614	1.004	0.408	0.054	0.838
COL14	19	-1.583	0.821	0.518	0.06	0.984
Mean	**20.4**	**-2.376**	**1.205**	**0.499**	**0.055**	
SD	**3.9**	**1.081**	**0.527**	**0.106**	**0.023**	

Notes: ID = participant identification; MBEA = Montreal Battery of Evaluation of Amusia (ranges between 0-31); CA-$BAT_{ability}$ = Perceptual judgment of the beat ability estimate (computed from the underlying item response model, ranges between -4 and 4, approximately); CA-BAT_{sem} = standard error of measurement for the ability estimate. BAT_{mean} = inter-tap mean measured in seconds; BAT_{sd} = inter-tap standard deviation measured in seconds; BAT_{corr} = Correlation between inter-tap mean and target inter-tap intervals adjusting for tactus. SD = standard deviation.

The CA-BAT had a high correlation coefficient with the learning estimate for the control items (r = 0.6), but not with the rhythmic learning estimates (r = 0.2), suggesting that PWA with better beat perception also had higher learning estimates for the control items. However, this correlation did not reach significance after Bonferroni correction. None of the other rhythmic performance variables showed strong correlation coefficients with neither learning estimate (rhythmic-enhanced and control). See Figure 12 for rhythmic processing correlations.

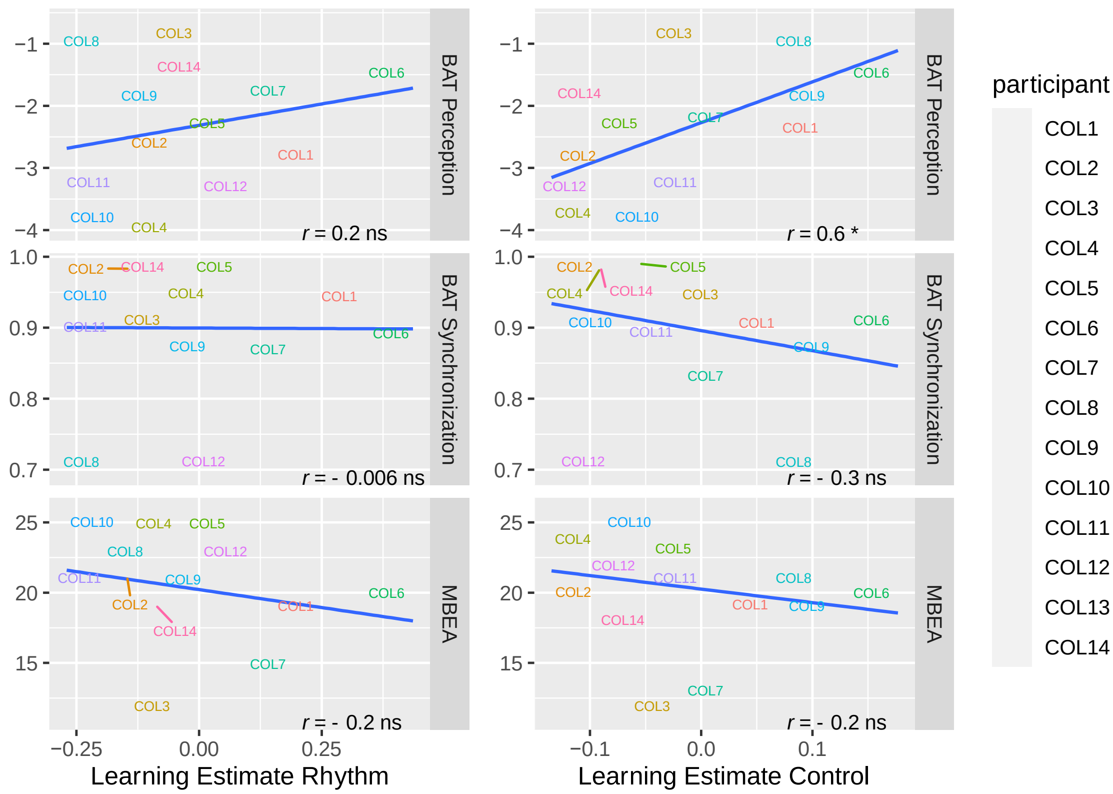

Figure 12 Correlations between rhythmic processing performance and learning estimates. MBEA = Montreal Battery of Evaluation of Amusia (ranges between 0-31); BAT perception = Perceptual judgment of the beat ability estimate from the CA-BAT (ranges between -4 and 4); BAT synchronization = correlation between BAT inter-tap mean and target inter-tap intervals adjusting for tactus. ns = not significant before or after Bonferroni correction; * = significance before Bonferroni correction. The correlation between BAT perception and the control learning Estimate did not reach significance after Bonferroni correction. Note that x-axis has different ranges.

5.1.3 Discussion: behavioral measures associated with scripted-sentence learning response

It is important to note that this study had a relatively small sample size, and that analyses for the second aim were exploratory. However, within the demographic variables, there was a wide range of variability in terms of education, age, and months post-onset, none of which showed a strong relationship with the learning estimates (for either rhythmic-enhanced or control items). Therefore, there is no evidence suggesting that the type of scripted-sentence learning used in the current study is only effective for some individuals. Additional work can further explore the possibility that this type of scripted-sentence learning is equally efficacious for people with chronic aphasia regardless of education, age, or their point in the chronic stage of stroke recovery.

5.1.3.1 Language performance and scripted-sentence learning

The current sample of Spanish speakers with aphasia showed a broad range of language impairments, varying from mild to very severe aphasia as measured by the WAB-AQ. Bearing in mind that all WAB scores are highly correlated, except for repetition, and that these are exploratory analyses, the evidence suggests a preliminary association between language ability and scripted-sentence learning estimates. Contrary to the a priori hypothesis, the correlation coefficients indicate that PWA with more severe language impairments (i.e., lower WAB scores) also had higher learning estimates. A similar finding from the script-training literature reported a negative correlation between WAB scores and production of script-related words (Lee et al., 2009). Conversely, evidence from word-level interventions suggests that PWA with milder language impairments have a higher response to treatment (e.g., Liang et al., 2001; Quique et al., 2019). A potential explanation of these conflicting findings could be found in the learning approach that these interventions use (errorless vs. effortful), the specific tasks these interventions target (propositional vs. rehearsed language production), and in that PWA with more severe impairments started with a lower learning intercept, giving them more room for improvement.

In aphasia, both errorless and errorful approaches have been equivalently effective for object naming (Fillingham et al., 2005a, 2005b, 2006) and for action naming treatment (Conroy et al., 2009c). Also, both errorless and effortful (retrieval practice) approaches benefited PWA during naming treatment, but effortful generated superior benefits during follow up testing (Middleton et al., 2015). Recently, Schuchard and Middleton (2018a, 2018b) posited that effortful approaches could be predominantly helpful for PWA with semantic impairments (i.e., higher-level impairments in mapping meaning or concepts onto words) while errorless could be helpful for PWA with either semantic and/or phonological impairments (i.e., lower level impairments in mapping words onto sounds). In short, both errorless and effortful approaches seem to benefit naming treatment, and specific aphasia profiles could guide which one to choose for treatment. The current study used errorless, effortless approaches given that unison production of speech provides a strong support for language production (minimizing error) while having low language processing demands (minimizing effort). This support provided by speech entrainment likely minimized both error and effort making it especially beneficial for people with more severe language impairments.

Furthermore, rather than training propositional language, scripted training targets chains of rehearsed language (a.k.a., formulaic). Formulaic language is defined as "a sequence, continuous or discontinuous, of words or other meaning elements which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar" (Wray & Perkins, 2000, p. 1). The current study used chains of rehearsed language, *scripted sentences,* which relied on overtraining the memory encoding system. Although it could be the case that all participants learned the scripted-sentences as formulaic language, as it is expected in script training, it could be the case that PWA with milder language impairments had more "interference" from propositional language, which led to the production of words that were related, but that deviated from the target scripted sentence. In contrast, PWA with more severe impairments learned the scripted-sentences as formulaic language, but without “interference” from their highly impaired propositional language. It is also important to recall that PWA with more severe impairments also started with a lower learning intercept, which gave them more opportunities for improvement.

In short, the learning approach (errorless and effortless) and the specific task used in this study (rehearsed language production) can clarify the correlations between language impairment measures and learning estimates. In parallel to the study from Lee et al. (2009), the results from the current study suggest that *when learning scripted sentences*, errorless approaches confer more benefit to PWA with greater language impairments given that they learn and access these sentences as rehearsed language, rather than propositional language. The correlations in the rhythmic-enhanced conditions and control condition showed similar negative slopes and similar correlation coefficients. These similarities suggest that PWA with more severe language impairments had higher performance in all conditions. A shallower slope for the rhythmic-enhanced conditions would have indicated that the use of rhythm for scripted-sentence learning was beneficial for PWA regardless of their language impairment level, but this is yet to be established. Future research should examine whether adding rhythm can reduce the influence of pre-treatment language-impairment status on response to scripted-sentence learning.

An additional and distinct measure from the WAB scores is repetition. The almost flat line representing the correlation between the repetition WAB score and learning estimates may suggest differences among repetition, the retrieval demands generated by probes, and speech entrainment tasks. During probes, participants were asked to produce the learned sentences in response to a cue (the related image unique to each sentence). Therefore, probes were a demanding and effortful task in terms of a longer-term verbal memory, distinct from an immediate repetition task where the phonological loop and short-term verbal working memory have been recently activated (for repetition demands see Nozari, Kittredge, Dell, & Schwartz, 2010). From this perspective, the low correlation coefficients between learning estimates and the WAB repetition score is expected. In the current design, probes did not include a speech entrainment task; instead, PWA were asked to retrieve a sentence in response to its related image. Given this probe administration, the differences and distinct benefits that repetition and speech entrainment may have for scripted-sentence learning remain unclear.

Overall, the current findings encourage future work designed to examine the efficacy of scripted-sentence learning in PWA with a wide range of language impairment severities and profiles, using a broad range of language measures. It also encourages further work examining the differences between repetition and speech entrainment in scripted-sentence learning.

5.1.3.2 Attentional performance and scripted-sentence learning

In the elevator counting subtest from the TEA, half of the participants performed within the probable impaired range (<7). This is similar to a prior report in which approximately half of the participants with aphasia (17 out of 39) performed below 7 (Murray, 2012). The group mean of the current group of participants (mean = 6.2) was also strikingly similar to the same previous study (mean = 6.1; Murray, 2012). Furthermore, in the elevator counting with distraction, the group mean (mean = 6.6, SD = 2.1) was slightly below a previous report in aphasia (mean = 7; Murray, 2012), indicating low performance on sustained auditory attention. Anecdotally, the TEA was challenging for PWA, and it required multiple trials with practice items before the test could be administered. In sum, the group of participants with aphasia in the current study had relatively low performance on auditory attention —selective and sustained— consistent with previous reports of concomitant attentional deficits in aphasia (Murray, 1999, 2012; Schumacher, Halai, & Lambon Ralph, 2019).

It was hypothesized that TEA scores would be negatively associated with learning performance in the rhythmic-enhanced condition given that the beats used in the stimuli could help PWA in focusing on stressed syllables (Attentional Bounce Hypothesis; Pitt & Samuel, 1990) and metronomic beats (Dynamic Attending Hypothesis; Large & Jones, 1999) improving scripted-sentence learning. However, the attentional measures did not have strong correlation coefficients with the learning estimates. This result was due in part to the poor performance and low variability in auditory attention performance shown by the current sample of PWA, the possibility that beats did not provide additional attentional support to learn scripted sentences, or that the current manipulations did not provide disproportionate help to PWA with low attention.

Consistent with previous reports, the current study found evidence of attentional deficits in PWA (Murray, 1999, 2012), but it did not find evidence about how these deficits were associated with scripted-sentence learning. Inconsistent with the predictions of the Attentional Bounce Hypothesis and the Dynamic Attending Hypothesis, there was no evidence to support that the beats added to the sentences (see Figure 3) helped direct attentional resources to the stress syllables and metronomic beats for learning of scripted sentences. In other words, the attentional measures did not show strong correlations with the rhythmic-enhanced and control conditions, which obscures the role that beats have on attentional resources. Although the current results do not suggest it, future research can investigate whether this type of scripted-sentence learning is only effective for some attentional profiles of PWA.

5.1.3.3 Rhythmic processing performance and scripted-sentence learning

Approximately 80% of the current sample performed below the cut-off in the rhythm subtest of the Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al., 2003). This performance indicates reduced abilities in discriminating whether two given rhythms were the same or different. As it will be explained in the following paragraph, there are at least three reasons why the current group of participants had rhythmic perception difficulties.

First, stroke survivors have shown difficulties in music perception abilities as measured by the MBEA (e.g., Sihvonen et al., 2016); the current group could have been a representative sample of these difficulties. Prior evidence reported that in a cohort of stroke survivors (n = 53), MBEA scores were bimodally distributed such that one set of participants (60% a week post-stroke and 43% three months post-stroke) had a distinctive low performance, below the cut-off, and the other had performance above it (Särkämö et al., 2009). Second, and specifically to aphasia, some PWA have shown low performance in the MBEA with some participants performing below the cut-off, evidence of reduced abilities in discriminating whether two given rhythms were the same or different (Racette, Bard, & Peretz, 2006). Third, it could be the case that Spanish speakers in general have low performance in this test, perhaps due to music-specific traits that are not included in this test (e.g., none of the excerpts includes Latin American music). Consistent with this latter explanation, neurotypical Spanish speakers have shown lower performance in the MBEA in comparison to the norms established by Peretz et al., 2003 (Toledo-Fernández & Salvador-Cruz, 2015).

In sum, it is not surprising that the current group of participants had shown low scores in a rhythm discrimination task given that both stroke survivors and stroke survivors with aphasia may have concomitant difficulties in music perception abilities as measured by the MBEA (e.g., Racette et al., 2006; Särkämö et al., 2009). Additionally, it is necessary to examine the psychometric properties of the MBEA in Colombia developing specific norms for this population as it has been done in other places in Latin America (e.g., Nunes-Silva & Haase, 2012) and/or to develop culturally-adapted tests that can also measure rhythm discrimination. Lastly, the current evidence available in Spanish by Toledo-Fernández and Salvador-Cruz (2015) suggests that even neurotypical controls had a lower performance in comparison to the norms established by Peretz et al., 2003.

Similar to the low performance in the rhythm subtest of the MBEA, the current group of participants also had a low performance in the perceptual judgment of the beat task (CA-BAT). The group mean of the current group of participants was -2.4 with a range between -0.6 and -4, indicating poor perceptual judgment of the beat. This low performance can be attributed, at least in part, to the prevalent difficulties in beat perception in PWA mentioned above. For example, a previous report suggested that PWA significantly differ from controls in identifying accurately similar and different rhythmic patterns (Zipse et al., 2014). Additionally, low performance in this specific test could also be associated with the high cognitive demands of the CA-BAT.

The CA-BAT relies strongly on cognitive abilities such as attention to fine perceptual details, working memory (Harrison & Müllensiefen, 2018b), and executive functions. As a reminder, musical excerpts are presented twice in this task, once with the superimposed beats arranged in synchrony with the excerpt, and once with the beats either too fast or too slow in relation to the excerpt; participants are asked to indicate in which of the presentations the beats were in synchrony with the excerpt. This means that participants need to be closely attending to the excerpts in order to perceive detailed information about the beats; they have to hold information in their working memory to compare the two presentations; inhibit responses until after hearing the second presentation; and finally, select and communicate an answer. Consistent with this cognitive load account, the current group of participants also showed attentional deficits (e.g., TEA scores). In short, it is possible that participants had a low performance in this task due to the high cognitive demands of the CA-BAT.

Discriminating whether two given rhythms are the same or different (as measured by the rhythm task of the MBEA) require rhythmic processing abilities, whereas the perceptual judgment of beats (as measured by the CA-BAT) require beat processing abilities. Although the concepts of rhythm and beats are certainly related, they are not the same. In music cognition, *rhythm* refers to the "relative timing of the intervals between note onsets" (Levitin, Grahn, & London, 2018, p. 53). Thus, rhythm is related to the relative length of sounds in terms of timing patterns. On the other hand, *beat* refers to a "perceived periodic pulse that structures the perception of musical rhythm and which serves as a framework for synchronized movement to music" (Patel & Iversen, 2014, p. 1). Thus, beat is a mental representation of a steady pulse (see also Murton, Zipse, Jacoby, & Shattuck-Hufnagel, 2017). However, given these definitions, it is worth noting that participants had low performance in both beat and rhythm processing measures. Therefore, regardless of whether beat or rhythm perception contributed more strongly to how participants perceived the rhythmic-enhanced stimuli in the two conditions, the current participants were relatively impaired in their ability to perceive the beat or the rhythmic regularities in the stimuli in both conditions. Consistent with this, participants appeared to learn very similarly in the two rhythmic-enhanced conditions.

Regarding synchronization with the beat of music, as measured by the BAT, participants showed strong abilities in having a motor expression in response to a beat. It is not surprising that people performed poorly in the rhythm/beat perception tasks (e.g., MBEA, CA-BAT) but had good performance in synchronizing with the beat of music. Synchronization to music has an early development in life (e.g., Phillips-Silver & Trainor, 2005; Zentner & Eerola, 2010), and it seems to be present in the general population (for reviews see Repp, 2010; Repp & Su, 2013) in behaviors such as clapping, tapping their feet, or nodding to the music. Additionally, cases reported in the literature have shown a partial independence between rhythm perception and a motor response to rhythm. Specifically, some people have shown impairments in motor responses to rhythm with preserved rhythm perception (Sowiński & Dalla Bella, 2013). Contrary to this, some people have shown impairments in rhythmic perception but not in the motoric response to rhythm (Bégel et al., 2017). The current study provides evidence in support of the latter, in that the majority of participants had poor performance in rhythm/beat perception tasks while having preserved abilities to synchronize to the beat of music.

Beyond performance in rhythmic and beat processing measures, of foremost interest for the current study was the association between performance in these measures and the learning estimates. Prior studies showed that PWA (n=3) with higher rhythmic abilities, such as performance in the synchronization to the beat of music from the BAT, also benefited from an additional rhythmic cue (tapping) in response to melodic intonation therapy, while PWA with lower rhythmic abilities benefited more from the absence of this rhythmic cue (Curtis, Nicholas, Pittmann, & Zipse, 2020; Curtis & Zipse, 2017). In accordance with these findings, it was hypothesized that people with better rhythmic performance would also have higher learning estimates in the rhythmic-enhanced conditions relative to the control condition. Such findings would have suggested that higher abilities in processing rhythm (Phillips-Silver et al., 2010) could amplify the ability to learn scripted sentences when using metronomic and stress-aligned beats.

Contrary to the hypothesized results, most of the rhythmic processing measures used in the study did not show strong correlation coefficients with either of the learning estimates. The lack of meaningful associations between the rhythmic processing performance measures (e.g., MBEA) and learning estimates could be explained in part due to the lack of variability in performance. This lack of variability makes it difficult to establish a specific pattern in relation to the learning estimates. The only strong correlation coefficient was found between the CA-BAT performance and the learning estimate for the control items, indicating that PWA with better beat perception also had higher learning estimates in the control condition. Although this isolated finding requires further investigation, its interpretation here warrants major caution as it could be a case of Type I error. A cautious interpretation is that better abilities in perceiving the beat helped PWA take advantage of speech entrainment to learn scripted sentences. Consistent with this possibility, lower abilities in perceiving the beat of music are perhaps not as important when rhythmic support is provided, as in the case of the rhythmic-enhanced conditions.

As noted above, most participants in the current study showed poor performance in rhythm/beat perception tasks. However, this performance did not interfere with their ability to do speech entrainment successfully. A prior study showed that people with marked difficulties in beat perception (a.k.a. beat-deaf) also had difficulties in entrainment to speech (Lagrois, Palmer, & Peretz, 2019). This was not the case in the current group of participants. As seen in Table 2, participants benefited from the speech entrainment task equitably across different conditions achieving a high syllable-level performance in the independent production step during training (above 70% correct) and also during probes (see also Table 3). This suggests that despite poor performance in rhythm/beat perception tasks, PWA benefited from speech entrainment to achieve a higher production of scripted sentences.

In summary, none of the demographic variables, despite a wide range of variability, showed a strong relationship with the learning estimates. In terms of language performance, more severe language impairments were associated with higher learning estimates (i.e., people with more severe aphasia showed faster learning). This association was interpreted in terms of PWA with more severe language impairments starting at a lower learning intercept (giving them more improvement space), the specific learning approach (errorless) and the task used in the study (rehearsed language production). Specifically, errorless approaches confer more benefit to PWA with more severe language impairments, *when learning scripted sentences*, given that they learn and access these sentences as rehearsed language with no “interference” from their highly impaired propositional language. In contrast, PWA with milder aphasia showed lower learning estimates because their the ability to produce propositional language allowed for retrieval interference in probe performance. Regarding attentional performance, PWA showed a low performance on auditory attention —selective and sustained—in line with previous reports (Murray, 1999, 2012), but there was no evidence for how these deficits were associated with scripted-sentence learning. Finally, participants showed poor performance in rhythm/beat perception tasks without impairments in the synchronization with the beat of music. The only strong correlation between performance on these tasks and learning estimates indicated that higher beat perception was associated with higher learning estimates in the control condition. Further research using additional rhythmic-processing tests is needed to identify which aphasia profiles benefit more from rhythmic support when learning scripted sentences.

5.2 Cortical tracking and its association with scripted-sentence learning

As mentioned above, identifying individual electrophysiological measures associated with scripted-sentence learning may also improve treatment customization. One such potential measure is cortical tracking of the speech envelope, the coupling of low-frequency *neural oscillations* and the rhythmic properties of the *speech envelope* (Ding & Simon, 2014). Cortical tracking reflects an alignment between oscillatory brain responses and rhythms present in linguistic input, thereby enabling efficient processing of acoustic and linguistic features (Peelle & Davis, 2012). For scripted-sentence learning, cortical tracking may index individual perception of speech-rhythmic properties in connected sentences (Hamilton & Huth, 2018), which could be associated with an individual's ability to learn sentences using speech entrainment. In other words, cortical tracking provides a measure of perception of speech rhythmic features that can be associated with scripted sentence learning, particularly given the use of rhythmic enhanced features during speech entrainment. This study will be the first to investigate cortical tracking of the speech envelope in stroke survivors with aphasia; therefore, the aim of investigating its association with scripted-sentence learning is entirely exploratory.

5.2.1.1 Cortical tracking: phase reset and phase-lock for processing efficiency.

Cortical tracking of the speech envelope receives multiple names in the literature. Some of them are speech-brain entrainment, neural entrainment, cortical entrainment, neural tracking, and envelope entrainment. All of these names refer to the same coupling or synchronization between neural oscillations and the speech envelope (Ding & Simon, 2014). For the purpose of this study, we will use the term cortical tracking.

To better conceptualize cortical tracking as an interaction between neural oscillations and sensory rhythms like speech (Lakatos, Gross, & Thut, 2019), it is necessary to review i) what are neural oscillations and ii) what is the rhythmic information contained in the speech envelope. The next paragraphs attempt to introduce neural oscillations, followed by a section that reviews the speech envelope's information. The evidence presented here will be brief and selectively focused on the aim of exploring the potential association of cortical tracking with scripted-sentence learning in PWA.

Neural oscillations are a characteristic of brain activity that suggests that rhythm is an underlying feature of neural activity (Lakatos et al., 2019). This characteristic is a likely consequence of the highly rhythmic auditory environment that surround human beings (Kösem & van Wassenhove, 2017) such as speech, music, environmental sounds, internal rhythms (heartbeat, breathing), and related auditory inputs (Hickok, Farahbod, & Saberi, 2015). Furthermore, humans have a natural tendency to externally synchronize with auditory rhythmic inputs by clapping, walking at a certain pace, or dancing (e.g., Tranchant, Vuvan, & Peretz, 2016). This tendency to synchronize is also reflected internally such that neural oscillations entrain to (i.e., track) rhythmic sensory input, to the level that it seems to be a fundamental characteristic of the auditory system (Kösem & van Wassenhove, 2017).

In an oversimplification of what occurs in the brain, a group of neurons exhibits synchronous fluctuations in response to sensory stimuli (Assaneo et al., 2019; Poeppel & Assaneo, 2020). These synchronous fluctuations can be measured directly from the brain using electrocorticography (ECoG) or using surface-level measures such as electroencephalography (EEG) and magnetoencephalography (MEG). The synchrony of these fluctuations indicates excitatory shifts in a group of neurons, such that when the onset of sensory input is detected, neuronal populations reset, or better adjust their oscillatory phase, to align with a given rhythmic input (Lakatos et al., 2019; Peelle & Davis, 2012). Thus, detecting a rhythmic stimulus (e.g., speech, music) results in the phase reset of the oscillatory activity in a population of neurons. See Figure 13 for a schematic representation of this phase reset adjustment. Furthermore, after the phase reset, neural populations display *phase-lock activity*, or *cortical tracking*, to the received input (Peelle & Davis, 2012). This tracking represents the "adjustment of one quasiperiodic system (neural oscillations) to *match the phase* of an external periodic or quasiperiodic stimulus (e.g., speech rhythm)" (Peelle & Davis, 2012, p.4).

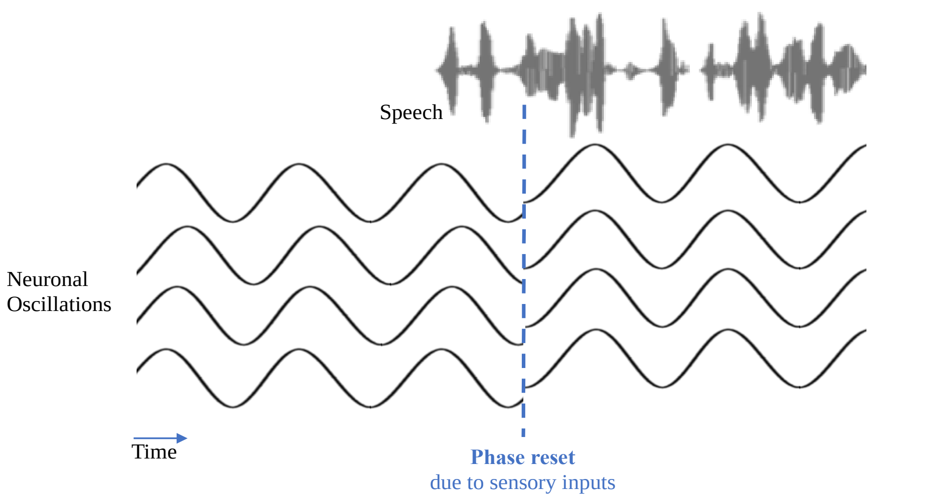

Figure 13 Schematic representation of a neuronal oscillation phase shift (reset) and phase-lock as a result of speech rhythm.Temporal regularity in sensory stimuli generates a phase shift on neuronal oscillations to allow information (e.g., speech) to be processed more efficiently. Figure adapted from Peelle & Davis (2012).

Neuronal oscillations are an indicator of excitatory activity and are present in both resting and stimulus processing. Importantly, sensory input with a regular structure is processed more efficiently when neural populations align with it at a time where the neuronal group is at a high excitability phase (Lakatos et al., 2019; Peelle & Davis, 2012). For example, animal models shows that both visual and auditory information reaching a neuronal group at an excitatory phase is processed more efficiently (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Lakatos et al., 2009; Womelsdorf, Fries, Mitra, & Desimone, 2006). A schematic representation of the relationship between processing efficiency and oscillatory neuronal activity is presented in Figure 14. Overall, neural oscillations, specifically low-frequency oscillations, are essential in packing information into temporal units for an efficient online processing of auditory stimuli (Giraud & Poeppel, 2012; Poeppel & Assaneo, 2020).

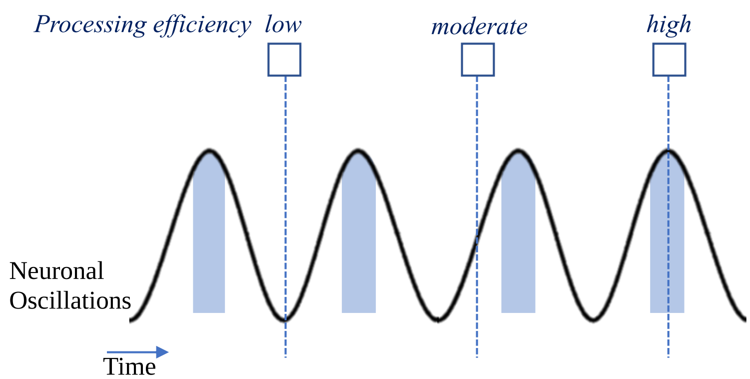

Figure 14 Schematic representation of the relationship between processing efficiency and neuronal oscillations phase.

Light blue vertical rectangles represent higher neuronal excitability. Sensory stimuli that are processed at a higher excitability phase are processed more efficiently. Adapted from Peelle & Davis (2012).

Notably, low-frequency oscillatory activity in the auditory cortex synchronizes to the rhythmic features of the speech input through cortical tracking (Ding, Chatterjee, & Simon, 2014; Notbohm & Herrmann, 2016; Poeppel & Assaneo, 2020). In other words, the brain responds to the temporal structure of speech by performing brain-to-stimulus synchronization (Poeppel & Assaneo, 2020) that is likely to facilitate the processing of acoustic information. Furthermore, the auditory cortex does not only synchronize to the temporal features of speech, but it also shows a "preference" for frequencies that correspond to speech characteristics such as the syllabic rate (Boemio, Fromm, Braun, & Poeppel, 2005; Giraud et al., 2000; Liégeois-Chauvel, Lorenzi, Trébuchon, Régis, & Chauvel, 2004). The following section provides more details regarding the rhythmic features found in the speech envelope to which the brain responds with cortical tracking.

5.2.1.2 A note on the speech envelope

Rosen (1992) posits that the speech signal encompasses three principal temporal features: periodicity, fine structure, and envelope (see Figure 15 for an overview of their primary characteristics). *Periodicity* refers to the fluctuations of a sound wave with periodic or aperiodic characteristics. Periodicity conveys linguistic information such as voicing and manner, and also information about intonation as it pertains to patterns of syllabic, word, and sentence stress. *Fine-structure* refers to variations of the wave shape within a periodic sound, or within a short interval of an aperiodic sound. The primary auditory correlate of fine-structure is timbre, and it mainly conveys linguistic properties about the place of articulation and vowel identity; it also carries information about voicing and the manner of articulation (Rosen, 1992).

The last temporal feature is the *speech envelope,* which "refers to the acoustic power at a given time in a given frequency range" (Peelle & Davis, 2012; p. 2). The speech envelope is generally associated with the temporal information of a speech sound wave and can be described with acoustic features such as intensity, duration, rise and fall time; and their auditory correlates, loudness, length, attack, and decay. The envelope represents a smooth curve over the contour of the fluctuations in the acoustic energy of speech or any other oscillatory signal (see a depiction of a speech envelope in Figure 16). The relative amplitude of the envelope indicates the onset and offset of linguistic units (e.g., syllables; Greenberg, Carvey, Hitchcock, & Chang, 2003). Given that the envelope contains durational information, it also contains speech rhythm variations (word and sentence stress) and tempo, as it relates to rate variations in segmental and prosodic information (Goswami & Leong, 2013; Peelle & Davis, 2012).

Linguistic Characteristics			*Temporal Features*		
			Periodicity	Fine-structure	Envelope
	segmental	voicing	⬤	●	•
	segmental	manner	●	•	⬤
	segmental	place		⬤	
	segmental	voice quality		⬤	•
	prosodic	tempo, rhythm			⬤
	prosodic	syllabicity			⬤
	prosodic	stress	⬤		•
	prosodic	intonation	⬤		

Figure 15 Temporal features and linguistic characteristics conveyed by the envelope, periodicity, and fine-structure of speech.

Note that the size of the circle indicates the extent to which the temporal feature conveys the linguistic contrast; blank spaces indicate a weak or non-existent cue. Adapted from Rosen (1992).

Low-frequency variations of the speech envelope (2-7 Hz) are particularly important for the current study as they convey speech segmental information related to manner of articulation, word stress, and syllabification (Rosen, 1992). These low-frequency oscillations can arise from the jaw's repetitive movement in combination with voicing (Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009). Despite that speech is produced by multiple articulators, each with precise and distinct movements, their kinematics are not working independently, but rather coordinate to achieve speech production. In consequence, the speech motor system has a strong regularity in the time domain that ranges from 2-7 Hz (Poeppel & Assaneo, 2020). This regularity corresponds to the syllable rate, and this is why *syllabic information –including syllabification and word stress– has a substantial influence on the speech envelope* (Greenberg et al., 2003). Notably, a syllable rate between ~3-7 Hz has been found across multiple languages (e.g., English, French, Italian, Spanish; Pellegrino, Coupé, & Marsico, 2011). Therefore, these low-frequency oscillations will also be present in Spanish, the target language of the current study.

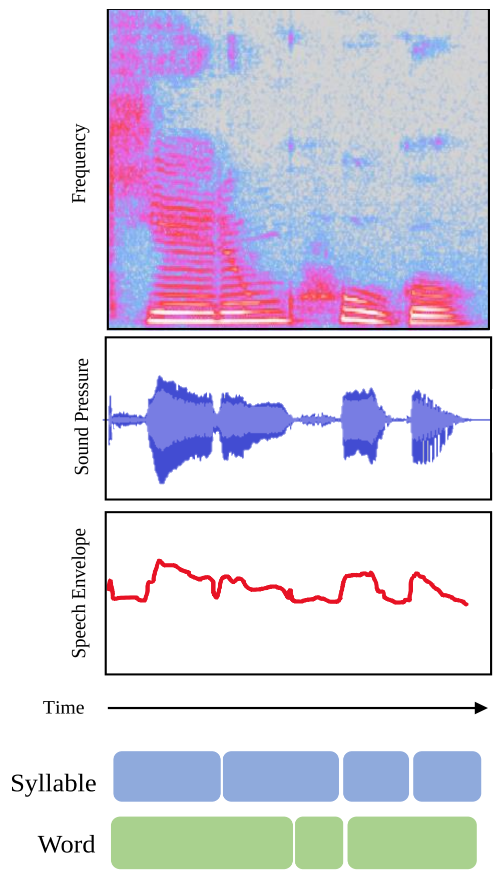

Figure 16 Multiple representations of acoustic and linguistic information of a sentence.

Note: The represented sentence is "Quiero un jugo" [*I want a smoothie*]. Figure adapted from Peelle & Davis (2012).

To summarize, the speech signal encompasses three principal temporal features: periodicity, fine structure, and envelope. The low-frequency variations of the speech envelope represent syllabic and word stress characteristics of speech (Rosen, 1992). Particularly, these low-frequency variations correspond to the syllabic range (~3-7 Hz) found in multiple languages, including Spanish (Pellegrino et al., 2011; Poeppel & Assaneo, 2020). Given that syllables and other rhythmic features of speech are prominent in the speech envelope, the brain capitalizes on them for speech processing (Poeppel & Assaneo, 2020). Specifically, neural oscillations synchronize with rhythmic features in the speech signal through a phenomenon called cortical tracking of the speech envelope (Ding, Melloni, Zhang, Tian, & Poeppel, 2016; Ding & Simon, 2014; Kubanek, Brunner, Gunduz, Poeppel, & Schalk, 2013). In fact, cortical tracking is most observed at frequencies below 10 Hz (Giraud & Poeppel, 2012; Gross et al., 2014). The next section will expand on frequency ranges of neural oscillations and their association with acoustic processing.

5.2.1.3 Cortical tracking in the delta, theta, and alpha bands

Neural oscillations are classified in five broad frequency ranges: delta (< 4 Hz), theta (∼4–8 Hz), alpha (∼8–13 Hz), beta (∼13–30 Hz), and gamma (> 30 Hz). It is known that the auditory cortex shows an enhanced neural oscillatory response for low frequencies, mainly related to syllabic rate (2-8 Hz; Boemio et al., 2005; Giraud et al., 2000; Liégeois-Chauvel et al., 2004), which broadly corresponds to the theta band frequencies (∼4–8 Hz). Given this enhancement in the theta band, and that the interest of this study circles around the rhythmic features of speech (e.g., word stress), the analysis focus will have an especial interest on the theta band, although we will also examine the delta and alpha bands. It is important to note, however, that the specific functions that each of these bands play in information processing is still under current examination.

The following paragraphs will provide a *brief introduction* about the role of neural oscillations in each of the bands, in order to set the ground for the current exploratory analysis.

The *delta band* (~1 to 4 Hz) aligns closely with the onset of words in connected speech. Some studies have shown an association of cortical tracking in this band with high-order language processing (Molinaro & Lizarazu, 2018). For example, it was shown that beyond following the syllabic rate in the theta band, native speakers also showed a significant oscillatory cortical activity at the sentence and phrasal rates in the delta band (e.g., Ding et al., 2016). Therefore, comprehension would seem to be crucial for cortical tracking in the delta band. Recent evidence suggests that cortical tracking in this band may aid in speech comprehension in noise (Etard & Reichenbach, 2019). Still, the specific link between neural oscillations in the delta frequency band and speech comprehension is an area of active study (e.g., Weissbart, Kandylaki, & Reichenbach, 2019).

The *alpha band* (~8 to 12 Hz) has been found to be modulated by attention such that there is lower cortical tracking in this band to unattended auditory stimuli (Frey, Ruhnau, & Weisz, 2015; Kerlin, Shahin, & Miller, 2010). In an additional cognitive domain, alpha oscillations have been associated with working memory such that participants that have good performance in anticipating auditory distractors also show an increased oscillatory responses in the alpha band (Bonnefond & Jensen, 2012). Furthermore, speech intelligibility, and perhaps speech comprehension, may be associated with alpha-frequency oscillatory responses (Obleser & Weisz, 2012). Despite the growing evidence regarding the alpha band contributions to comprehension, the precise association between cortical tracking of the speech envelope and comprehension is yet to be established (McHaney, Gnanateja, Smayda, Zinszer, & Chandrasekaran, 2020).

As mentioned before, the *theta band* (~4 to 8 Hz) aligns closely to the syllabic rate (~3-7 Hz) found in multiple languages, including Spanish. Some evidence has shown that theta-band oscillations help in the segmentation of auditory information (Teng, Tian, Doelling, & Poeppel, 2018). This segmentation is likely to organize acoustic input into chunks that facilitate their processing (Teng et al., 2018). Specifically, theta oscillations would help process chunks of acoustic information related to phonological and syllabic information (Kösem & van Wassenhove, 2017). Moreover, the auditory cortex has shown a cortical tracking enhancement for frequencies that correspond to the syllabic rate (2-8 Hz), which closely aligns with the theta band (Boemio et al., 2005; Giraud et al., 2000; Liégeois-Chauvel et al., 2004).

Cortical tracking of the speech envelope in the theta band has been observed in unintelligible speech, indicating that although it is associated with parsing information, it may not be directly related to comprehension (e.g., Howard & Poeppel, 2010; Pefkou, Arnal, Fontolan, & Giraud, 2017). For example, speakers that do not understand Chinese show responses to the syllabic acoustic rhythm of Chinese sentences (Ding et al., 2016). Consistent with these findings, Etard and Reichenbach (2019) presented listeners with stimuli in different temporal lags and found a significant role of the theta band encoding acoustic-related properties of the signal rather than higher-level speech comprehension. Nonetheless, when comparing Spanish or English as a native vs. foreign language, a higher synchronization in the theta-band is found for listening to a native language (Pérez, Carreiras, Gillon-Dowens, & Duñabeitia, 2015). Recent findings have also shown that current stimulation in the theta-band enhances speech comprehension compared to sham stimulation (Keshavarzi, Kegler, Kadir, & Reichenbach, 2020). In sum, it is known that neural oscillations in the theta band aligns with the syllabic rate and that it helps to segment auditory input; however, its top-down and bottom-up role in acoustic and linguistic processing still under study.

Before closing this section, it is important to highlight that cortical tracking can reflect both bottom-up and top-down information processing (Ding et al., 2014; Poeppel & Assaneo, 2020). The evidence that discusses these types of processing is extensive and is under current debate. The present study does not aim to provide evidence in favor of top-down or bottom-up processes; thus, there will not be an extensive discussion on this topic. Some studies have suggested that cortical tracking reflects the processing of rhythmic-acoustic information independent of high-level linguistic information. In support of this *bottom-up processing*, there have been i) observations of cortical tracking to musical stimuli (e.g., Doelling & Poeppel, 2015). Moreover, some researchers have reported ii) similar degrees of cortical tracking in rhythmically analogous music and speech stimuli, particularly in cases where people have had less musical training (Harding, Sammler, Henry, Large, & Kotz, 2019). Finally, iii) cortical tracking has been observed in unintelligible speech (Howard & Poeppel, 2010; Millman, Johnson, & Prendergast, 2015). Overall, some evidence suggests that cortical tracking may reflect acoustic processing rather than higher-order linguistic processing.

On the other hand, evidence in favor cortical tracking supporting top-down speech processing comes from i) modeling approaches in which linguistic information contributes significantly to the degree of cortical tracking in both isolated sentences (Di Liberto, O'Sullivan, & Lalor, 2015) and natural connected speech (Di Liberto & Lalor, 2017). In addition, some evidence has shown ii) that cortical tracking depends on attentional resources to speech (Ding & Simon, 2012) such that the degree of cortical tracking increases in response to attended speech (Hambrook & Tata, 2014). Finally, there have been observations of iii) a significant enhancement of cortical tracking to intelligible sentences compared to unintelligible sentences with equivalent speech envelopes (Peelle et al., 2013). Together, these findings suggest that cortical tracking also may also play a role in facilitating the processing of higher-order features of the spoken language.

5.2.1.4 Cortical tracking and aging

Most of the literature in cortical tracking has been done in healthy adults, with only a handful of studies examining older adults. The existent studies have reported significant differences in the patterns of cortical tracking in older adults; particularly an enhanced oscillatory activity without the benefits of improving auditory processing, when compared to younger adults (e.g., Brodbeck, Presacco, Anderson, & Simon, 2018). Potential explanations for these recent observations are described below. As mentioned above, to the best of our knowledge, this is the first report of cortical tracking in stroke-related aphasia. As such, it is unknown whether PWA will have cortical tracking responses resembling older adults (although this study does not have a control group of neurologically healthy adults to drive a direct comparison), or if due to the language system impairments, cortical tracking would display unique responses with distinct characteristics unknown until now.

Cortical tracking, particularly in the 4 Hz amplitude modulation (i.e., theta band), has been shown to be enhanced by approximately 30% in older adults when compared to younger people (Herrmann, Buckland, & Johnsrude, 2019). Studies using Auditory Steady-State Response (ASSR; a measure that reflects synchronization of brain oscillatory activity with acoustic input that coincide in frequency) reported higher theta synchronization in older adults when compared to young adults with clinically normal hearing (Goossens, Vercammen, Wouters, & van Wieringen, 2016). This higher synchronization in older adults has been seen at the cortical level, in both quiet and noise conditions, despite a significantly lower amplitude responses at the midbrain level

(Presacco, Simon, & Anderson, 2016a, 2016b). In the alpha band oscillatory activity, older adults have shown an enhancement when listening to unattended stimuli (passive listening), while young adults show attenuation of the oscillatory response in this task (Henry, Herrmann, Kunke, & Obleser, 2017). In contrast, older adults showed a significant decrease in alpha tracking during active listening in comparison to young adults (Henry et al., 2017). Therefore, it seems that both theta and alpha oscillatory activity change substantially with age.

Fuglsang, Märcher-Rørsted, Dau, and Hjortkjær (2020) reported an amplified cortical tracking of the speech envelope when older adults with age-related hearing loss listened to speech. In this study, Fuglsang et al. (2020) compared cortical tracking in older adults with and without hearing loss using a two competing talker condition and a single talker condition. Older adults with hearing loss showed a significant enhancement of cortical tracking in both the single-talker and the two-talker listening conditions. Fuglsang et al. (2020) hypothesized that this enhancement could be a consequence of peripheral or central hearing damage, which makes the auditory system hyper-excitable (e.g., Heeringa & van Dijk, 2019; Kale & Heinz, 2010).

Some additional explanations for this enhancement have been explored. For example, older adults may recruit higher-order cortical regions to compensate for age-related deficits resulting in an enchantment of cortical tracking (Peelle, 2018; Reuter-Lorenz & Cappell, 2008; Wingfield & Grossman, 2006). Similarly, Brodbeck et al. (2018) posited that older adults may exhibit an enhanced cortical tracking due to amplified representations of any auditory features in the auditory cortex. Goossens et al. (2016) interpreted an observed increase in oscillatory synchronization in the theta band in older adults in light of predictive coding concepts, which states that the brain generates expectations about changes in sensory input based on prior knowledge (Friston, 2005; Sohoglu, Peelle, Carlyon, & Davis, 2012); an increase in brain oscillatory response may be the result of "the presence of a larger predictive error component in the neural response" (Goossens et al., 2016). It has also been proposed that older adults might show an increase in listening effort leading to an enhance cortical response (Peelle, 2018). Overall, multiple potential interpretations have been posited for the observed increment in cortical tracking responses in older adults, but an ultimate explanation of this phenomenon or its specific effects in language is still unknown.

In sum, studies examining cortical tracking in older adults have shown converging evidence of an enhanced oscillatory activity in this population. This oscillatory activity pattern does not seem to improve auditory processing of the speech signal (e.g., Brodbeck et al., 2018). The enhancement in oscillatory activity has been observed in the theta band (Goossens et al., 2016; Herrmann et al., 2019) and in passive listening tasks in the alpha band (Henry et al., 2017). Some hypotheses have been built to explain the enhancement of cortical tracking in this population. For example, age-related hearing damage could increase excitability in the auditory system (Fuglsang et al., 2020), compensatory responses to age-related difficulties can lead to the recruitment of higher-order cortical regions (Peelle, 2018; Reuter-Lorenz & Cappell, 2008; Wingfield & Grossman, 2006), and an increase in listening effort could also result in an enhanced cortical response to auditory stimuli (Peelle, 2018). It is currently unknown whether these enhanced responses are observed in people with stroke-related aphasia; similarly, it is unknown if cortical tracking would be related to language performance or scripted-sentence learning in this population. In other words, cortical tracking may be increased in PWA (similar to the observations reported for older adults) without an additional improvement to auditory processing, or it may be associated with scripted-sentence learning. There are reasons to predict both, but it is an open question.

5.2.1.5 Cortical tracking and scripted-sentence learning in aphasia

Cortical tracking may index PWA's perception and processing of speech-rhythmic features. An efficient processing of these features can in turn facilitate speech entrainment, the unison production of sentences by patient and clinician (Fridriksson et al., 2015; Fridriksson et al., 2012). As mentioned above, speech entrainment relies on the detection, integration, and production of rhythmic features (e.g., Phillips-Silver et al., 2010). Therefore, the perception and processing of speech rhythmic features as provided by the speech envelope may support speech entrainment, and consequently, scripted-sentence learning. In the current study, the perception of these speech-rhythmic features was examined by comparing cortical tracking to the speech envelope of an audiobook with randomly generated cortical tracking obtained from mismatched permutations of EEG responses and the envelopes of the audiobook. This comparison provided key evidence of the association between cortical tracking and a specific speech input rather than a chance response.

In addition to perceiving and processing speech-rhythmic features, cortical tracking may also indicate an efficient processing of higher-order linguistic features (Di Liberto & Lalor, 2017; Di Liberto et al., 2015; Peelle & Davis, 2012), which could be associated with PWA's ability to learn sentences. In aphasia, successful tracking of these higher-order features should promote efficient language processing and therefore, scripted-sentence learning. We will compare cortical tracking estimates with WAB comprehension scores and WAB-AQ to explore if language comprehension ability or overall aphasia severity are associated with cortical tracking. Also, as described in the methods below, PWA listened to a narrative and answered questions about it. Additional analyses examined whether individual PWA's cortical tracking was related to their accuracy in answering these comprehension questions.

Notably, cortical tracking may index individual perception of speech-rhythmic properties in connected sentences (Hamilton & Huth, 2018) that are more closely related to naturalistic speech that PWA commonly hear (as opposed to oversimplified and highly controlled stimuli). In neurolinguistics and neurosciences, the step towards natural stimuli has been a revolutionary approach that has improved the understanding of how the brain processes language (Hamilton & Huth, 2018). The current study uses a narrative (e.g., an audiobook) as a stimulus for the cortical tracking measure. Following Hamilton & Huth (2018), a narrative story has a more naturalistic approach in comparison to controlled isolated sentences given that this is a stimulus that people might reasonably be exposed to outside of an experimental setting, it appears in the same context as it would in real life, and there is an existent motivation for perceiving and understanding this type of stimulus, outside of experimental purposes (though see Alexandrou et al., 2018, for a critique of using audiobooks as an analog to naturalistic speech).

Lastly, it is important to note that the aim of identifying individual characteristics in cortical tracking associated with learning response is exploratory. The existing cortical tracking evidence in neurotypicals guided the procedures and predictions of this aim, but the current study is the first to examine cortical tracking in stroke survivors with aphasia. As mentioned before, it is unknown whether cortical tracking may be increased in PWA (as it happens in older adults) without an additional benefit for auditory processing, or if cortical tracking may be associated with scripted-sentence learning response. Simply determining whether PWA show evidence of cortical tracking is informative.

5.2.2 Methods: cortical tracking associated with scripted-sentence learning response

5.2.2.1 Stimuli: audiobook and Auditory Long Latency Responses.

Participants heard a Spanish adaptation of the audiobook "Who was Albert Einstein" (Brallier, 2002), which has a reading level between third and fourth grade. The story was read by a male native Spanish speaker using natural prosody and clear speech (Krause & Braida, 2004; Uchanski, 2005). To avoid long pauses, silences larger than 500 ms were truncated using Audacity® software, version 2.3.3 (Audacity Team, 2019). This process resulted in a 9-minute audiobook excerpt, which was split into ~1-minute segments. Each segment began and ended with complete sentences and had a mean syllable rate of ~6.3 Hz (see Figure 17). To ensure attention, participants answered one yes/no comprehension question after each segment. The audiobook and questions were presented using PsychoPy version 3.2.4 (Peirce et al., 2019) using the USB audio interface Behringer U-Phoria UMC404HD. The nine segments were heard chronologically preserving the storyline. A similar approach was recently used to capture cortical tracking in people with primary progressive aphasia (Dial, Zinszer, Chandrasekaran, & Henry, 2018).

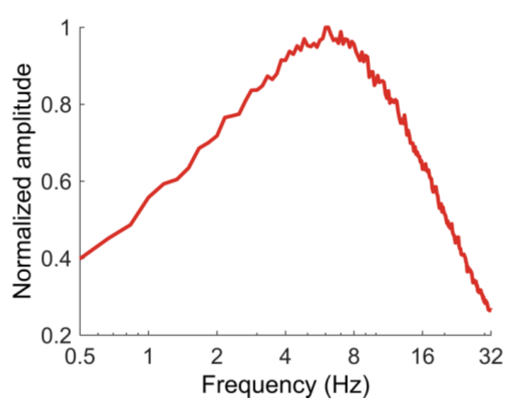

Figure 17 Mean envelope modulation spectrum of the audiobook used in the current study. This modulation spectrum was analyzed based on Ding et al., (2017), and it and shows a peak at ~6.3 Hz.

Preceding the audiobook excerpt, participants listened to 100 repetitions of a 1000 Hz pure tone and a 6 Hz sinusoidally amplitude modulated tone. Each tone had a duration of 1 second. Participants were instructed to listen to the tones without any additional tasks. The purpose of using pure tones was to examine Auditory Long Latency Responses (ALLRs), evoked potentials generated by the primary auditory cortex typically observed between ~50 to 300 ms after stimuli onset (Davis & Zerlin, 1966; Vaughan & Ritter, 1970). The modulated tone was presented to pseudorandomly to engage the participant's attention. These responses, and their P1, N1, P2, and N2 components, should be observed in response to any auditory stimuli including speech (Burkard, Eggermont, & Don, 2007) in individuals with intact auditory cortex. In the current study, ALLRs were used with two main purposes: i) to test the recording quality of the EEG system; and ii) to confirm that the EEG was capturing the expected brain responses to auditory stimuli. Therefore, ALLRs results will be reported, but not analyzed in relation to scripted-sentence learning.

5.2.2.2 EEG acquisition

The audiobook and ALLRs stimuli were presented binaurally using insert earphones (ER-3A; Etymotic Research) at 84 dB SPL, following similar amplification used previously in aphasia (e.g., Becker & Reinvang, 2007; Robson et al., 2014). Participants were instructed to listen attentively to the story, and to the best of their ability answer the yes/no questions presented after each ~1-minute story segment. They were also asked to refrain from extraneous movements and look at a fixation point marked on the computer screen. They were told that the total duration of the audiobook would be around nine minutes. Electrophysiological responses to ALLRs and continuous speech were recorded at a neurologic institute in Medellin, Colombia using a Nihon Kohden 32-channel clinical EEG system. Electrodes were placed on the scalp following the International 10-20 system montage (Klem, Lüders, Jasper, & Elger, 1999). Electrode impedance was less than 20 kΩ for all channels, and the recording sampling rate was 5000 Hz. As the EEG system was designed for recording auditory evoked potentials time-locked to the stimuli, a custom setup was used to achieve time-alignment. The analog electrical audio signal from one of the outputs of the audio-interface was directly plugged into the EEG amplifier with appropriate attenuation. Thus, the audio-signal that was played simultaneously recorded the participant's EEG and was used to get the timestamps of stimulus onsets and offsets.

5.2.2.3 Raw EEG preprocessing

The EEG data was preprocessed using EEGLAB v2019.1 (Delorme & Makeig, 2004) in MATLAB R2019a (MathWorks Inc). The data was preprocessed following Makoto's pipeline (Swartz Center for Computational Neuroscience). The first step in the pipeline was *resampling* down to 128 Hz to reduce the data size and to improve computational efficiency. The second step was *filtering* to reduce noise and to improve data quality (de Cheveigné & Nelken, 2019). Minimum-phase non-causal Sinc FIR filters were used with a high pass cut-off frequency of 1 Hz, filter order of 846; and a low-pass cut-off frequency of 15 Hz, filter order of 212. Third, the filtered EEG data was *re-referenced* to the average of the two mastoid channels (Di Liberto et al., 2015; O'Sullivan et al., 2017). Fourth, large movement artifacts were suppressed using Artifact Subspace Reconstruction (ASR; Mullen et al., 2013; Plechawska-Wojcik, Kaczorowska, & Zapala, 2019). For this step, clean sections of the data were identified with an approximate duration of 1 minute. These sections were then used as calibration data for ASR. After cleaning the data with ASR, it was separated into epochs from -5 seconds to 70 seconds (onset of each track), resulting in a total of 9 epochs. Lastly, Independent Component Analysis (Jung et al., 2001) was performed on the epoched data with the purpose of removing eye-movement and muscle artifacts. The independent components were then inspected visually for time-course, topography, and spectrum and were only removed if they corresponded to ocular and muscular activity. To extract the delta, theta, and alpha bands from the EEG, the data was filtered using windowed sinc filters (Di Liberto et al., 2015) with the following parameters: delta (lf = 1 Hz, order = 846; hf = 4 Hz, order = 424), theta (lf = 4 Hz, order = 424; hf = 8 Hz, order = 424), and alpha (lf = 8 Hz, order = 424, hf = 15 Hz, order = 424).

5.2.2.4 Estimating cortical tracking of the speech envelope

Preprocessed EEG in response to the audiobook was analyzed using the multivariate linear regression (Di Liberto et al., 2015), with the multivariate Temporal Response Function (mTRF) MATLAB toolbox (Crosse, Di Liberto, Bednar, & Lalor, 2016). The mTRF was used to estimate how well the EEG activity was predicted by speech envelope properties, providing an individual measure of cortical tracking (see Figure 18 for a schematic representation of the mTRF). *First*, the Hilbert transform was used to estimate temporal envelopes from the waveform signal (i.e., from the audiobook). Specifically, the multiband speech envelope was analyzed using Hilbert decomposition of the output of sixteen gamma tone filters equally spaced on the equivalent rectangular bandwidth scale (Slaney, 1998). To match the sampling rate of the EEG, the multiband envelopes were also down-sampled to 128 Hz.

Second, the mTRF was used to predict EEG envelopes via ridge regressions, using the preprocessed EEG and the temporal envelopes obtained from the Hilbert decomposition in the first step. To reduce overfitting, a regularization parameter during model estimation was applied. This regularization parameter was optimized from 2^0 to 2^{20} using cross validation in order to obtain the best model fit across electrodes and participants (Crosse et al., 2016). Model fit was evaluated via Pearson's correlation between the EEG predicted data and the observed EEG. Lastly, these correlation values were taken as an estimate of the participants' degree of cortical tracking, because they reflect how well each participant's neural activity was explained by the rhythmic features of the speech envelope.

To ensure that cortical tracking corresponds to the coupling between neural oscillations and a given speech envelope, the observed cortical tracking estimates were compared against randomly generated cortical tracking values. These random values were obtained from one hundred mismatched permutations of EEG responses and the envelopes of the audiobook. In other words, random cortical responses from the same participant are paired with different envelopes multiple times to obtain a randomly generated cortical tracking. Comparing the observed cortical tracking against the random cortical tracking provides evidence regarding the association between neural oscillations as a result of a *specific* auditory input as presented by the speech envelope rather than a chance response.

The correlation between cortical tracking estimates and individual learning estimates (rhythmic-enhanced and control, derived from participant-level random effects in the short-term learning paradigm (see analysis section 5.1.1.1 above) was assessed. In addition, this study also examined the correlation between cortical tracking estimates and the language, attentional, and rhythmic processing measures to trace the relationship between cortical tracking and higher cognitive-linguistic factors. Lastly, a correlation between comprehension questions performance and cortical tracking was used to explore the association between comprehension accuracy and cortical tracking.

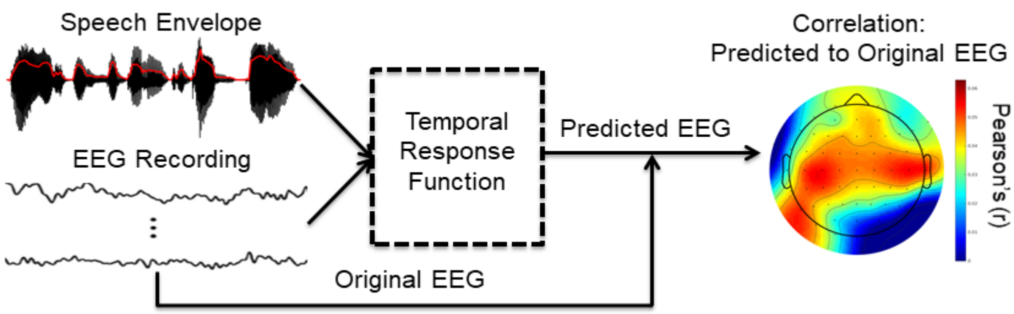

Figure 18 Schematic representation of the mTRF model.

Unpublished figure, reproduced with permision from the Sound Brain Lab at the University of Pittsburgh.

5.2.2.5 Potential outcomes of cortical tracking in aphasia

Although examining cortical tracking in aphasia, and its association with scripted-sentence learning, has an exploratory nature, some potential outcomes can be expected. It could be the case that PWA show cortical tracking responses similar to the ones reported in older adults (e.g., Brodbeck et al., 2018): enhanced cortical tracking without the benefits of improved performance. For example, PWA may increase their listening effort (see Peelle, 2018 for neurotypical responses) in order to compensate for their language-impairments by recruiting additional cortical resources (see Peelle, 2018; Reuter-Lorenz & Cappell, 2008; Wingfield & Grossman, 2006 for neurotypical responses). Consistent with the aging literature, despite an enhanced cortical activity, PWA with stronger cortical tracking measures may have lower learning estimates due to an inefficient processing of both speech rhythmic features and higher-order linguistic features.

Conversely, PWA with stronger cortical tracking measures may have higher learning estimates in both the rhythmic-enhanced and control conditions, indicating that an efficient processing of both rhythmic speech features and higher-order linguistic features facilitate sentence learning (Peelle & Davis, 2012; Peelle et al., 2013). Regardless of the potential outcome, the findings in sections 5.1.2.2 and 5.1.3.1 show that scripted-sentence learning estimates are influenced by the presence of aphasia (e.g., aphasia severity and overall language-impairment measures); therefore, it is important to examine the association between cortical tracking and overall measures of aphasia severity in order to reveal the full set of relationships between cortical tracking, language impairments, and scripted-sentence learning estimates. Moreover, it is unknown whether there is an association between cortical tracking and attention and rhythmic processing performance. Thus, the current findings are informative regardless of whether a relationship between cortical tracking response and scripted-sentence learning is found.

5.2.3 Results: cortical tracking associated with scripted-sentence learning response

Auditory Long Latency Responses (ALLRs) were used to confirm adequate EEG recording quality by capturing auditory cortex responses that are expected to *any* auditory stimuli. As anticipated, ALLRs and their P1, N1, P2, N2 components were present in all participants. See Figure 19 for a depiction of ALLRs per participant and their corresponding N1 scalp topography. These responses provide evidence that the subsequent cortical tracking measure captures brain responses associated with the auditory stimuli (i.e., the speech envelope of each audiobook segment).

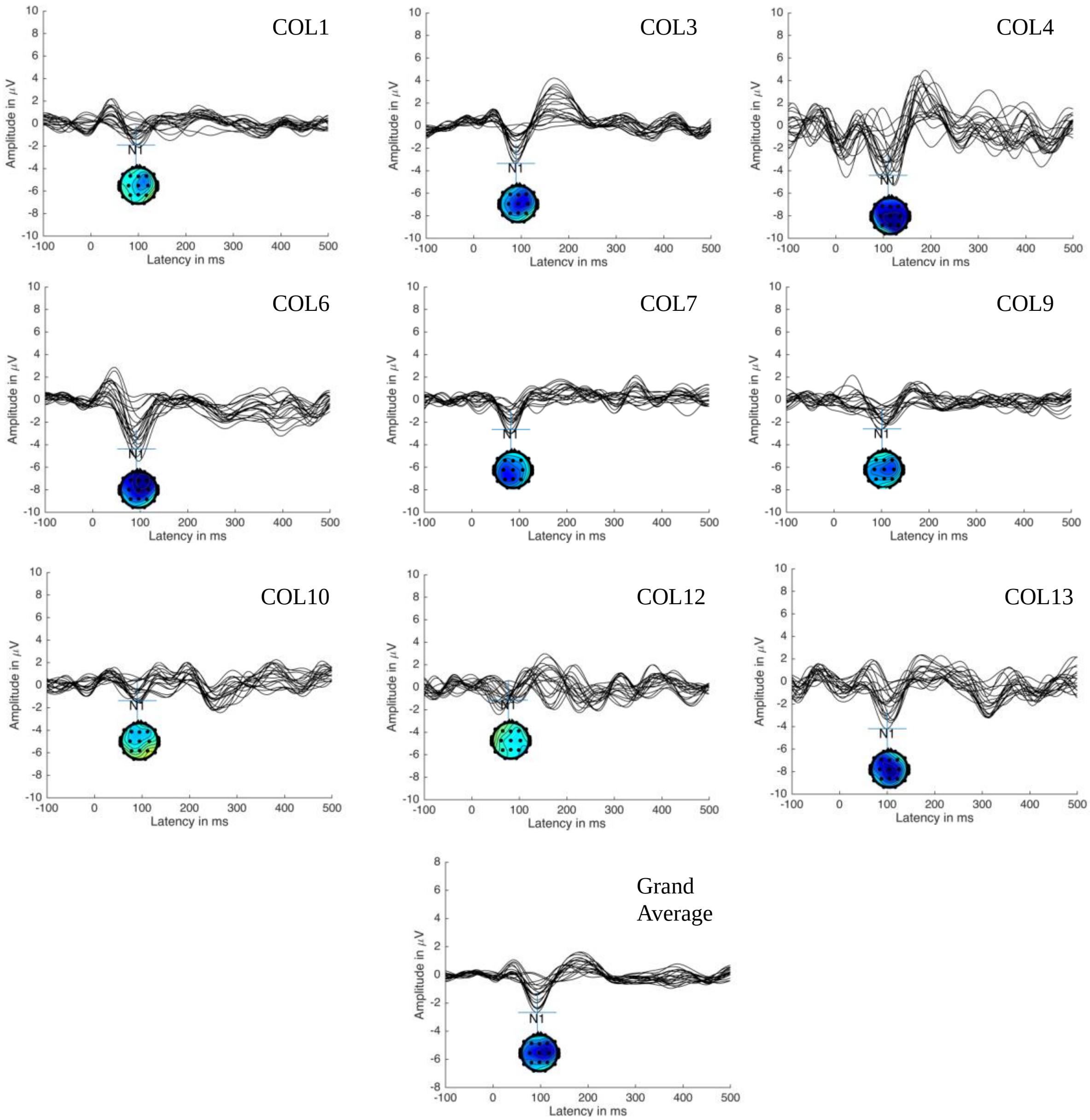

Figure 19 Auditory Long Latency Responses for each participant.

The x-axis represents latency in milliseconds; the y-axis represents amplitude in microvolts.

Each figure includes the N1 component and its related topography. The bottom figure represents the grand average of ALLRs across participants.

Estimates of cortical tracking in the delta (< 4 Hz), theta (~4–8 Hz), alpha (~8–13 Hz) bands were obtained per participant using the mTRF. The strength of these cortical tracking values is consistent with previous reports (e.g., Di Liberto et al., 2015). These estimates were compared to randomly generated cortical tracking measures using a Wilcoxon test. The observed cortical tracking measures were significantly higher than random cortical tracking in all bands ($z = -2.88$, $p = 0.004$), indicating an association between neural oscillations and a *specific* auditory input (i.e., speech envelope) rather than a random response (see Figure 20 for a depiction of observed versus random cortical tracking measures per participant).

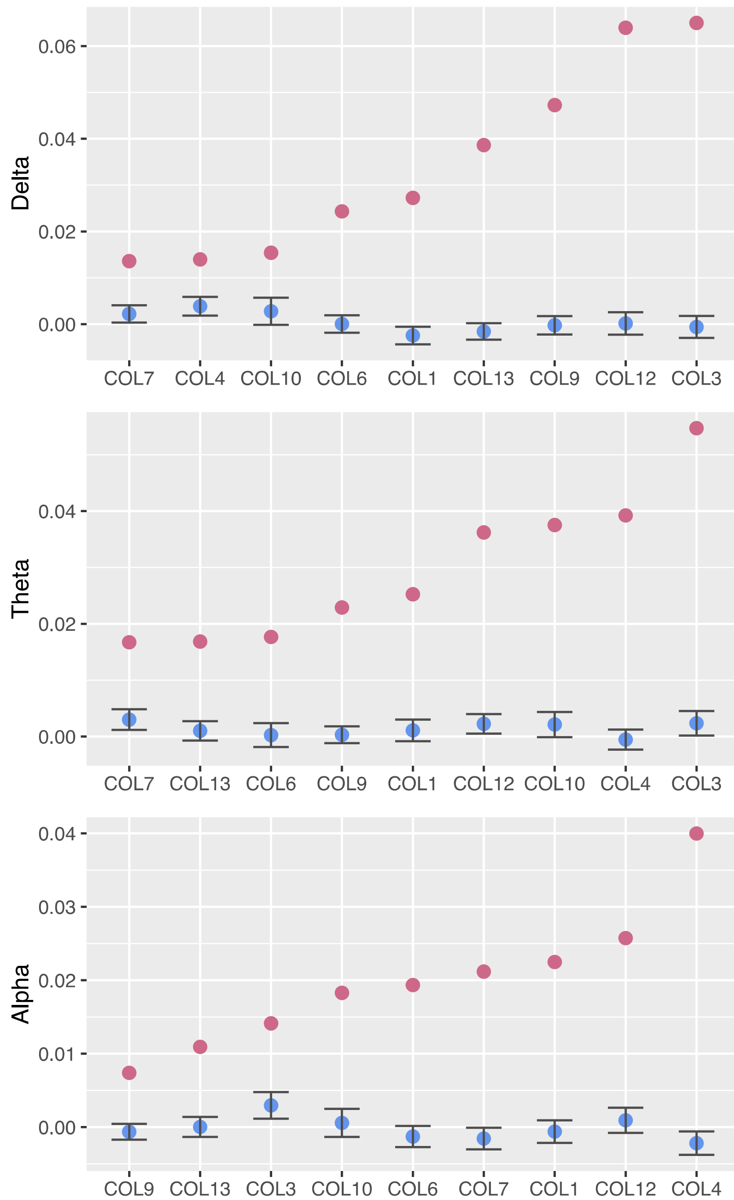

Figure 20 Observed (light pink) vs. Random (light blue) Cortical tracking measures.

The x-axis is organized from lower to higher observed cortical trakcing per participant; the y-axis represents cortical tracking measures as obtained from the mTRF. Error bars represent 95% confidence intervals of chance distribution.

Correlations were used to explore the association between individual learning estimates (rhythmic-enhanced and control) and cortical tracking in the delta, theta, and alpha bands. None of the EEG bands correlated significantly with the learning estimates (either before or after Bonferroni correction), indicating a lack of statistically robust associations between cortical tracking and scripted-sentence learning. Nonetheless, the correlation coefficients in the theta band had absolute values of 0.68 (control learning estimate) and 0.62 (rhythm learning estimate), and the scatterplot shows a linear trend that associates cortical tracking in this band and learning estimates (see Figure 21). As depicted, this association would indicate that PWA with higher cortical tracking also had lower learning estimates for scripted sentences. This preliminary finding will be discussed below in its association with language impairment measures. Further investigation is required to determine if these cortical tracking estimates are consistent with the recently published evidence of an increased cortical tracking in older adults without the benefits of improving auditory processing (e.g., Brodbeck et al., 2018; Henry et al., 2017; Herrmann et al., 2019).

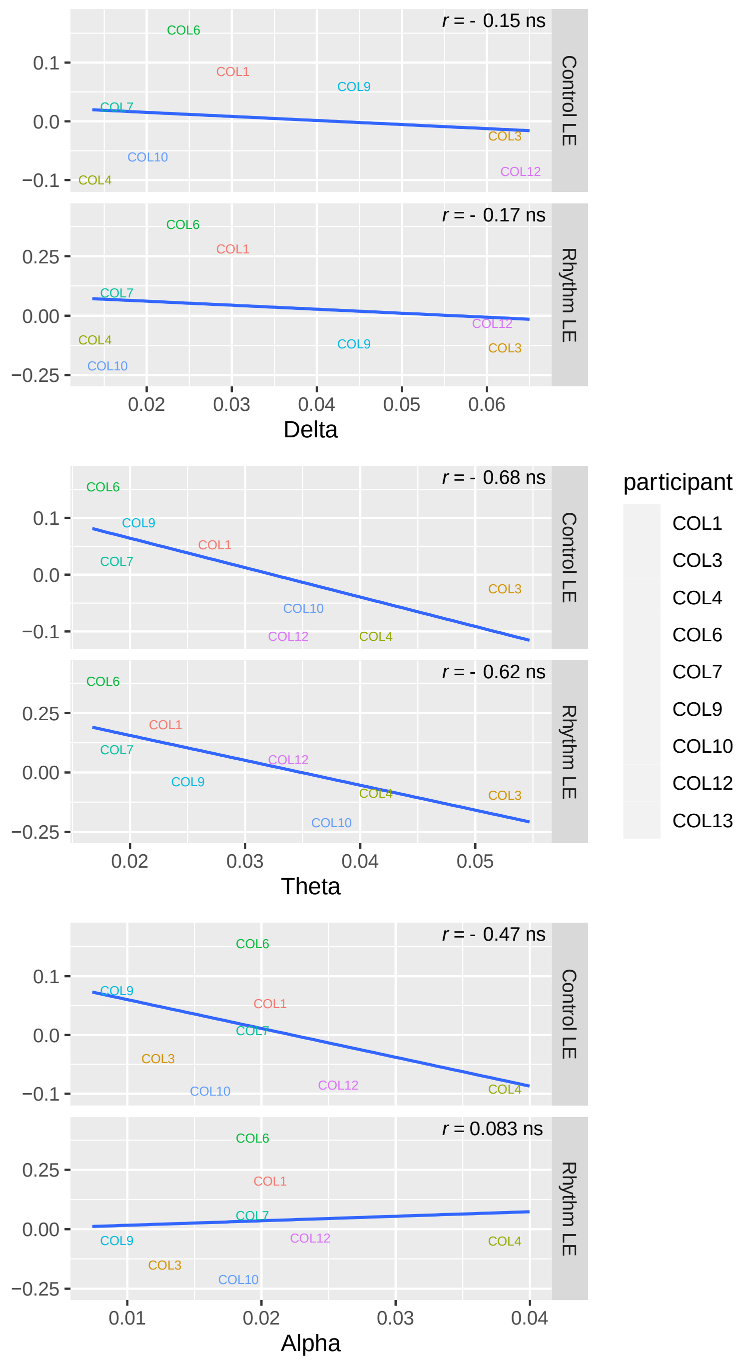

Figure 21 Correlations between cortical tracking in the different EEG bands and learning estimates. The x-axis corresponds to observed cortical trakcing values as obtained from the mTRF; the y-axis represents learning estimates derived from the random effects of the group-level analyses; ns = not significant either before or after Bonferroni correction.

Correlations were used to examine the association between cortical tracking estimates and the language, attentional, and rhythmic processing measures (see Table 7). Three significant correlations before Bonferroni correction were found: i) between the theta band and the Auditory Verbal Comprehension score from the WAB, indicating that PWA with higher cortical tracking also had better language comprehension abilities; ii) between the theta band and the Naming and Word Finding score from the WAB, indicating that PWA with higher cortical tracking also had better naming abilities; and iii) between the delta band and the Repetition score from the WAB, indicating that PWA with lower cortical tracking also had higher repetition abilities (perhaps due to the use of a nonlexical repetition route; Hanley, Dell, Kay, & Baron, 2004). None of these correlations reached significance after Bonferroni corrections.

Table 7 Correlations between cortical tracking and language, attention, and rhythmic processing measures

Measure \ Band	alpha	delta	theta
WAB Comprehension	0.22	0.10	0.68*
WAB Naming	0.50	-0.28	0.69*
WAB Repetition	0.42	-0.74*	-0.19
WAB spontaneous	0.50	-0.20	0.58
WAB-AQ	0.56	-0.34	0.62
TEA Elevator	0.07	0.37	0.01
TEA Distraction	0.55	-0.04	0.14
MBEA	0.33	-0.37	-0.13
BAT perception	-0.54	0.37	-0.04
BAT synchronization	0.21	-0.51	0.26

WAB = Western Aphasia Battery; WAB Comprehension = Auditory Verbal Comprehension Score; WAB Naming = Naming and Word Finding Score; WAB Repetition = Repetition Score; WAB spontaneous= Spontaneous Speech Score; WAB-AQ = WAB Aphasia Quotient; TEA Elevator = elevator counting subtest, Test of Everyday Attention; TEA Distraction = age scaled-score for the elevator counting with distraction subtest; MBEA = rhythmic task of the Montreal Battery of Evaluation of Amusia; BAT perception = Perceptual judgment of the beat ability estimate from the CA-BAT; BAT synchronization = correlation between BAT inter-tap mean and target inter-tap intervals adjusting for tactus; * = significant before Bonferroni correction. After Bonferroni correction none of these correlations reached significance.

Regarding comprehension accuracy, participants listened to the audiobook attentively and responded nine yes/no questions provided after each ~1-minute segment. Accuracy in answering these questions was low (group mean = 53.09%, SD = 19.86). As Figure 22 depicts, all participants, except for COL10 had accuracy scores around or below 50%. There were no significant correlations between performance in answering these yes/no questions with cortical tacking in any of the bands.

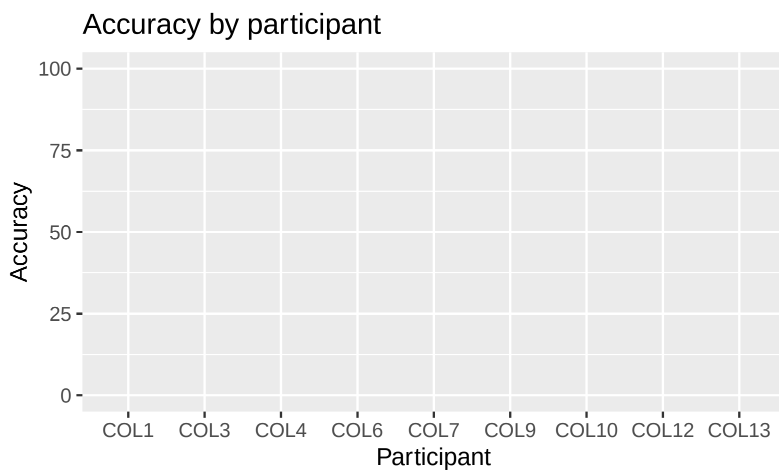

Figure 22 Question accuracy by participant.

y-axis reflects percentage correct. x-axis represents each participant.

5.2.4 Discussion: cortical tracking associated with scripted-sentence learning response

The aim of identifying individual characteristics in cortical tracking associated with scripted-sentence learning in aphasia was *exploratory*. The rationale for this exploration was based on the claim that the rhythmic features of speech entrainment facilitate scripted-sentence learning (Fridriksson et al., 2015; Fridriksson et al., 2012). The findings in section 4.2.3, demonstrating that rhythmic-enhanced speech results in better scripted-sentence learning, are consistent with this claim. Consequently, a better perception and processing of speech-rhythmic features (as indexed by cortical tracking) could be related to the ability to do speech entrainment (i.e., produce speech simultaneously with another speaker) and therefore to learn scripted sentences. Still, this study is the first report of cortical tracking in stroke survivors with aphasia, which underlines its exploratory nature but also highlights the relevance of the findings beyond a direct association between cortical tracking and scripted-sentence learning.

Auditory Long Latency Responses (ALLRs) confirmed an adequate response from the auditory cortex to both pure tones (1000 Hz) and sinusoidally amplitude-modulated tones (6 Hz). Also, ALLRs showed the expected P1, N1, P2, and N2 components (Burkard et al., 2007). These findings are evidence of a suitable EEG recording system in Colombia and of an adequate capturing of brain responses to auditory stimuli (Davis & Zerlin, 1966; Vaughan & Ritter, 1970).

Cortical tracking estimates were obtained from the mTRF (Crosse et al., 2016; Di Liberto et al., 2015) for the delta (< 4 Hz), theta (~4–8 Hz), and alpha (~8–13 Hz) bands. The magnitude of cortical tracking estimates are consistent with previous reports (e.g., Di Liberto et al., 2015), and were found to be significantly different from randomly generated cortical tracking measures. This finding indicates an association between PWA's neural oscillations and the speech envelope of the audiobook used, rather than a random brain oscillatory response.

Cortical tracking did not correlate significantly with the scripted-sentence learning estimates (rhythmic-enhanced and control). However, the correlation coefficients for the theta band and learning estimates were relatively high (control, $r = -0.68$; rhythmic-enhanced, $r = -0.62$) and the scatterplot indicated a clear linear trend (Figure 21), both of which encourage future investigation. This trend suggested that PWA with higher cortical tracking in the theta band also had lower learning estimates. Before aiming to interpret this preliminary finding, it is important to highlight that language ability (as measured by WAB comprehension and naming) was positively associated with theta band cortical tracking ($r = 0.68$ and $r = 0.69$, respectively), indicating that milder PWA had higher perception of the rhythmic properties of the speech signal.

This preliminary finding is consistent with the association between language capacities and scripted-sentence learning estimates (see sections 5.1.2.2 and 5.1.3.1) such that PWA with higher language ability showed lower learning estimates. In this case, the evidence also suggests that PWA with higher cortical tracking also showed evidence of lower learning estimates. A cautious reading of these preliminary findings is that PWA with more preserved language ability also show higher cortical tracking in the theta band, suggesting that at an acoustic level, they may be segmenting syllabic information appropriately (Teng et al., 2018). Consistent with this cautious reading, people with more preserved language ability did not show a special benefit in learning scripted sentences given that they started at a higher learning intercept (less opportunity for improvement), and also because they experienced more “interference” from their propositional language when producing the trained stimuli (rehearsed language).

As discussed in section 5.1.3.1, PWA with more severe language impairments may benefit more from an errorless approach (Fillingham et al., 2005a, 2005b, 2006) as they learn and retrieve scripted sentences as rehearsed or formulaic language (Wray & Perkins, 2000) without “interference” from their highly impaired propositional language. On the contrary, PWA with lower language impairments may experience "interference" from their preserved propositional language, leading to more word diversity in their production. This diversity would result in larger deviations from the target scripted sentences. In other words, PWA with better language ability seemed to have higher cortical tracking and also lower learning estimates, due to the special benefit that speech entrainment, and this type of sentences, have for more severely impaired PWA. A three-way interaction model that includes language impairment measures, cortical tracking, and learning estimates would be the optimal analytical approach to examine these variables' association. Although the current study did not have sufficient power to implement such analysis, these preliminary findings encourage future exploration of these interactions in larger samples.

Participants in the current sample may have displayed increased cortical tracking activity, but the lack of a control group of neurologically healthy older adults does not allow a direct comparison in these two populations. Recent studies have reported enhanced cortical tracking in older adults, without the added benefits of improved auditory processing (e.g., Brodbeck et al., 2018; Henry et al., 2017; Herrmann et al., 2019), and it is likely that some of this age-related enhancement may be present in PWA. Also, acoustically challenging circumstances (e.g., diminished hearing acuity, background noise) cause listeners to increase their listening effort to meet the auditory processing demands (Peelle, 2018). Thus, it would be expected that due to their language impairments, PWA show an increase in listening effort in order to process auditory information successfully (Richie, 2015). This increased effort could be expected to be similar or higher than what is reported in older adults (Peelle, 2018). In short, the current study does not allow a direct examination of an enhancement in cortical tracking in PWA, but some degree of enhancement could be expected due to age-related processing patterns and listening effort; future studies could compare cortical tracking in neurologically healthy older adults and PWA, and also examine its association with linguistic and non-linguistic learning.

Lastly, the WAB repetition score and cortical tracking in the delta band were negatively correlated (r = -0.74). This correlation suggests that PWA with lower cortical tracking in the delta band also had higher repetition abilities. One preliminary interpretation of this finding is that PWA in the current sample may have been engaging a nonlexical repetition route to do the WAB repetition task (Hanley et al., 2004). Previous findings associate the delta band with higher-order language processing (Molinaro & Lizarazu, 2018), including facilitating speech comprehension in noise (Etard & Reichenbach, 2019). Successful speech comprehension reflects successful access to the lexical-semantic system, meaning that higher cortical tracking in the delta band may index more successful access to the semantic system. Therefore, a preliminary interpretation of the current findings is that lower cortical tracking in the delta band could be associated with repetition abilities that are unrelated to the semantic system. In the absence of access to the semantic system, it is likely that PWA have repeated the words and sentences in the WAB repetition task using a nonlexical route. This interpretation is preliminary, but it would explain the observed negative relationship between delta band activity and repetition performance. Nonetheless, a recent report in older adults suggested that listeners with lower cortical tracking in the delta band had higher accuracy in comprehending yes/no questions about an audiobook (McHaney et al., 2020). Thus, it is necessary to examine the association between lexical/nonlexical routes of repetition and cortical tracking in the delta band in a larger sample of PWA. This examination is required before reaching any conclusions about the association between cortical tracking in this band and repetition abilities.

A relevant future direction that expands the current study's purpose is to compare cortical tracking to a narrative *without* and *with* rhythmic enhancements (stress-aligned and metronomic beats) as employed in the current study. As seen in the behavioral results, PWA learned significantly more scripted sentences in the rhythmic-enhanced conditions when compared to the control condition. Comparing cortical tracking to these different stimuli, and investigating its association with learning in aphasia, can help to determine whether rhythmic enhancements can improve PWA's perception and processing of speech-rhythmic features that result in behavioral learning gains.

In summary, ALLRs confirmed both a suitable EEG recording and an adequate response from the auditory cortex in the current sample of participants. Cortical tracking of the speech envelope was obtained using the mTRF (Crosse et al., 2016; Di Liberto et al., 2015), and responses in each band (delta, theta, and alpha) were significantly different from random brain oscillatory responses. There were no statistically robust correlations between cortical tracking in either band and the scripted-sentence learning estimates; however, the association between theta and learning estimates had high correlation coefficients, and its scatterplot depicts a clear negative linear trend.

This preliminary finding suggests that PWA with higher cortical tracking also showed lower learning estimates (poorer perception of rhythmic properties in the speech signal), which is consistent with the earlier finding that PWA with higher language ability showed lower learning estimates. It is plausible that better language ability, higher cortical tracking, and lower learning estimates reflect that milder PWA started at a higher learning intercept giving them less improvement space. It is also likely that errorless learning and formulaic language (i.e., scripted sentences) confer a special benefit to people with more severe aphasia. A statistical model with a three-way interaction between language impairment measures, cortical tracking, and learning estimates will be an appropriate analytical approach for future studies with larger samples. Lastly, the negative correlation between the delta band and repetition abilities may be related to using the nonlexical route for repetition (Molinaro & Lizarazu, 2018). Overall, cortical tracking may index PWA's perception and processing of speech-rhythmic features that may support speech entrainment, and consequently, scripted-sentence learning.

6.0 General Discussion

Aphasia rehabilitation depends on evidence-based treatments, but it is unclear how some of these treatments work (i.e., their mechanisms of action) and for which PWA they work best (e.g., individual characteristics associated with treatment response). Understanding the "how" and "for whom" of evidence-based treatments will improve both aphasia rehabilitation and treatment personalization. Using an adaptation of script training for Spanish speakers with aphasia, the current study investigated mechanisms and individual characteristics associated with scripted sentence learning.

Adapting script training to Spanish speakers with aphasia is important because although there are ~450 million native Spanish speakers globally (compared to ~360 million English native speakers), over 85% of published aphasia treatment research has focused on English (Beveridge & Bak, 2011). Therefore, the current adaptation contributes to aphasia rehabilitation knowledge for this large portion of the world population. Specifically, the evidence reported in this study has increased current knowledge in three domains: i) demonstrating that script training can be adapted for Spanish-speaking PWA; ii) providing evidence that identifies speech entrainment as a key mechanism for script training; and iii) providing preliminary evidence identifying some individual characteristics (behavioral and cortical tracking) associated with scripted-sentence learning.

Script training is an evidence-based treatment approach for PWA that trains the production of personally relevant sentences (Hubbard, Nelson, & Richardson, 2020). Using this treatment method, PWA learn to produce functional sentences to communicate in daily life situations. In contrast to many aphasia treatments that focus on the retrieval and production of single words (e.g., Kendall et al., 2008; Martin, Fink, Renvall, & Laine, 2006; Wisenburn & Mahoney, 2009), script training focuses on scripted sentences as a minimal learning unit. Although several studies have examined this intervention in PWA (e.g., Cherney et al., 2008; Cherney, Halper, et al., 2011; Cherney et al., 2015; Cherney et al., 2014; Goldberg et al., 2012; Holland, Halper, & Cherney, 2010; Lee et al., 2009), its mechanisms of action are not well understood, and it remains unclear what characteristics of PWA are associated with better treatment response.

To examine *script training mechanisms of action*, we first considered its typical learning procedure, which includes listening, repetition, speech entrainment, and independent production (e.g., Cherney et al., 2014; Kaye & Cherney, 2016). The listening, repetition, and independent production steps are commonly investigated in multiple aphasia treatments (e.g., Fillingham et al., 2006; Nozari et al., 2010), but less is known about speech entrainment and its role in scripted-sentence learning (e.g., Fridriksson et al., 2015; Fridriksson et al., 2012). Entrainment relies on the detection, integration, and production of rhythmic features (Phillips-Silver et al., 2010); in other words, rhythmic features are the basis over which speech entrainment takes place, but it is unclear exactly how these features can facilitate learning of scripted sentences. Therefore, we examined the rhythmic features of speech entrainment as a potential mechanism of action for scripted-sentence learning.

It was hypothesized that i) unison production of scripted sentences enabled PWA to align to word stress, supporting lexical retrieval (e.g., Soto-Faraco et al., 2001), and ii) unison production of scripted-sentences enhanced sentence memorization via chunking (e.g., Purnell-Webb & Speelman, 2008; Stahl et al., 2011). To test these hypotheses, we assessed the effects of speech entrainment to two enhanced rhythmic features in speech, stress-aligned (which should support lexical retrieval) vs. metronomic beats (which should enhance memorization via chunking), and compared them to a control condition without rhythmic enhancements. We also compared the two types of rhythmic cues (supporting lexical access vs. supporting memorization via chunking) against each other to test for potential differential contributions to scripted-sentence learning.

Analyses showed a significant interaction effect between session and a contrast that compared the rhythmic-enhanced conditions (stress-aligned and metronomic) against the control condition. This interaction effect indicated that as the number of sessions increased, the difference between the rhythmic-enhanced and control conditions also increased. As entrainment depends on the detection, integration, and production of rhythmic features (Phillips-Silver et al., 2010). This finding provides evidence supporting speech entrainment is a key mechanism to facilitate scripted sentence learning (Stahl et al., 2013; Stahl et al., 2011). However, it is important to note that this finding does not imply that the speech entrainment (and its underlying rhythmic structure) is the only mechanism for scripted-sentence learning, but rather that the rhythmic features examined in this study are at least one key active ingredient.

There were no significant interaction effects between session and the contrast that compared the two rhythmic conditions against each other (stress-aligned vs. metronomic). This finding suggests that both rhythmic manipulations, one that aims to enhance lexical access (e.g., Cutler, 2005; Soto-Faraco et al., 2001) and one that aims to improve memorization via chunking (e.g., Purnell-Webb & Speelman, 2008; Stahl et al., 2011) may facilitate scripted-sentence learning, but their differential contributions could not be established. In other words, from the current data, there is no evidence that one is more beneficial than the other. Another potential interpretation of this finding is that, given that entrainment depends on the integration of rhythmic features (Phillips-Silver et al., 2010), *any* rhythmic enhancement to speech entrainment could facilitate scripted sentence learning. This account would have to be directly tested using additional variations of rhythmic cues not directly tied to the learning mechanisms that were the focus in this study. Lastly, the lack of differences between the rhythmic conditions might have been due to some PWA benefiting more from one rhythmic enhancement than the other, but in the current study none of the participants showed an unequivocal response to one of the rhythmic-enhanced conditions than the other.

Importantly, collapsing across conditions (stress-aligned, metronomic, and control), the odds of producing a correct syllable were significantly higher with each additional training session. This result has multiple implications: it indicates successful scripted-sentence learning in the sample of this study, it provides further evidence of script training as an efficacious treatment for aphasia (e.g., Cherney, Halper, et al., 2011; Cherney et al., 2014; Lee et al., 2009), it encourages the use of speech entrainment and its underlying rhythmic features for scripted-sentence learning (e.g., Fridriksson et al., 2015; Fridriksson et al., 2012), and it demonstrates cross-linguistic script training benefits for Spanish speakers with aphasia.

Although the current evidence supports the cross-linguistic benefits of script training, it is important to mention that the rhythmic-enhanced manipulations could potentially have different results in speakers with aphasia in other languages, such as English. In terms of the *word-stress condition*, there are a number of stress assignment differences between the English and Spanish languages that suggest that this condition could be particularly facilitative for Spanish speakers compared to English speakers. English speakers use more segmental information for word recognition than suprasegmental information (Fear et al., 1995). In contrast, Spanish speakers directly use suprasegmental information for lexical retrieval (Soto-Faraco et al. 2001), particularly because unstressed syllables are never reduced in this language (Cutler, 2012). For example, pairs of words in Spanish can be segmentally identical but suprasegmentally different, constraining speakers of this language to use word stress information to distinguish between words and access lexical information (Baqué, 2017). Given that word stress plays a more critical role in accessing lexical information in the Spanish language, this condition could be more helpful for scripted-sentence learning in PWA that speak this language and relatively less useful for languages like English, where segmental information rather than word stress is primarily used for word recognition (Fear et al., 1995). Nevertheless, this hypothesis needs to be directly tested in future studies.

Moreover, the *metronomic beat condition* could bring a greater benefit for stress-timed languages, such as English, when compared to the stress-aligned condition used in this study. This benefit is supported by evidence from the Speech Cycling Paradigm (Cummins & Port, 1998), where a tendency to align phrases with a provided rhythmic structure has been reliably observed. This tendency provides an inherent rhythmic structure over which information can be grouped/chunked and then recalled (e.g., Purnell-Webb & Speelman, 2008). An enhancement in memorization via chunking provided by metronomic beats would therefore apply to a great variety of languages. In fact, evidence from the Speech Cycling Paradigm has come from other stress-timed languages (e.g., Arabic) and even mora-timed languages (e.g., Japanese) (Tajima et al., 2000), which supports cross-linguistic observations of the tendency to align linguistic information with rhythmic structures. Nonetheless, evidence supporting the use of metronomic beats for scripted-sentence learning in aphasia has thus far come from German (Stahl et al., 2013; Stahl et al., 2011) and now from Spanish; therefore, it needs to be replicated and expanded in other languages, including English.

Regarding *individual characteristics associated with scripted-sentence learning*, the current study examined behavioral (language, attention, and rhythmic processing performance) and neuro-electrophysiological (cortical tracking) measures associated with learning estimates. In terms of *language performance*, PWA were given the Western Aphasia Battery (WAB; Kertesz, 1982) in its Spanish adaptation (Kertesz & Pascual-Leone, 2000). Based on previous reports from word-finding treatments for aphasia (e.g., Liang et al., 2001; Quique et al., 2019), it was hypothesized that WAB-AQ would be positively associated with learning performance, such that higher scores (milder aphasia severity) would be associated with better scripted-sentence learning. Contrary to this hypothesis, it was found that WAB scores, except for the WAB repetition subscale, correlated negatively with the learning estimates. This correlation was statistically significant for WAB-AQ and the control-condition learning estimates after Bonferroni correction. Thus, this study found that PWA with more severe language impairments (i.e., lower WAB scores) also had higher scripted-sentence learning estimates, which is consistent with at least one other study in script training, Lee et al. (2009), who also found negative correlations between WAB scores and production of script-related words.

A potential explanation for a negative correlation between WAB scores and scripted-sentence learning estimates can be found in errorless learning (e.g., Fillingham et al., 2005a, 2005b; Fillingham et al., 2006) and in the specific tasks that script training aims to achieve, rehearsed language production (e.g., Stahl, 2013; Van Lancker Sidtis, 2012). Specifically, speech entrainment provides strong support for language production (minimizing error) while having low language processing demands (minimizing effort). This support offered by speech entrainment likely minimized both error and effort making it especially beneficial for more severely impaired PWA. Similarly, script training relies on overtraining the memory encoding system so that scripted sentences are stored and then retrieved automatically as formulaic language (i.e., rehearsed language), rather than being generated at the moment, as propositional language (see Logan, 1988; Wray & Perkins, 2000). It is likely that PWA that were more severely impaired learned scripted sentences as formulaic language given that their propositional language had higher levels of impairment; conversely, PWA with less language impairments would have more "interference" from propositional language, which led to the production of words that were related, but that deviated from the target scripted sentence.

Notably, PWA with more severe language impairments also showed lower performance at the beginning of the short-term learning paradigm. In other words, more severe PWA started with lower performance in scripted-sentence learning, which gave them more opportunity to improve. Consistent with this finding, PWA with more severe impairments did not have interference from their highly impaired propositional language when retrieving the target sentences during probes, which resulted in higher learning of scripted-sentences. Overall, in the current study PWA with more severe language impairments also showed better scripted-sentence learning estimates. It is important to note that these results are preliminary, and that future research should aim to examine whether using rhythm for scripted-sentence learning is beneficial for PWA regardless of their level of language impairment.

Regarding *attentional performance*, PWA were assessed with the Elevator Counting and Elevator Counting with Distraction subtests of the Test of Everyday Attention (TEA; Robertson et al., 1994). It was expected that the rhythmic enhancements would direct attentional resources towards the stressed syllables and metronomic beats (Large & Jones, 1999; Pitt & Samuel, 1990), helping to overcome the concomitant attentional deficits that negatively affect PWA's language production (Murray, 1999, 2012); therefore, it was hypothesized that poor attention performance in the TEA would be negatively associated with learning performance in the rhythmic-enhanced conditions. However, none of the attention performance variables strongly correlated with either of the learning estimates (rhythmic-enhanced and control), and the scatterplots depicting these correlations did not reveal meaningful patterns. This lack of association between attentional performance and learning estimates could be due to poor performance and low variability in the attentional performance of the current sample of participants. Still, PWA performed strikingly similar to previous studies, consistent with evidence that shows attentional deficits in PWA (e.g., Murray, 1999; Murray, 2012). Another possibility for this lack of association is that rhythmic enhancements, as used in this study, do not provide additional attentional benefits to scripted sentence learning.

Regarding *rhythmic processing performance*, PWA were given the rhythm subtest of the Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al., 2003), the adaptive version of the perceptual judgment of the beat task (CA-BAT; Harrison & Müllensiefen, 2018b), and the synchronization with the beat of music task from the Beat Alignment Test (BAT; Iversen & Patel, 2008).

Participants showed poor performance in the rhythm and beat perception tasks. Specifically, participants showed low abilities in discriminating whether two given rhythms were the same or different (as measured by the MBEA), and also poor abilities in beat perception (as measured by the CA-BAT). The low performance in these perception tasks is consistent with previous results reported in stroke survivors with and without aphasia (e.g., Racette et al., 2006; Särkämö et al., 2009; Zipse et al., 2014). Conversely, in the rhythm production task, participants showed good synchronization with the beat of music (as measured by the BAT). It is not unusual to observe low performance in rhythm/beat perception tasks (e.g., MBEA, CA-BAT) while displaying higher abilities in rhythmic production tasks (e.g., Bégel et al., 2017; Sowiński & Dalla Bella, 2013). Perception and production rhythmic abilities seem to be partially independent (Sowiński & Dalla Bella, 2013), and it is expected that the general population has strong production rhythmic abilities given their early development in life (e.g., Phillips-Silver & Trainor, 2005; Zentner & Eerola, 2010) and their vast manifestation in the population (for reviews see Repp, 2010; Repp & Su, 2013).

It was hypothesized that people with better rhythmic processing performance would also have higher learning estimates in the rhythmic-enhanced conditions relative to the control condition. Previous studies reported observations in which PWA with higher rhythmic abilities responded better to rhythmic cues (e.g., hand tapping) in melodic intonation therapy (Curtis et al., 2020; Curtis & Zipse, 2017). Contrary to these predictions, most rhythmic processing measures did not show strong associations with scripted sentence learning estimates (rhythmic-enhanced and control). The only significant correlation was found between the CA-BAT and the learning estimates for the control items, indicating that PWA with better beat perception had higher learning estimates; however, this correlation did not reach significance after Bonferroni correction. Also, performance on this task was low and with low variability in this sample of PWA, and the scatterplot depicting this association does not show a clear linear pattern. Thus, this finding warrants further examination using larger samples. A cautious preliminary interpretation is that higher abilities in beat perception may be important when the rhythmic cues are absent (control condition) but are less critical when rhythmic support is provided, as in the rhythmic-enhanced conditions. This interaction will need to be directly tested in future studies.

Lastly, a measure of cortical tracking of the speech envelope (Ding & Simon, 2014) was obtained using the mTRF (Crosse et al., 2016; Di Liberto et al., 2015). The association between cortical tracking (in the delta, theta, and alpha bands) and scripted sentence learning estimates was examined using correlations. Similarly, the association between cortical tracking and language, attention, and rhythmic abilities were examined using correlations. There were no statistically robust correlations between cortical tracking and the scripted sentence learning estimates. However, the correlation coefficients for the theta band were high, and the scatterplot depicted a clear negative linear trend, suggesting that PWA with higher cortical tracking in this band showed lower learning estimates (poorer perception of rhythmic features present in the speech signal). Also, language ability (as measured by WAB comprehension and naming) was positively associated with the theta band cortical tracking.

These preliminary findings are consistent with the negative association between language capacity and scripted-sentence learning estimates described in sections 5.1.2.2 and 5.1.3.1. Recall that this association suggested that PWA with higher language ability showed lower learning estimates. That being the case, PWA with higher cortical tracking also showed evidence of lower learning estimates. A cautious reading of these findings is that PWA with more preserved language ability and higher cortical tracking did not show a benefit in learning scripted sentences because they started at a lower learning intercept, and given the learning approach (errorless and effortless) and the type of formulaic language used in this study (i.e., scripted sentences). A three-way interaction model on a future adequately power sample, including language impairment measures, cortical tracking, and learning estimates, would be an optimal approach to test this cautious reading of the current findings.

Cortical tracking in the delta band was strongly negatively correlated with the WAB repetition score, suggesting that PWA with lower cortical tracking had better repetition abilities. However, this association was not statistically robust after a Bonferroni correction. Previous findings have associated the delta band with higher-order language processing (Molinaro & Lizarazu, 2018), including facilitating speech comprehension in noise (Etard & Reichenbach, 2019). Although the specific association between cortical tracking in this band and speech comprehension is still an area of active study (e.g., Weissbart et al., 2019), it seems that it has an association with top-down processing and therefore with access to the lexical-semantic system. Given this relationship, a preliminary interpretation of the current findings is that lower cortical tracking in the delta band could be associated with repetition abilities that are unrelated to accessing the semantic system (i.e., using the nonlexical route to repetition; Hanley et al., 2004). Future studies can directly examine the association between lexical/nonlexical routes of repetition and cortical tracking in a larger sample of PWA.

7.0 Conclusion

The current study aimed to: i) identify the effects of speech entrainment to two rhythm types on learning scripted sentences; and to ii) identify individual characteristics in behavioral measures and cortical tracking associated with scripted-sentence learning response. These aims were designed to help characterize mechanisms (i.e., how it works) and individual characteristics associated with scripted sentence learning (i.e., for whom it works) in Spanish speakers with aphasia.

It was found that the rhythmic-enhanced conditions engendered greater scripted-sentence learning over time compared to the control condition. As entrainment depends on the detection and integration of rhythmic features (Phillips-Silver et al., 2010), highlighting these features facilitated speech entrainment, and in turn, scripted sentence learning. This finding is consistent with the rhythmic features of speech entrainment being a facilitative mechanism for scripted sentence learning. Also, learning in the two rhythmic-enhanced conditions did not differ over time. This suggests that both rhythms, supporting word stress (aiding lexical retrieval; see Soto-Faraco et al., 2001) and metronomic beats (improving memorization via chunking; see Purnell-Webb & Speelman, 2008; Stahl et al., 2011) may have benefits for scripted-sentence learning, but their differential contributions remain unknown.

Regarding individual characteristics associated with scripted-sentence learning response, there was variability in language, attention, and rhythmic performance. Participants with more severe *language* impairments showed higher learning estimates. This finding is consistent with script training evidence (e.g., Lee et al., 2009) and suggests that using speech entrainment for training scripted sentences (rehearsed language) could be especially beneficial for PWA with more severe language impairments (reducing error and effort for language production). Consistent with previous studies (e.g., Murray, 1999; Murray, 2012), there was evidence of *attentional* deficits, which were not associated with scripted sentence learning. Participants also showed low performance in *rhythm/beat* perception tasks but good performance in synchronizing with the beat of music, neither strongly associated with scripted sentence learning.

A negative linear trend suggested that PWA with higher cortical tracking in the theta band had lower learning estimates. This trend was consistent with the findings indicating that PWA with higher language ability had lower learning estimates. Taken together, a preliminary interpretation of these findings is that people with milder aphasia had higher cortical tracking, and lower learning estimates, due to the particular benefit that speech entrainment and this type of highly formulaic sentences have for more severely impaired PWA. A future analytical approach to directly test this cautious preliminary interpretation is a three-way interaction model that accounts for language impairment measures, cortical tracking, and learning estimates.

Given the role of enhancing rhythmic features in speech entrainment for scripted-sentence learning, future studies should examine various rhythmic cues, and the related learning mechanisms they support, in their association with scripted-sentence learning. Similarly, future studies with larger samples can explore additional individual characteristics that may be important for the mechanisms identified as critical in the current study. For example, working memory may be essential in holding verbal information active while simultaneously engaging in speech entrainment. Individual measures of retention capacity may also be necessary to identify how PWA carry over scripted sentences, successfully practiced during speech entrainment, to their daily life. Finally, cortical tracking of the speech envelope to a narrative with and without rhythmic enhancements (e.g., stress-aligned and metronomic beats) may help determine whether highlighting rhythmic features of speech improves speech perception and processing resulting in behavioral learning gains.

Future directions also include further cross-linguistic investigations of the rhythmic-enhanced manipulations used in this study. As discussed above, given that Spanish speakers rely strongly on suprasegmental information for lexical retrieval (Soto-Faraco et al. 2001), information highly marked in the language (Hualde, 2012), it would be expected that this manipulation is more beneficial for Spanish than for English speakers. One possibility to test this assumption is to compare the effects of enhancing word-stress in stressed-timed (e.g., English or German) and syllable-timed languages (e.g., Spanish or Italian) on scripted-sentence learning, and to explore whether lexical retrieval contributes to the learning outcomes. Similarly, it would be expected that the metronomic beat manipulation would be helpful for information chunking and recall (e.g., Purnell-Webb & Speelman, 2008) regardless of the language spoken, because of the tendency of speakers to align their spoken productions with rhythmic structures across many different language types (e.g., Cummins & Port, 1998; Tajima et al., 2000). A future study that compares this metronomic manipulation in a variety of languages could also provide direct evidence about this assumption.

In summary, the current findings again demonstrate that there is heterogeneity in response to scripted sentence learning, as has been previously reported in other script training studies (e.g., Cherney, Halper, et al., 2011; Goldberg et al., 2012; Lee et al., 2009). However, these findings provide evidence that rhythmic features, supporting speech entrainment, are a key mechanism for scripted sentence learning. Furthermore, these findings provide preliminary evidence for individual characteristics associated with scripted sentence learning: language ability, some rhythmic processing measures, and cortical tracking. Importantly, this is the first study to examine scripted-sentence learning in Spanish speakers with aphasia, demonstrating the cross-linguistic benefits of script training interventions.

References

Abel, S., Schultz, A., Radermacher, I., Willmes, K., & Huber, W. (2005). Decreasing and increasing cues in naming therapy for aphasia. *Aphasiology, 19*(9), 831-848.

Abercrombie, D. (1982). *Elements of general phonetics*. Chicago: University Press.

Alexandrou, A. M., Saarinen, T., Kujala, J., & Salmelin, R. (2018). Cortical entrainment: what we can learn from studying naturalistic speech perception. *Language, Cognition and Neuroscience*, 1-13.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica, 66*(1-2), 46-63.

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*(3), 351-373.

Assaneo, F. M., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., & Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature Neuroscience*.

Audacity Team. (2019). Audacity software is copyright © 1999-2019 (Version 2.3.3): https://audacityteam.org/.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*(1), 177-189.

Ballard, K., Wambaugh, J., Duffy, J., Layfield, C., Maas, E., Mauszycki, S., & McNeil, M. (2015). Treatment for Acquired Apraxia of Speech: A Systematic Review of Intervention Research Between 2004 and 2012. *American Journal of Speech-Language Pathology*, *24*(2), 316-337.

Baqué, L. (2017). Lexical stress contrast marking in fluent and non-fluent aphasia in Spanish: The relationship between acoustic cues and compensatory strategies. *Clinical Linguistics & Phonetics, 31*(7-9), 642-664.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang, 68*(3).

Bate, A. J., Mathias, J. L., & Crawford, J. R. (2001). Performance on the Test of Everyday Attention and Standard Tests of Attention following Severe Traumatic Brain Injury. *The Clinical Neuropsychologist, 15*(3), 405-422.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Becker, F., & Reinvang, I. (2007). Successful syllable detection in aphasia despite processing impairments as revealed by event-related potentials. *Behavioral and Brain Functions, 3*(1), 6.

Bégel, V., Benoit, C. E., Correa, A., Cutanda, D., Kotz, S. A., & Dalla Bella, S. (2017). "Lost in time" but still moving to the beat. *Neuropsychologia, 94*, 129-138.

Beveridge, M. E., & Bak, T. H. (2011). The languages of aphasia research: Bias and diversity. *Aphasiology, 25*(12), 1451-1468.

Bilda, K. (2011). Video-based conversational script training for aphasia: A therapy study. *Aphasiology, 25*(2), 191-201.

Blomert, L., Kean, M. L., Koster, C., & Schokker, J. (1994). Amsterdam—Nijmegen everyday language test: construction, reliability and validity. *Aphasiology, 8*(4), 381-407.

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience, 8*(3), 389-395.

Bonnefond, M., & Jensen, O. (2012). Alpha Oscillations Serve to Protect Working Memory Maintenance against Anticipated Distracters. *Current Biology, 22*(20), 1969-1974.

Borrie, S. A., Barrett, T. S., Willi, M. M., & Berisha, V. (2019). Syncing Up for a Good Conversation: A Clinically Meaningful Methodology for Capturing Conversational Entrainment in the Speech Domain. *Journal of Speech, Language, Hearing Research, 62*(2), 283-296.

Borrie, S. A., & Liss, J. M. (2014). Rhythm as a Coordinating Device: Entrainment With Disordered Speech. *Journal of Speech, Language, and Hearing Research, 57*(3), 815-824.

Borrie, S. A., Lubold, N., & Pon-Barry, H. (2015). Disordered speech disrupts conversational entrainment: a study of acoustic-prosodic entrainment and communicative success in populations with communication challenges. *Frontiers in Psychology, 6*, 1187.

Brallier, J. (2002). *Who was Albert Einstein?* : Penguin Workshop.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*(2), B13-B25.

Brendel, B., & Ziegler, W. (2008). Effectiveness of metrical pacing in the treatment of apraxia of speech. *Aphasiology, 22*(1), 77-102.

Brennan, S. E., & Clark, H. H. (1996). Conceptual Pacts and Lexical Choice in Conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1482-1493.

Breslow, N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American statistical Association, 88*(421), 9-25.

Brodbeck, C., Presacco, A., Anderson, S., & Simon, J. Z. (2018). Over-representation of speech in older adults originates from early response in higher order auditory cortex. *Acta Acust United Acust, 104*(5), 774-777.

Burkard, R. F., Eggermont, J. J., & Don, M. (2007). *Auditory evoked potentials: basic principles and clinical application*: Lippincott Williams & Wilkins.

Butler, R. A., Lambon Ralph, M. A., & Woollams, A. (2014). Capturing multidimensionality in stroke aphasia: Mapping principal behavioural components to neural structures. *Brain: a journal of neurology, 137*(12).

Chan, R. (2000). Attentional deficits in patients with closed head injury: a further study to the discriminative validity of the test of everyday attention. *Brain Injury, 14*(3), 227-236.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The Natural Statistics of Audiovisual Speech. *PLOS Computational Biology, 5*(7), e1000436.

Chen, H.-C., Koh, C.-L., Hsieh, C.-L., & Hsueh, I. P. (2013). Test of Everyday Attention in patients with chronic stroke: Test–retest reliability and practice effects. *Brain Injury, 27*(10), 1148-1154.

Cherney, L. R., & Halper, A. S. (2008). Novel technology for treating individuals with aphasia and concomitant cognitive deficits. *Topics in Stroke Rehabilitation, 15*(6), 542-554.

Cherney, L. R., Halper, A. S., Holland, A. L., & Cole, R. (2008). Computerized script training for aphasia: Preliminary results. *American Journal of Speech-Language Pathology, 17*(1), 19-34.

Cherney, L. R., Halper, A. S., & Kaye, R. C. (2011). Computer-based script training for aphasia: Emerging themes from post-treatment interviews. *Journal of Communication Disorders, 44*(4), 493-501.

Cherney, L. R., Kaye, R. C., Lee, J. B., & Van Vuuren, S. (2015). Impact of personal relevance on acquisition and generalization of script training for aphasia: A preliminary analysis. *American Journal of Speech-Language Pathology, 24*(4), S913-S922.

Cherney, L. R., Kaye, R. C., & van Vuuren, S. (2014). Acquisition and Maintenance of Scripts in Aphasia: A Comparison of Two Cuing Conditions. *American Journal of Speech-Language Pathology*, *23*, S343–S360

Cherney, L. R., Patterson, J. P., & Raymer, A. M. (2011). Intensity of aphasia therapy: Evidence and efficacy. *Current neurology and neuroscience reports, 11*(6), 560.

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009a). A comparison of word versus sentence cues as therapy for verb naming in aphasia. *Aphasiology*, *23*(4), 462-482.

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009b). The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia. *Aphasiology, 23*(6), 707-730.

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009c). Errorless and errorful therapy for verb and noun naming in aphasia. *Aphasiology*, *23*(11), 1311-1337.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of Lexical Stress on Lexical Access in English: Evidence from Native and Non-native Listeners. *Language and Speech, 45*(3), 207-228.

Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Front. Hum. Neurosci.*, *10*(604).

Crosson, B., Rodriguez, A. D., Copland, D., Fridriksson, J., Krishnamurthy, L. C., Meinzer, M., . . . Leff, A. P. (2019). Neuroplasticity and aphasia treatments: new approaches for an old problem. *Journal of Neurology, Neurosurgery & Psychiatry*, jnnp-2018-319649.

Cruice, M., Worrall, L., & Hickson, L. (2006). Perspectives of quality of life by people with aphasia and their family: Suggestions for successful living. *Topics in Stroke Rehabilitation, 13*(1), 14-24.

Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica, 32*, 133-143.

Cummins, F. (2002). *Speech rhythm and rhythmic taxonomy.* Paper presented at the Speech Prosody 2002, Aix-en-Provence, France.

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26*(2), 145-171.

Curtis, S., Nicholas, M. L., Pittmann, R., & Zipse, L. (2020). Tap your hand if you feel the beat: differential effects of tapping in melodic intonation therapy. *Aphasiology, 34*(5), 580-602.

Curtis, S., & Zipse, L. (2017). *Tap Your Hand if You Feel the Beat: Differential Effects of Tapping in Melodic Intonation Therapy.* Paper presented at the Academy of Aphasia 55th Annual Meeting, Baltimore, MD.

Cutler, A. (1989). Auditory lexical access: Where do we start? In W. Marslen-Wilson (Ed.), *Lexical Representation and Process* (pp. 342-356). Cambridge, Massachusetts: Cambridge: MIT Press.

Cutler, A. (1990). Exploiting prosodic probabilities in speech segmentation. In *Cognitive models of speech processing: Psycholinguistic and computational perspectives.* (pp. 105-121). Cambridge, MA, US: The MIT Press.

Cutler, A. (2005). Lexical Stress. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 264-289). United Kingdom Blackwell Publishing Ltd.

Cutler, A. (2012). *Native Listening. Language Experience and the Recognition of Spoken Words*: Mit Press.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language, 2*(3), 133-142.

Cutler, A., Mister, E., Norris, D., & Sebastian-Gallés, N. (2004). Chapitre 3. La perception de la parole en espagnol: un cas particulier? In *Psycholinguistique Cognitive* (pp. 57-74). Louvain-la-Neuve: De Boeck Supérieur.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human perception and performance, 14*(1), 113-121.

Dasher, R., & Bolinger, D. (1982). On pre-accentual lengthening. *Journal of the International Phonetic Association, 12*(2), 58-71.

Davis, H., & Zerlin, S. (1966). Acoustic relations of the human vertex potential. *J Acoust Soc Am, 39*(1), 109-116.

Davis, L., & Copeland, K. (2006). Computer use in treatment of aphasia—A survey of practice patterns and opinions. *Contemporary Issues in Communication Sciences and Disorders, 33*, 138-146.

de Bree, E., Janse, E., & van de Zande, A. M. (2007). Stress assignment in aphasia: Word and non-word reading and non-word repetition. *Brain and Language*, *103*(3), 264-275.

de Cheveigné, A., & Nelken, I. (2019). Filters: When, Why, and How (Not) to Use Them. *Neuron*, *102*(2), 280-293.

Delattre, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics in Language Teaching, 7*(4), 295-326.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods, 134*(1), 9-21.

Di Liberto, G. M., & Lalor, E. C. (2017). Indexing cortical entrainment to natural speech at the phonemic level: Methodological considerations for applied research. *Hear Res, 348*, 70-77.

Di Liberto, Giovanni M., O'Sullivan, James A., & Lalor, Edmund C. (2015). Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing. *Current Biology, 25*(19), 2457-2465.

Dial, H., Zinszer, B., Chandrasekaran, B., & Henry, M. (2018). *Neural encoding of phonetic features relates to phonological processing in PPA.* Paper presented at the 56th Annual Meeting of the Academy of Aphasia., Montreal, Canada.

Ding, N., Chatterjee, M., & Simon, J. Z. (2014). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *NeuroImage, 88*, 41-46.

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience, 19*(1), 158-164.

Ding, N., & Simon, J. Z. (2012). *Emergence of neural encoding of auditory objects while listening to competing speakers.* Paper presented at the National Academy of Sciences

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. *Frontiers in human neuroscience, 8*, 311.

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences, 112*(45), E6233.

Doyle, P., McNeil, M., Hula, W., & Mikolic, J. (2003). The Burden of Stroke Scale (BOSS): Validating patient-reported communication difficulty and associated psychological distress in stroke survivors. *Aphasiology, 17*(3), 291-304.

Doyle, P. J., McNeil, M. R., Mikolic, J. M., Prieto, L., Hula, W. D., Lustig, A. P., . . . Elman, R. J. (2004). The Burden of Stroke Scale (BOSS) provides valid and reliable score estimates of functioning and well-being in stroke survivors with and without communication disorders. *Journal of Clinical Epidemiology, 57*(10), 997-1007.

Duffy, J. R. (2005). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management* (2nd ed.). New York: Mosby.

Duffy, J. R., Strand, E. A., & Josephs, K. A. (2014). Motor Speech Disorders Associated with Primary Progressive Aphasia. *Aphasiology*, *28*(8-9), 1004-1017.

Dworkin, J. P., Abkarian, G. G., & Johns, D. F. (1988). Apraxia of Speech. *Journal of Speech and Hearing Disorders, 53*(3), 280-294.

Edmonds, L. A., Mammino, K., & Ojeda, J. (2014). Effect of verb network strengthening treatment (VNeST) in persons with aphasia: Extension and replication of previous findings. *American Journal of Speech-Language Pathology, 23*(2), S312-S329.

Edmonds, L. A., Nadeau, S. E., & Kiran, S. (2009). Effect of Verb Network Strengthening Treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia. *Aphasiology, 23*(3), 402-424.

Edmonds, L. A., Obermeyer, J., & Kernan, B. (2015). Investigation of pretreatment sentence production impairments in individuals with aphasia: Towards understanding the linguistic variables that impact generalisation in Verb Network Strengthening Treatment. *Aphasiology*, *29*(11), 1312-1344.

Etard, O., & Reichenbach, T. (2019). Neural Speech Tracking in the Theta and in the Delta Frequency Band Differentially Encode Clarity and Comprehension of Speech in Noise. *The Journal of Neuroscience, 39*(29), 5750.

Fear, B. D., Cutler, A., & Butterfield, S. (1995). The strong/weak syllable distinction in English. *Journal of the Acoustical Society of America, 97*(3), 1893-1904.

Fillingham, J., Hodgson, C., Sage, K., & Lambon Ralph, M. (2003). The application of errorless learning to aphasic disorders: A review of theory and practice. *Neuropsychological Rehabilitation, 13*(3), 337-363.

Fillingham, J., Sage, K., & Lambon Ralph, M. (2005a). Further explorations and an overview of errorless and errorful therapy for aphasic word-finding difficulties: The number of naming attempts during therapy affects outcome. *Aphasiology*, *19*(7), 597-614.

Fillingham, J., Sage, K., & Lambon Ralph, M. (2005b). Treatment of anomia using errorless versus errorful learning: are frontal executive skills and feedback important? *International Journal of Language & Communication Disorders, 40*(4), 505-523.

Fillingham, J., Sage, K., & Lambon Ralph, M. (2006). The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation, 16*(2), 129-154.

Finch, E., & Hill, A. J. (2014). Computer use by People with Aphasia: A Survey Investigation. *Brain Impairment, 15*(2), 107-119.

Flamand-Roze, C., Falissard, B., Roze, E., Maintigneux, L., Beziz, J., Chacon, A., . . . Denier, C. (2011). Validation of a new language screening tool for patients with acute stroke. The Language Screening Test (LAST). *Stroke, 42*(5), 1224-1229.

Frey, J. N., Ruhnau, P., & Weisz, N. (2015). Not so different after all: The same oscillatory processes support different types of attention. *Brain Res, 1626*, 183-197.

Fridriksson, J., Basilakos, A., Hickok, G., Bonilha, L., & Rorden, C. (2015). Speech entrainment compensates for Broca's area damage. *Cortex, 69*, 68-75.

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*(Pt 12), 3815-3829.

Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 360*(1456), 815-836.

Fuglsang, S. A., Märcher-Rørsted, J., Dau, T., & Hjortkjær, J. (2020). Effects of Sensorineural Hearing Loss on Cortical Synchronization to Competing Speech during Selective Attention. *The Journal of Neuroscience, 40*(12), 2562-2572.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*: Cambridge university press.

Gilmore, N., Meier, E. L., Johnson, J. P., & Kiran, S. (2019). Non-linguistic cognitive factors predict treatment-induced recovery in chronic post-stroke aphasia. *Archives of Physical Medicine and Rehabilitation, 100*(1), 1251-1258.

Giraud, A. L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., & Kleinschmidt, A. (2000). Representation of the temporal envelope of sounds in the human brain. *J Neurophysiol, 84*(3), 1588-1598.

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci, 15*(4), 511-117.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*(6), 236-243.

Goldberg, S., Haley, K. L., & Jacks, A. (2012). Script Training and Generalization for People With Aphasia. *American Journal of Speech-Language Pathology, 21*(3), 222-238.

Goossens, T., Vercammen, C., Wouters, J., & van Wieringen, A. (2016). Aging Affects Neural Synchronization to Speech-Related Acoustic Modulations. *Front Aging Neurosci, 8*, 133.

Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology, 4*, 67-92.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology, 7*(515-546).

Grasso, S. M., Cruz, D. F., Benavidez, R., Peña, E. D., & Henry, M. L. (2019). Video-Implemented Script Training in a Bilingual Spanish–English Speaker With Aphasia. *Journal of Speech, Language, and Hearing Research, 62*(7), 2295-2316.

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech—a syllable-centric perspective. *Journal of Phonetics, 31*(3), 465-485.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2014). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLOS Biology, 11*(12), e1001752.

Gueguen, N., Jacob, C., & Martin, A. (2009). Mimicry in social interaction: Its effect on human judgment and behavior. *European Journal of Social Sciences, 8*(2), 253-259.

Gutierrez-Bravo, R. (2002). Focus, word order variation and intonation in Spanish and English: An OT account. In *Romance phonology and variation* (pp. 39-53): John Benjamins.

Halai, A. D., Woollams, A. M., & Lambon Ralph, M. A. (2017). Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural correlates of speech fluency, phonology and semantics. *Cortex, 86*, 275-289.

Haley, K. L., & Jacks, A. (2019). Word-level prosodic measures and the differential diagnosis of apraxia of speech. *Clinical Linguistics & Phonetics*, *33*(5), 479-495.

Hambrook, D. A., & Tata, M. S. (2014). Theta-band phase tracking in the two-talker problem. *Brain and Language*, *135*, 52-56.

Hamilton, L. S., & Huth, A. G. (2018). The revolution will not be controlled: Natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience*, 1-10.

Hanley, R. J., Dell, G. S., Kay, J., & Baron, R. (2004). Evidence for the involvement of a nonlexical route in the repetition of familiar words: A comparison of single and dual route models of auditory repetition. *Cognitive Neuropsychology*, *21*(2-4), 147-158.

Harding, E. E., Sammler, D., Henry, M. J., Large, E. W., & Kotz, S. A. (2019). Cortical tracking of rhythm in music and speech. *NeuroImage, 185*, 96-101.

Harrison, P. M. C. (2019a). cabat: R Implementation of the Computerised Beat Alignment Test (CA-BAT) (Version 0.8.0 R package).

Harrison, P. M. C. (2019b). psychTestR: Framework for delivering psychological tests via Shiny (Version 2.9.0 R package). Retrieved from https://github.com/pmcharrison/psychTestR

Harrison, P. M. C., & Müllensiefen, D. (2018a). Computerised Adaptive Beat Alignment Test (CA-BAT), psychTestR implementation. *Zenodo*.

Harrison, P. M. C., & Müllensiefen, D. (2018b). Development and Validation of the Computerised Adaptive Beat Alignment Test (CA-BAT). *Scientific reports, 8*(1), 12395-12395.

Hebb, D. O. (1949). *The Organization of Behavior: a Neuropsychological Theory*. New York: Wiley.

Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory* (T. F. Group Ed.): Psychology Press.

Heeringa, A. N., & van Dijk, P. (2019). Neural coding of the sound envelope is changed in the inferior colliculus immediately following acoustic trauma. *Eur J Neurosci, 49*(10), 1220-1232.

Henry, M. J., Herrmann, B., Kunke, D., & Obleser, J. (2017). Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. *Nature Communications, 8*(1), 15801.

Herrmann, B., Buckland, C., & Johnsrude, I. S. (2019). The effects of aging on neural signatures of temporal regularity processing in sounds. *bioRxiv*, 522375.

Hickok, G. (2014). The architecture of speech production and the role of the phoneme in speech processing. *Language and cognitive processes, 29*(1), 2-20.

Hickok, G., Farahbod, H., & Saberi, K. (2015). The Rhythm of Perception: Entrainment to Acoustic Rhythms Induces Subsequent Perceptual Oscillation. *Psychological science, 26*(7), 1006-1013.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal Grouping Effects in Immediate Recall: A Working Memory Analysis. *The Quarterly Journal of Experimental Psychology Section A, 49*(1), 116-139.

Holland, A. L., Halper, A. S., & Cherney, L. R. (2010). Tell me your story: Analysis of script topics selected by persons with aphasia. *American Journal of Speech-Language Pathology, 19*(3), 198-203.

Howard, D., & Nickels, L. (1999). Effects of lexical stress on aphasic word production. *Clinical Linguistics & Phonetics, 13*(4), 269-294.

Howard, D., & Smith, K. (2002). The effects of lexical stress in aphasic word production. *Aphasiology, 16*(1-2), 198-237.

Howard, M. F., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of Neurophysiology, 104*(5), 2500-2511.

Hualde, J. I. (2012). Stress and Rhythm. In *The Handbook of Hispanic Linguistics* (pp. 153-171).

Hubbard, H. I., Nelson, L. A., & Richardson, J. D. (2020). Can Script Training Improve Narrative and Conversation in Aphasia across Etiology? *Seminars in speech and language, 41*(01), 99-124.

Huber, W., Poeck, K., Weniger, D., & Willmes, K. (1983). *Aachener Aphasie Test (AAT): Handanweisung*: Verlag für Psychologie Hogrefe.

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLOS Medicine, 2*(8), e124.

Ioannidis, J. P. A. (2019). What Have We (Not) Learnt from Millions of Scientific Papers with P Values? *The American Statistician, 73*(sup1), 20-25.

Iversen, J., & Patel, A. (2008). *The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population.* Paper presented at the 10th International Conference on Music Perception, and Cognition (ICMPC 10), Sapporo, Japan.

Jung, T. P., Makeig, S., McKeown, M. J., Bell, A. J., Lee, T. W., & Sejnowski, T. J. (2001). Imaging Brain Dynamics Using Independent Component Analysis. *Proc IEEE Inst Electr Electron Eng, 89*(7), 1107-1122.

Kale, S., & Heinz, M. G. (2010). Envelope coding in auditory nerve fibers following noise-induced hearing loss. *J Assoc Res Otolaryngol, 11*(4), 657-673.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test*: Pro-ed.

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

Kaye, R. C., & Cherney, L. R. (2016). Script templates: A practical approach to script training in aphasia. *Topics in language disorders, 36*(2), 136.

Kendall, D. L., Rosenbek, J. C., Heilman, K. M., Conway, T., Klenberg, K., Gonzalez Rothi, L. J., & Nadeau, S. E. (2008). Phoneme-based rehabilitation of anomia in aphasia. *Brain Lang, 105*(1), 1-17.

Kerlin, J. R., Shahin, A. J., & Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a "cocktail party". *The Journal of neuroscience : the official journal of the Society for Neuroscience, 30*(2), 620-628.

Kershenbaum, A., Nicholas, M. L., Hunsaker, E., & Zipse, L. (2017). Speak along without the song: what promotes fluency in people with aphasia? *Aphasiology*, 1-24.

Kertesz, A. (1982). *Western aphasia battery test manual*: Psychological Corp.

Kertesz, A., & Pascual-Leone, Á. (2000). *Batería de afasías Western (The western aphasia battery en versión y adaptación castellana)*: Nau llibres.

Keshavarzi, M., Kegler, M., Kadir, S., & Reichenbach, T. (2020). Transcranial alternating current stimulation in the theta band but not in the delta band modulates the comprehension of naturalistic speech in noise. *NeuroImage, 210*, 116557.

Kimelman, M. (1991). The role of target word stress in auditory comprehension by aphasic listeners. *Journal of Speech, Language, and Hearing Research, 34*(2), 334-339.

Kimelman, M. D. (1999). Prosody, Linguistic Demands, and Auditory Comprehension in Aphasia. *Brain and Language, 69*(2), 212-221.

Kimelman, M. D. Z., & McNeil, M. R. (1987). An investigation of emphatic stress comprehension in adult aphasia: A replication. *Journal of Speech, Language, Hearing Research, 30*(3), 295-300.

Kiran, S., & Thompson, C. K. (2019). Neuroplasticity of Language Networks in Aphasia: Advances, Updates, and Future Challenges. *Front. Neurol, 10*(295).

Klem, G., Lüders, H., Jasper, H., & Elger, C. (1999). The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl, 52*, 3-6.

Kösem, A., & van Wassenhove, V. (2017). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Language, Cognition and Neuroscience, 32*(5), 536-544.

Krause, J. C., & Braida, L. D. (2004). Acoustic Properties of Naturally Produced Clear Speech at Normal Speaking Rates. *The Journal of the Acoustical Society of America, 115*(1), 362-378.

Kubanek, J., Brunner, P., Gunduz, A., Poeppel, D., & Schalk, G. (2013). The tracking of speech envelope in the human cortex. *PLOS ONE, 8*(1), e53398-e53398.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13).

Lagrois, M.-É., Palmer, C., & Peretz, I. (2019). Poor Synchronization to Musical Beat Generalizes to Speech. *Brain sciences, 9*(7), 157.

Lakatos, P., Gross, J., & Thut, G. (2019). A New Unifying Account of the Roles of Neuronal Entrainment. *Current Biology, 29*(18), R890-R905.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science, 320*(5872), 110-113.

Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., & Schroeder, C. E. (2009). The Leading Sense: Supramodal Control of Neurophysiological Context by Attention. *Neuron, 64*(3), 419-430.

Lambon Ralph, M., Snell, C., Fillingham, J., Conroy, P., & Sage, K. (2010). Predicting the outcome of anomia therapy for people with aphasia post CVA: both language and cognitive status are key predictors. *Neuropsychol Rehabil, 20*(2), 289-305.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119.

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.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral And Brain Sciences, 22*(1), 1-38.

Levitin, D. J., Grahn, J. A., & London, J. (2018). The Psychology of Music: Rhythm and Movement. *Annual review of psychology, 69*.

Liang, C. L., Chang, H. W., Lu, K., Lee, T. C., Liliang, P. C., Lu, C. H., & Chen, H. J. (2001). Early prediction of aphasia outcome in left basal ganglia hemorrhage. *Acta Neurologica Scandinavica, 103*(3), 148-152.

Liégeois-Chauvel, C., Lorenzi, C., Trébuchon, A., Régis, J., & Chauvel, P. (2004). Temporal envelope processing in the human left and right auditory cortices. *Cereb Cortex, 14*(7), 731-740.

Llisterri, J., Machuca, M., de la Mota, C., Riera, M., & Ríos, A. (2003). *The perception of lexical stress in Spanish.* Paper presented at the Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain

Llisterri, J., Machuca, M. J., de la Mota, C., Riera, M., & Ríos, A. (2005). La percepción del acento léxico en español. *Filología y lingüística. Estudios ofrecidos a Antonio Quilis, 1*, 271-297.

Logan, G. D. (1988). Automaticity, resources, and memory: theoretical controversies and practical implications. *Hum Factors, 30*(5), 583-598.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive science, 36*(8), 1404-1426.

Lüdecke, D. (2020). sjPlot: Data visualization for statistics in social science. *R package version 2.8.3, 2*(1).

Manheim, L. M., Halper, A. S., & Cherney, L. (2009). Patient-reported changes in communication after computer-based script training for aphasia. *Archives of Physical Medicine and Rehabilitation, 90*(4), 623-627.

Martin, N., Fink, R. B., Renvall, K., & Laine, M. (2006). Effectiveness of contextual repetition priming treatments for anomia depends on intact access to semantics. *J Int Neuropsychol Soc, 12*(6), 853-866.

MathWorks Inc. MATLAB R2019 a. Massachusetts, USA.

McElhinney, M., & Annett, J. M. (1996). Pattern of Efficacy of a Musical Mnemonic on Recall of Familiar Words over Several Presentations. *Perceptual and Motor Skills, 82*(2), 395-400.

McHaney, J. R., Gnanateja, G. N., Smayda, K. E., Zinszer, B. D., & Chandrasekaran, B. (2020). Cortical Tracking of Speech in Delta Band Relates to Individual Differences in Speech in Noise Comprehension in Older Adults. *Ear Hear*.

Middleton, E. L., Schwartz, M. F., Rawson, K. A., & Garvey, K. (2015). Test-enhanced learning versus errorless learning in aphasia rehabilitation: testing competing psychological principles. *Journal of experimental psychology. Learning, memory, and cognition, 41*(4), 1253-1261.

Millman, R. E., Johnson, S. R., & Prendergast, G. (2015). The Role of Phase-locking to the Temporal Envelope of Speech in Auditory Perception and Speech Intelligibility. *Journal of Cognitive Neuroscience, 27*(3), 533-545.

Molinaro, N., & Lizarazu, M. (2018). Delta(but not theta)-band cortical entrainment involves speech-specific processing. *Eur J Neurosci, 48*(7), 2642-2650.

Mora, E. (1998). Acústica del acento español en su variedad venezolana. *Lengua y Habla, 3*(1), 70-78.

Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., . . . Jung, T. (2013, 3-7 July 2013). *Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG.* Paper presented at the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

Murdock, B. B. (1995). Developing TODAM: Three models for serial-order information. *Memory & Cognition, 23*(5), 631-645.

Murray, L. L. (1999). Review Attention and aphasia: theory, research and clinical implications. *Aphasiology, 13*(2), 91-111.

Murray, L. L. (2002). Attention Deficits in Aphasia: Presence, Nature, Assessment, and Treatment. *Seminars in Speech and Language, 23*(2), 107-116.

Murray, L. L. (2012). Attention and Other Cognitive Deficits in Aphasia: Presence and Relation to Language and Communication Measures. *American Journal of Speech-Language Pathology, 21*(2), S51-S64.

Murton, O., Zipse, L., Jacoby, N., & Shattuck-Hufnagel, S. (2017). Repetition and a Beat-Based Timing Framework: What Determines the Duration of Intervals Between Repetitions of a Tapping Pattern? *Timing & Time Perception, 5*(3-4), 244-259.

Nadeu, M. (2014). Stress- and speech rate-induced vowel quality variation in Catalan and Spanish. *Journal of Phonetics, 46*, 1-22.

National Institute of Neurological Disorders and Stroke. (2018). Aphasia information. Retrieved from https://www.ninds.nih.gov/Disorders/All-Disorders/Aphasia-Information-Page

Nobis-Bosch, R., Springer, L., Radermacher, I., & Huber, W. (2011). Supervised Home Training of Dialogue Skills in Chronic Aphasia: A Randomized Parallel Group Study. *Journal of Speech, Language and Hearing Research (Online), 54*(4), 1118-1136A.

Nolan, F., & Jeon, H.-S. (2014). Speech rhythm: a metaphor? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 369*(1658), 20130396.

Notbohm, A., & Herrmann, C. S. (2016). Flicker Regularity Is Crucial for Entrainment of Alpha Oscillations. *Frontiers in human neuroscience, 10*(503).

Nozari, N., Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2010). Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of Memory and Language, 63*(4), 541-559.

Nunes-Silva, M., & Haase, V. G. (2012). Montreal Battery of Evaluation of Amusia: Validity evidence and norms for adolescents in Belo Horizonte, Minas Gerais, Brazil. *Dementia & neuropsychologia, 6*(4), 244-252.

O'Sullivan, J., Chen, Z., Herrero, J., McKhann, G. M., Sheth, S. A., Mehta, A. D., & Mesgarani, N. (2017). Neural decoding of attentional selection in multi-speaker environments without access to clean sources. *Journal of Neural Engineering, 14*(5), 056001.

Obleser, J., & Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cerebral cortex (New York, N.Y. : 1991), 22*(11), 2466-2477.

Ortega-Llebaria, M. (2006). *Phonetic cues to stress and accent in Spanish.* Paper presented at the 2nd conference on laboratory approaches to spanish phonetics and phonology, Somerville, MA.

Ortega-Llebaria, M., Gu, H., & Fan, J. (2013). English speakers' perception of Spanish lexical stress: Context-driven L2 stress perception. *Journal of Phonetics, 41*(3), 186-197.

Ortega-Llebaria, M., Olson, D. J., & Tuninetti, A. (2019). Explaining Cross-Language Asymmetries in Prosodic Processing: The Cue-Driven Window Length Hypothesis. *Language and Speech, 62*(4), 701-736.

Ortega-Llebaria, M., & Prieto, P. (2011). Acoustic Correlates of Stress in Central Catalan and Castilian Spanish. *Language and Speech, 54*(1), 73-97.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in systems neuroscience, 8*, 57-57.

Peelle, J. E. (2018). Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear and Hearing, 39*(2).

Peelle, J. E., & Davis, M. H. (2012). Neural Oscillations Carry Speech Rhythm through to Comprehension. *Frontiers in Psychology, 3*(320), 1-17.

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral cortex, 23*(6), 1378-1387.

Pefkou, M., Arnal, L. H., Fontolan, L., & Giraud, A. L. (2017). θ-Band and β-Band Neural Activity Reflects Independent Syllable Tracking and Comprehension of Time-Compressed Speech. *J Neurosci, 37*(33), 7930-7938.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., . . . Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195-203.

Pellegrino, F., Coupé, C., & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language, 87*(3), 539-558.

Peñaloza, C., Benetello, A., Tuomiranta, L., Heikius, I.-M., Järvinen, S., Majos, M. C., . . . Rodríguez-Fornells, A. (2015). Speech segmentation in aphasia. *Aphasiology, 29*(6), 724-743.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Ann N Y Acad Sci, 999*, 58-75.

Pérez, A., Carreiras, M., Gillon-Dowens, M., & Duñabeitia, J. A. (2015). Differential oscillatory encoding of foreign speech. *Brain and Language, 147*, 51-57.

Phillips-Silver, J., Aktipis, A., & Bryant, G. (2010). The Ecology of Entrainment: Foundations of Coordinated Rhythmic Movement. *Music Perception: An Interdisciplinary Journal, 28*(1), 3-14.

Phillips-Silver, J., & Trainor, L. (2005). Feeling the beat: movement influences infant rhythm perception. *Science, 308*(5727), 1430.

Pike, K. L. (1945). *The Intonation of American English*: University of Michigan Press.

Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 564-573.

Plechawska-Wojcik, M., Kaczorowska, M., & Zapala, D. (2019, 2019//). *The Artifact Subspace Reconstruction (ASR) for EEG Signal Correction. A Comparative Study.* Paper presented at the Information Systems Architecture and Technology: Proceedings of 39th International Conference on Information Systems Architecture and Technology – ISAT 2018, Cham.

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research, 9*(3), 233-265.

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience, 21*(6), 322-334.

Port, R. F. (2003). Meter and speech. *Journal of Phonetics, 31*(3-4), 599-611.

Presacco, A., Simon, J. Z., & Anderson, S. (2016a). Effect of informational content of noise on speech representation in the aging midbrain and cortex. *Journal of Neurophysiology, 116*(5), 2356-2367.

Presacco, A., Simon, J. Z., & Anderson, S. (2016b). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *Journal of Neurophysiology, 116*(5), 2346-2355.

Price, C. J., Seghier, M. L., & Leff, A. P. (2010). Predicting language outcome and recovery after stroke: the PLORAS system. *Nature Reviews Neurology, 6*, 202.

Purnell-Webb, P., & Speelman, C. P. (2008). Effects of Music on Memory for Text. *Perceptual and Motor Skills, 106*(3), 927-957.

Quique, Y., Dresang, H., & Dickey, M. W. (2018). Semantic memory for objects, actions, and events in Colombian people with aphasia. *Aphasiology, 32*(sup1), 172-175.

Quique, Y., Swiderski, A. M., Hula, W., & Dickey, M. W. (2018). Meta-analysis of Treatment of Underlying Forms: dosage-related and person-level predictors of acquisition and generalization response. *Front. Hum. Neurosci*.

Quique, Y. M., Evans, W. S., & Dickey, M. W. (2019). Acquisition and Generalization Responses in Aphasia Naming Treatment: A Meta-Analysis of Semantic Feature Analysis Outcomes. *American Journal of Speech-Language Pathology, 28*(1S), 230-246.

R Core Team. (2020). R: A Language and Environment for Statistical Computing (Version 3.6.3): R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Racette, A., Bard, C., & Peretz, I. (2006). Making non-fluent aphasics speak: sing along! *Brain, 129*(10), 2571-2584.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265-292.

Repp, B. (2010). Sensorimotor synchronization and perception of timing: effects of music training and task experience. *Hum Mov Sci, 29*(2), 200-213.

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review, 20*(3), 403-452.

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive Aging and the Compensation Hypothesis. *Current Directions in Psychological Science, 17*(3), 177-182.

Rhodes, N. C., & Isaki, E. (2018). Script training using telepractice with two adults with chronic non-fluent aphasia. *International journal of telerehabilitation, 10*(2), 89.

Richie, C. (2015). Speech recognition and listening effort across various conditions in adults with aphasia. *The Journal of the Acoustical Society of America, 137*(4), 2236-2236.

Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia naming test: Scoring and rationale. *Clinical aphasiology, 24*, 121-133.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). The test of everyday attention (TEA). In. England: Thames Valley Test Company.

Robson, H., Cloutman, L., Keidel, J. L., Sage, K., Drakesmith, M., & Welbourne, S. (2014). Mismatch negativity (MMN) reveals inefficient auditory ventral stream function in chronic auditory comprehension impairments. *Cortex, 59*, 113-125.

Rochon, E., Laird, L., Bose, A., & Scofield, J. (2005). Mapping therapy for sentence production impairments in nonfluent aphasia. *Neuropsychological Rehabilitation, 15*(1), 1-36.

Romanelli, S., & Menegotto, A. C. (2015). English speakers learning Spanish: Perception issues regarding vowels and stress. *Journal of language teaching and research, 6*(1), 30-42.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London, 336*(1278), 367-373.

RStudio Team. (2016). RStudio: Integrated Development Environment for R (Version 1.1.456). Boston, MA: RStudio, Inc. Retrieved from http://www.rstudio.com/

Särkämö, T., Tervaniemi, M., Soinila, S., Autti, T., Silvennoinen, H. M., Laine, M., & Hietanen, M. (2009). Cognitive deficits associated with acquired amusia after stroke: A neuropsychological follow-up study. *Neuropsychologia, 47*(12), 2642-2651.

Schuchard, J., & Middleton, E. (2018a). The roles of retrieval practice versus errorless learning in strengthening lexical access in aphasia. *Journal of Speech, Language, and Hearing Research*, 1-18.

Schuchard, J., & Middleton, E. (2018b). Word repetition and retrieval practice effects in aphasia: Evidence for use-dependent learning in lexical access. *Cognitive Neuropsychology*, 1-17.

Schumacher, R., Halai, A. D., & Lambon Ralph, M. A. (2019). Assessing and mapping language, attention and executive multidimensional deficits in stroke aphasia. *Brain, 142*(10), 3202-3216.

Schwartz, M. F., Saffran, E. M., Fink, R. B., Myers, J. L., & Martin, N. (1994). Mapping therapy: A treatment programme for agrammatism. *Aphasiology, 8*(1), 19-54.

Seghier, M. L., Patel, E., Prejawa, S., Ramsden, S., Selmer, A., Lim, L., . . . Price, C. J. (2016). The PLORAS Database: A data repository for Predicting Language Outcome and Recovery After Stroke. *NeuroImage, 124*, 1208-1212.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(4), 592-608.

Sihvonen, A. J., Ripollés, P., Leo, V., Rodríguez-Fornells, A., Soinila, S., & Särkämö, T. (2016). Neural Basis of Acquired Amusia and Its Recovery after Stroke. *The Journal of Neuroscience*, *36*(34), 8872-8881.

Simic, T., Chambers, C., Bitan, T., Stewart, S., Goldberg, D., Laird, L., . . . Rochon, E. (2020). Mechanisms underlying anomia treatment outcomes. *Journal of Communication Disorders, 88*, 106048.

Slaney, M. (1998). Auditory toolbox. *Interval Research Corporation, Tech. Rep, 10*(1998).

Slansky, B. L., & McNeil, M. R. (1997). Resource allocation in auditory processing of emphatically stressed stimuli in aphasia. *Aphasiology*, *11*(4-5), 461-472.

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive Top-Down Integration of Prior Knowledge during Speech Perception. *The Journal of Neuroscience, 32*(25), 815–836.

Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and Suprasegmental Mismatch in Lexical Access. *Journal of Memory and Language, 45*(3), 412-432.

Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia*, *51*(10), 1952-1963.

Sowiński, J., & Dalla-Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia, 51*(10), 1952-1963.

Stahl, B. (2013). *Treatment of non-fluent aphasia through melody, rhythm and formulaic language.* Max Planck Institute for Human Cognitive and Brain Sciences Leipzig,

Stahl, B., Henseler, I., Turner, R., Geyer, S., & Kotz, S. A. (2013). How to engage the right brain hemisphere in aphasics without even singing: evidence for two paths of speech recovery. *Frontiers in human neuroscience, 7*, 35.

Stahl, B., Kotz, S. A., Henseler, I., Turner, R., & Geyer, S. (2011). Rhythm in disguise: why singing may not hold the key to recovery from aphasia. *Brain*, *134*(10), 3083-3093.

Tajima, K., Zawaydeh, B. A., Kitahara, M., & Port, R. F. (2000). Finding acoustic evidence for speech rhythm across languages:A speech cycling approach. *Acoustical Science and Technology, 21*(5), 287-289.

Teng, X., Tian, X., Doelling, K., & Poeppel, D. (2018). Theta band oscillations reflect more than entrainment: behavioral and neural evidence demonstrates an active chunking process. *Eur J Neurosci, 48*(8), 2770-2782.

Thompson, C. K., Riley, E. A., den Ouden, D.-B., Meltzer-Asscher, A., & Lukic, S. (2013). Training verb argument structure production in agrammatic aphasia: behavioral and neural recovery patterns. *Cortex; a journal devoted to the study of the nervous system and behavior, 49*(9), 2358-2376.

Thompson, C. K., & Shapiro, L. P. (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of Underlying Forms. *Aphasiology, 19*(10-11), 1021-1036.

Tilsen, S., & Arvaniti, A. (2013). Speech rhythm analysis with decomposition of the amplitude envelope: Characterizing rhythmic patterns within and across languages. *The Journal of the Acoustical Society of America, 134*(1), 628-639.

Toledo-Fernández, A., & Salvador-Cruz, J. (2015). Exploración de las propiedades psicométricas de la Batería Montreal de Evaluación de Amusia en una muestra de sujetos con epilepsia de lóbulo temporal. *Salud mental, 38*, 311-319.

Tranchant, P., Vuvan, D. T., & Peretz, I. (2016). Keeping the Beat: A Large Sample Study of Bouncing and Clapping to Music. *PLOS ONE*, *11*(7), e0160178-e0160178.

Uchanski, R. M. (2005). Clear speech. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 207-235). Oxford;Malden, MA;: Blackwell Pub.

van de Sandt-Koenderman, W. M., van Harskamp, F., Duivenvoorden, H. J., Remerie, S. C., van der Voort-Klees, Y. A., Wielaert, S. M., . . . Visch-Brink, E. G. (2008). MAAS (Multi-axial Aphasia System): realistic goal setting in aphasia rehabilitation. *International Journal of Rehabilitation Research, 31*(4), 314-320.

Van Engen, K. J., Chandrasekaran, B., & Smiljanic, R. (2012). Effects of speech clarity on recognition memory for spoken sentences. *PLOS ONE, 7*(9), e43753.

Van Lancker Sidtis, D. (2012). Formulaic language and language disorders. *Annual Review of Applied Linguistics, 32*, 62-80.

Van Vuuren, S., & Cherney, L. R. (2014). *A virtual therapist for speech and language therapy.* Paper presented at the International Conference on Intelligent Virtual Agents.

Varela Merino, E., Moíno Sánchez, P., & Jauralde Pou, P. (2005). *Manual de métrica española*. Madrid: Castalia.

Vaughan, H. G., & Ritter, W. (1970). The sources of auditory evoked responses recorded from the human scalp. *Electroencephalogr Clin Neurophysiol, 28*(4), 360-367.

Vergis, M. K., Ballard, K. J., Duffy, J. R., McNeil, M. R., Scholl, D., & Layfield, C. (2014). An acoustic measure of lexical stress differentiates aphasia and aphasia plus apraxia of speech after stroke. *Aphasiology, 28*(5), 554-575.

Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1471.

Wambaugh, J. (2003). A comparison of the relative effects of phonologic and semantic cueing treatments. *Aphasiology, 17*(5), 433-441.

Wambaugh, J., Duffy, J., McNeil, M., Robin, D., & Rogers, M. (2006). Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of the evidence. *Journal of Medical Speech-Language Pathology, 14*(2), xv-xv.

Wambaugh, J. L., Doyle, P. J., Martinez, A. L., & Kalinyak-Fliszar, M. (2002). Effects of two lexical retrieval cueing treatments on action naming in aphasia. *Journal of rehabilitation research and development*, *39*(4), 455.

Wambaugh, J. L., & Martinez, A. L. (2000). Effects of rate and rhythm control treatment on consonant production accuracy in apraxia of speech. *Aphasiology, 14*(8), 851-871.

Weissbart, H., Kandylaki, K. D., & Reichenbach, T. (2019). Cortical Tracking of Surprisal during Continuous Speech Comprehension. *Journal of Cognitive Neuroscience, 32*(1), 155-166.

White, L., Mattys, S. L., & Wiget, L. (2012). Language categorization by adults is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language, 66*(4), 665-679.

Wickham, H. (2010). ggplot2: An Implementation of the Grammar of Graphics. In (R package version 2.2.1 ed.).

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., . . . Hester, J. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686.

Wickham, H., Romain, F., Lionel, H., & Müller, K. (2020). dplyr: A Grammar of Data Manipulation. *R package version 0.8.5. https://CRAN.R-project.org/package=dplyr*.

Wiley, R. W., & Rapp, B. (2019). Statistical analysis in Small-N Designs: using linear mixed-effects modeling for evaluating intervention effectiveness. *Aphasiology*, *33*(1), 1-30.

Wingfield, A., & Grossman, M. (2006). Language and the aging brain: patterns of neural compensation revealed by functional brain imaging. *J Neurophysiol, 96*(6), 2830-2839.

Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology, 23*(11), 1338-1352.

Womelsdorf, T., Fries, P., Mitra, P. P., & Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature, 439*(7077), 733-736.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication, 20*(1), 1-28.

Youmans, G., Holland, A., Muñoz, M., & Bourgeois, M. (2005). Script training and automaticity in two individuals with aphasia. *Aphasiology, 19*(3-5), 435-450.

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proc Natl Acad Sci U S A, 107*(13), 5768-5773.

Zipse, L., Worek, A., Guarino, A. J., & Shattuck-Hufnagel, S. (2014). Tapped out: do people with aphasia have rhythm processing deficits? *Journal of Speech, Language, and Hearing Research, 57*(6), 2234-2245.

Zubizarreta, M. L. (1998). *Prosody, Focus, and Word Order*: MIT.

Zumbansen, A., Peretz, I., & Hébert, S. (2014a). The Combination of Rhythm and Pitch Can Account for the Beneficial Effect of Melodic Intonation Therapy on Connected Speech Improvements in Broca's Aphasia. *Frontiers in human neuroscience, 8*, 592.

Zumbansen, A., Peretz, I., & Hébert, S. (2014b). Melodic Intonation Therapy: Back to Basics for Future Research. *Front Neurol, 5*, 7.

Appendix A Complete set of sentences

Appendix Table 1 Complete set of sentences with linguistic properties and word stress

Set	Sentence	Number of words	Number of syllables	Spanish tense	Spanish mode	Mean Lexical Freq.	Word Stress and Syllabification
Practice set	Vaso de leche					4.16	**Va**so de **le**che
	Jugo de fresa					3.78	**Ju**go de **fre**sa
	Jarra de vino					4.08	**Ja**rra de **vi**no
set 1	Un director me ayudó	4	7	Pretérito perfecto simple	indicativo	4.49	Un direc**tor** me ayu**dó**
	Un pescador me auxilió					3.53	Un pesca**dor** me auxi**lió**
	Un bailarín me invitó					4.17	Un baila**rín** me invi**tó**
set 2	El gato está dormido	4	7	Presente	indicativo	4.46	El **ga**to es**tá** dor**mi**do
	El mico está cansado					4.00	El **mi**co es**tá** can**sa**do
	El pollo está corriendo					4.47	El **po**llo es**tá** co**rrien**do
set 3	Estos papeles son importantes	4	10	Presente	indicativo	3.85	**Es**tos pa**pe**les **son** impor**tan**tes
	Estas noticias son alarmantes					3.49	**Es**tas no**ti**cias **son** alar**man**tes
	Esas camisas son elegantes					3.43	**Es**as ca**mi**sas **son** ele**gan**tes
set 4	La niña juega en la tienda	6	8	Presente	indicativo	4.71	La **ni**ña **jue**ga en la **tien**da
	El niño salta en el patio					4.57	El **ni**ño **sal**ta en el **pa**tio
	El joven canta en el barrio					4.56	El **jo**ven **can**ta en el **ba**rrio
set 5	Estoy subiendo las escaleras	4	10	Presente Continuo - perífrasis verbal	gerundio	4.02	Es**toy** su**bien**do las esca**le**ras
	Estás llevando los alimentos					4.02	Es**tás** lle**van**do los ali**men**tos
	Están mirando los animales					4.29	Es**tán** mi**ran**do los ani**ma**les
set 6	Cuándo pagan la deuda?	4	7	Presente	indicativo	4.00	**Cuán**do **pa**gan la **deu**da?
	Dónde queda la finca?					4.18	**Dón**de **que**da la **fin**ca?
	Cómo tocas el piano?					4.13	**Có**mo **to**cas el **pia**no?
set 7	No pongas la mano en el fuego	7	9	Presente	subjuntivo	4.93	No **pon**gas la **ma**no en el **fue**go
	No tires la salsa en la mesa					4.67	No **ti**res la **sal**sa en la **me**sa
	No mires la tele en la noche					4.93	No **mi**res la **te**le en la **no**che
set 8	El bombero arrastraba la manguera	6	12	Pretérito imperfecto	indicativo	3.61	El bom**be**ro arras**tra**ba la man**gue**ra
	El portero entrenaba su mascota					3.58	El por**te**ro entre**na**ba su mas**co**ta
	El joyero arreglaba la pulsera					3.47	El jo**ye**ro arre**gla**ba la pul**se**ra
set 9	No le gusta la comida fría	6	10	Presente	indicativo	4.77	No le **gus**ta la co**mi**da **frí**a
	No me sirve la buseta grande					4.56	No me **sir**ve la bu**se**ta **gran**de
	No te queda la corbata negra					4.59	No te **que**da la cor**ba**ta **ne**gra
set 10	Pasaremos la noche aquí el viernes	6	1	Futuro	indicativo	4.59	Pasa**re**mos la **no**che a**quí** el **vier**nes
	Lavaremos el auto allá el jueves					4.08	Lava**re**mos el **au**to a**llá** el **jue**ves
	Pagaremos la cuenta ahí el martes					4.34	Paga**re**mos la **cuen**ta a**hí** el **mar**tes

The table above contains the complete set of sentences used in the short-term learning paradigm. Each sentence set was matched for the number of words, the number of syllables, syntactic structure, and roughly matched for phonetic complexity and lexical frequency.

Phonetic complexity was considered in terms of both syllabic structure and articulatory complexity of individual sounds. Lexical frequency was obtained from a Spanish database (Cuetos et al., 2011), and it reflects the mean lexical frequency across words in the sentence. The last column (word stress and Spanish syllabification) depicts the word rhythm of each set of sentences.

As the alternating grey (stressed) and white (unstressed) syllables show, each of the sentences in the same set have the same rhythmic pattern by maintaining the precise sequence of alternating unstressed and stressed syllables. Numbers of syllables per sentence are counted according to Spanish poetic metrics (Varela et al., 2005) which considers oral resyllabification and the stress of the last word. This method aligns better with rhythmic perception.