Linguistic Predictors of Anomia Treatment Outcomes

Reva M. Zimmerman

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington
2020

Reading Committee:
Diane L. Kendall, Chair
Carolyn Baylor
Michael Walsh Dickey
JoAnn P. Silkes

Program Authorized to Offer Degree:
Speech and Hearing Sciences

University of Washington

Abstract

Linguistic Predictors of Anomia Treatment Outcomes

Reva M. Zimmerman

Chair of the Supervisory Committee:

Diane L. Kendall

Speech and Hearing Sciences

Many people with anomia (PWA), or word-finding difficulties, seek treatment to ameliorate their deficits; however, it is unclear why some PWA experience significant acquisition, generalization, and maintenance of the skills they learn in treatment, while others experience very limited effects. The purpose of this retrospective study was to determine the predictive value of specific pre-treatment linguistic skills to acquisition, generalization, and maintenance of anomia treatment skills in 58 PWA who were randomized to one of two intensive anomia interventions: Phonomotor Treatment (PMT, n = 28) or Semantic Feature Analysis (SFA, n = 30). Mixed-effect logistic regressions were used to determine whether participants' baseline scores on measures of auditory comprehension, input phonological processing, input-output phonological processing, and semantic and phonological impairment levels predicted naming of items representing acquisition and generalization immediately and three months post-treatment. Higher baseline input-output phonological processing was the most robust predictor of greater acquisition, generalization and maintenance for all 58 PWA combined, the PMT group, and the SFA group. Semantic and phonological impairment also generally predicted outcomes in the combined and PMT groups, such that greater baseline impairment related to higher acquisition, generalization, and maintenance. These results suggest that input-output processing – and the skills that underlie it – may be critical for favorable response to anomia treatment in general. Furthermore, PMT may be a superior treatment for people who are more impaired prior to treatment than SFA.

Table of Contents

List of Figures and Tables

Tables

Figures

Dedication

To Dida for sustaining me body and soul,
Nini for being loving and patient beyond her years, and
Meeka for keeping me company during dark times.
To Bae for keeping it REAL REAL,
to my Play-mom for reminding me to take care of myself,
and to my Dad for supporting me, even though he still has no idea what I do.
To my stans worldwide.

To my PhD "Saturday" study group. We got this, y'all.

Acknowledgements

I acknowledge that this work was carried out on the traditional land of the first people of Seattle, the Duwamish People – past, present, and future – and honor with gratitude the land itself and the Duwamish Tribe.

I deeply appreciate Dr. Diane Kendall, Dr. Stephen Nadeau, Megan Moldestad, Wesley Allen, Kerry Lam, Janaki Torrence, and all the other Merit 2 team members whose hard work made this study possible, and the Veteran's Health Administration (VA RR&D C6572R) for funding the main study.

I thank Dr. Diane Kendall, Dr. Carolyn Baylor, Dr. Michael Walsh Dickey, and Dr. JoAnn Silkes for their guidance, support, and mentorship. Many thanks to Dr. Malcolm McNeil, Dr. Nadine Martin, and other experts in the field for their encouragement and mentorship. And of course, thank you to the members of the UW Aphasia Research Laboratory, past and present, and Dr. Michael Burns for their friendship and insights.

I am indebted to Dr. Kevin Cain and James Proudfoot for their statistical expertise and patience with me while I cried through my analyses. (Special thank you to Madeline, Clara, and June.)

I am grateful to the University of Washington and the Speech and Hearing Sciences Department for their support, and to the many faculty, staff, and students who make this place special. I am especially thankful for the support provided by the University of Washington Retirement Association, who awarded me the Charles Chamberlain Scholarship in Aging.

Finally, I have unending gratitude for the participants and families who gave their time, energy, and expertise to help make Phonomotor Treatment and the UW Aphasia Lab the best possible versions of themselves.

Thank you,
Reva M. Zimmerman

Chapter 1: Introduction

Aphasia is a breakdown in normal language processing that can result from damage to the language centers of the brain, typically left-hemisphere stroke. Following the onset of aphasia, people with this disorder may experience any combination of impairments in oral language, written expression, auditory comprehension, and/or reading comprehension. Furthermore, some people with aphasia (PWA) demonstrate non-fluent and/or agrammatic language, while others exhibit fluent and/or meaningless speech (Lichtheim, 1885; Schuell, 1960). Whatever constellation of deficits people present with, an endemic feature of aphasia is *anomia*, or an inability to find the correct word when it is needed (Goodglass & Wingfield, 1997). Aphasia with anomia may impact PWA's ability to participate in vocational, social, and leisure activities, which can have devastating socio-emotional consequences: PWA may not be able to return to work, even when physical sequelae of the brain injury may have resolved; they may have trouble communicating with people in their communities and loved ones; and they may not be able to partake in hobbies such as reading or watching movies (Wallace et al., 2017). Limited participation in everyday life can negatively influence how PWA see themselves, leading to withdrawal, feelings of isolation, and depression (Døli, Helland, & Andersen Helland, 2017; Hadidi, Treat-Jacobson, & Lindquist, 2009; Morris, Eccles, Ryan, & Kneebone, 2017; Shehata, El Mistikawi, Risha, & Hassan, 2015). Far from being limited to those with the disorder, the effects of aphasia can also negatively impact the friends, families, and caregivers of PWA, leading to "third-party disability" characterized by feelings of isolation, depression, and caregiver fatigue/burnout (Davidson, Howe, Worrall, Hickson, & Togher, 2008; Grawburg, Howe, Worrall, & Scarinci, 2013; Threats, 2010).

Given the far-reaching effects of aphasia, PWA and their loved ones often seek treatment from speech-language pathologists (SLPs) to ameliorate their language deficits. Because anomia

is the most ubiquitous and, arguably, the most easily quantified aspect of aphasia, it is often the target of aphasia intervention. Anomia treatment approaches are highly variable, ranging from semantic (meaning-based) or phonological (sound-based) cueing hierarchies, to "language network strengthening," to teaching self-cueing strategies. Numerous anomia treatments are demonstrated to be efficacious on the whole (Nickels, 2002; Robey, 1994, 1998), as individuals have demonstrated statistically or clinically significant changes in (1) naming ability (Wisenburn & Mahoney, 2009), (2) reading abilities (C. E. Brookshire, Conway, Pompon, Oelke, & Kendall, 2014; Raymer, Thompson, Jacobs, & Legrand, 1993), (3) discourse abilities (Silkes, Fergadiotis, Hunting Pompon, Torrence, & Kendall, 2019), and (4) quality of life (Kendall, Moldestad, Allen, Torrence, & Nadeau, 2019).

Although therapy can yield improvements overall, when scrutinized in any detail, it quickly becomes apparent that some PWA are treatment "responders" and others are "non-responders." Responders typically achieve statistically significant acquisition of treatment effects over baseline, sometimes demonstrate generalization to untrained items, and may demonstrate maintenance of acquisition and/or generalization at a follow-up point. Non-responders, on the other hand, may show little or no change over baseline for trained and untrained items, and they may not maintain whatever minimal gains they do experience. For non-responders, the therapeutic process can feel frustrating or futile, with the potential to exacerbate negative emotions associated with aphasia (Fromm et al., 2011; Sinyor et al., 1986). For researchers attempting to promote language improvements in PWA, participants who demonstrate limited gains can inspire dissatisfaction on the participants' behalf, as well as pose theoretical or clinical quandaries about how best to treat aphasia. Thus, many aphasia researchers have sought to

answer the question, *why do some PWA respond well to treatment while others do not seem to respond at all?*

Predictors of Treatment Response

The search for the source of patient variability has led many aphasiologists to consider several variables that could predict treatment response. Predictors of treatment response from the past 40 years or so have included such variables as age at time of stroke, months-post onset (MPO) of aphasia, lesion characteristics (i.e., site and size), aphasia type, and aphasia severity prior to treatment. Age, MPO, and aphasia type have not been consistent, significant predictors in treatment of chronic (>6 months post-onset) aphasia; however, lesion characteristics and aphasia severity have been associated with degree of improvement experienced following treatment. Specifically, aphasia severity has been the most robust predictor of treatment response, regardless of the depth and breadth of the tools used to measure severity (Kiran, 2016; Quique, Evans, & Dickey, 2018; Robey, 1998). While this finding is interesting, because it speaks to a linguistically based mechanism of language change during treatment (discussed below), it is also frustratingly vague, because it gives no indication of *which* mechanism(s) might underlie treatment-driven change.

Measures of aphasia severity have varied across the predictor literature, with early work relying on tools such as the National Institute of Health Stroke Scale (National Institute of Neurological Disorders and Stroke, 2011), which uses a 0-4 rating scale, with 0 representing "no aphasia" and 4 representing "severe aphasia." With the introduction and adoption of the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1972) and the Western Aphasia Battery (WAB; Kertesz, 1982), aphasia assessments became more comprehensive, but often aphasia severity was still reported as a single score, with little insight into which specific aspects of

language were more or less affected. For example, one person with relatively spared comprehension ability but poor expression could have the same score as another person who has generally equal impairment across domains. These profiles are qualitatively different, yet the score does not reflect that.

To complicate the interpretability of omnibus severity scores, comprehensive aphasia tests often weight linguistic modalities differently, such that a "mild" impairment in one domain contributes to the overall score more than or as much as a "moderate" impairment in another area. For example, recent work suggests that the greatest proportion of the overall aphasia score, or Aphasia Quotient (AQ) on the WAB-Revised (Kertesz, 2007) derives from the Spontaneous Speech subtest (i.e., responses to egocentric questions and a picture description task; ~30%), followed by Repetition and Word Finding (both ~25%), then Auditory Verbal Comprehension (~20%; Ellis, Peach, & Rothermich, 2019). From a theoretical perspective, it is unclear why any of these domains should contribute more or less than other domains to language processing. From a functional perspective, one would think that auditory comprehension might be more meaningful than repetition. Owing to the obscurity of individual subtest scores and the possibility that various skills may differentially contribute to the overall score, it seems that reliance on an omnibus index of aphasia severity must be only a first step in understanding response to aphasia and anomia treatment.

Theoretical Underpinnings of Normal and Aphasic Language Performance

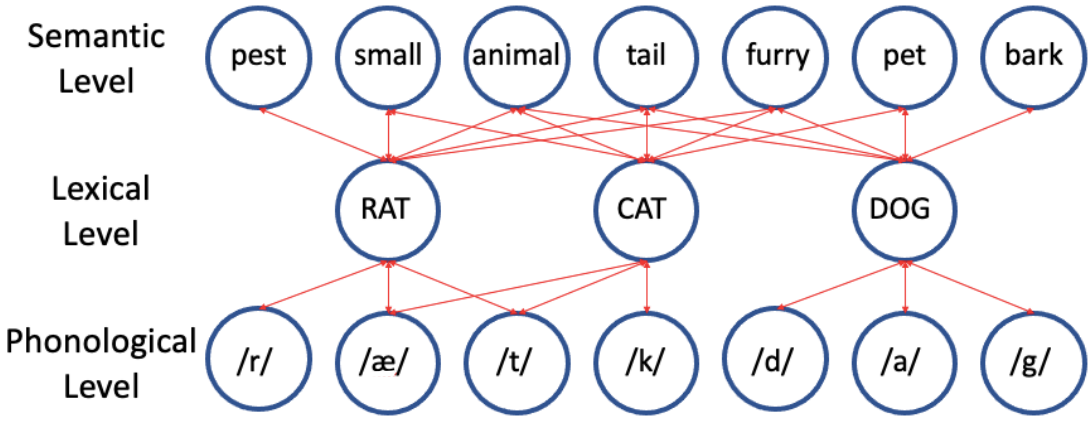
Psycholinguistic theory can clarify how language ability can be decomposed into sub-skills and why they would influence treatment. Although many theories of word and sentence production are available to explain aphasic language behaviors, a popular and useful model is the interactive activation theory of word production (henceforth, IA model; Dell, 1986; Dell &

O'Seaghdha, 1992). The IA model was initially proposed to explain speech errors produced by normal (i.e., non-brain injured) speakers. However, the model's ability to account for a large variety of errors lent itself to accounting for behaviors observed in aphasic speakers, and it was soon adopted for aphasiology (Dell, Lawler, Harris, & Gordon, 2004; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Foygel & Dell, 2000; Jokel, Rochon, & Leonard, 2004; Martin & Saffran, 1992, 1997).

The Interactive Activation Model of Language Production. The essential premise of the IA model is that, in a production task, activation is introduced into the "uppermost" layer of the network (semantics), then cascades to each of the two "lower" levels (lexical and phonological) via bidirectional links (see Figure 1). At each level, nodes exist as linguistic representations: semantic nodes encode features of every concept known to the speaker; every lexeme at the lexical level receives links from all of the associated features at the semantic level to encapsulate them into a single entity. The more conceptually rich a lexeme is, the more links it receives from the semantic level. For example, words associated with objects tend to have more semantic features than abstract words because objects not only contain semantic information related to definition (e.g., *cat* might be defined as *feline*, *companion*, *housebound*, *friendly*), but

Figure 1

Interactive activation model of language (adapted from Dell & O'Seaghdha, 1991)



also semantic information related to their physical characteristics (e.g., *cat* has *four legs*, *fur*, *tail*). The degree to which a lexeme can be associated with physical characteristics is known as its imageability.

Returning to network structure, lexemes not only receive information from and encapsulate semantic features, but they also provide instructions for morphosyntactic and phonological encoding. The phonological level contains phonemes for all of the sounds of the speaker's language, and each segment is linked to every lexeme that contains it. The links between levels develop and strengthen through experience: when nodes at different levels activate at the same time, they form links to each other; repeated co-activation strengthens the link (Dell, 1986; Hebb, 1949; Nadeau, 2001; Plaut, 1996). That connection strength, or weight, is experience-dependent and has important implications for both word learning and producing words of high or low frequency. Specifically, words that are encountered more often in everyday language have stronger connection weights between levels, meaning that they are accessed more quickly, and connections are more resistant to degradation (discussed below).

When a speaker attempts to produce a word, activation is introduced into the system at the semantic level, and then cascades downward through the network via the links. Although activation spreads through every available link, the target nodes receive the highest level of activation, because they receive activation from more nodes than nontarget items (resulting in a higher summation of the signal) and they have stronger links to the source nodes because of their frequent co-activation. Because activation decays over time and distance, the reverberation of activation through the system ensures that the target nodes maintain the highest activation level for the longest period of time. This mechanism allows time for the assembly of the target word for production.

The basic model described above was originally intended for production only (Dell, 1986). Generally speaking, it seems reasonable to assume that the network structure for comprehension is essentially the same as the production network, except that activation begins at the phonological level and spreads upward in the comprehension network (Gambi & Pickering, 2017; N. Martin & Saffran, 2002; R. C. Martin, Lesch, & Bartha, 1999). The larger issue has been in determining the relationship between comprehension and production: Does one network serve both processes, i.e., are the comprehension and production networks inseparable? Alternatively, does each process have its own network, such that the networks are completely separate? If the networks are separate, how and when do they communicate to produce the behavioral and imaging observations seen across the literature? Or is the best model perhaps situated somewhere between these extremes, where some levels of the network are shared, but others are separable?

A fair amount of behavioral, computational, and neuroimaging research favors the separable view (Dell & Chang, 2014; Heim, Opitz, Müller, & Friederici, 2003; Hickok, Houde, & Rong, 2011; Kittredge & Dell, 2016; N. Martin & Saffran, 2002; Nozari, Kittredge, Dell, & Schwartz, 2010; Walker & Hickok, 2015). Through analysis of performance on comprehension and production tasks, and the relationships between these tasks, behavioral and imaging studies generally suggest "representational parity" – or shared representations – for lexical and semantic network levels in comprehension and production tasks (Bock, Dell, Chang, & Onishi, 2007; Brennan & Clark, 1996; Gambi & Pickering, 2017; Heim et al., 2003; Pickering & Garrod, 2013; Schriefers, Meyer, & Levelt, 1990; Segaert, Menenti, Weber, Petersson, & Hagoort, 2012). However, representational parity does not seem to apply to phonology. Although some studies have suggested that phonological representations are shared (e.g., Fadiga, Craighero, Buccino, &

Rizzolatti, 2002; Pulvermuller et al., 2006), many behavioral, computational, and neuroimaging studies have found dissociations between "input" and "output" phonology. Namely, evidence of separate input and output buffers includes differential performance on production and repetition tasks (especially in PWA), divergent patterns of cortical activation during production versus comprehension tasks, and relatively successful computational network instantiations to model behavioral data (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Gambi & Pickering, 2017; Kittredge & Dell, 2016; Nozari et al., 2010; Walker & Hickok, 2015). Neuroanatomically-based processing models suggest both functional and spatial separations between input (in neurological terms: perceptual) and output (motor) phonological processing, and these models lend support for imaging findings (Hickok & Poeppel, 2004; Nadeau, 2001). Importantly, most of the investigations of phonological processing strongly support the notion of input and output buffers that are individually linked to the lexical level *and* to each other.

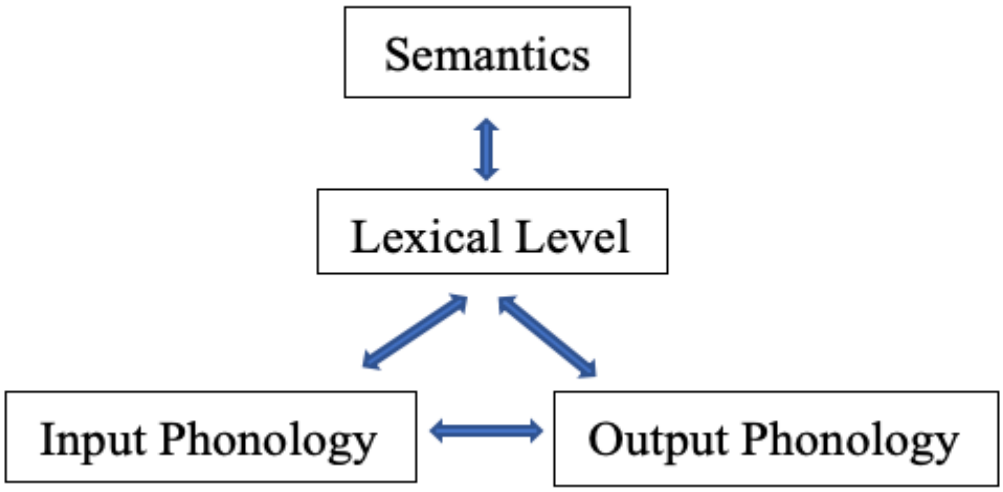
With specific regard to behavior, research clearly underscores the role of perception in learning to produce language. Specifically, typically developing children who have exposure to rich linguistic environments have better language abilities, while children who experience limited or degraded linguistic input – due to hearing loss, for example – often demonstrate poorer language abilities (Gervain & Mehler, 2010; Ingram, 1989; Jusczyk, 1997; Menyuk, Menu, & Silber, 1986). In neurologically healthy adults, participants learned synthetic phonotactic constraints during error monitoring tasks but not during passive listening tasks. When participants simply heard a string of nonsense syllables, they did not demonstrate evidence of the synthetic phonotactic constraints on lab-induced slips of the tongue. However, when participants monitored for correct production of a nonsense syllable, they acquired the phonological "rules" for these syllables, as evidenced through the types of errors they committed in their own

productions of these syllables (Kittredge & Dell, 2016; Warker, Xu, Dell, & Fisher, 2009). The authors of these studies took these findings as evidence for separate, but related, phonological networks.

In aphasia, people often experience more impairment in one domain than another (which supports a separate view); however, if one system is impaired to any degree, then the other system is highly likely to also be impaired (supporting a inseparable view; Butterworth, Howard, & McLoughlin, 1984; Goodglass & Baker, 1976; N. Martin & Saffran, 2002). In a study by Dell et al. (1997), participants demonstrated (1) better repetition than naming performance and (2) different error patterns in repetition than in naming, forcing the authors to concede that a single phonological network for both input and output phonology could not account for their findings. Later, Martin and Saffran (2002) specifically investigated the relationship between input and output phonology and found that although there was not a perfect, 1:1 correspondence between number of errors on input and output tasks, the relationship between these behaviors was significant. These findings led them to conclude that input and output phonology are at least "functionally linked" (p. 143), but that their results do not support a single network view.

Thus, following compelling evidence across different populations, this study relies on a linguistic network capable of accounting for comprehension and production abilities that features unitary semantic and lexical layers, and separate but linked input and output phonological networks as in Figure 2, which is adapted from Nozari et al. (2010). In this network structure, an example of single word production is a confrontation naming task, wherein a person is shown a picture of an object and asked to name it. If the picture is of a cat, then features such as *furry*, *pet*, *quadruped*, *small*, *friendly* might become activated at the semantic level, and that activation

Figure 2
Basic Structure of IA Model of Production and Comprehension



would pass to all of the linked nodes at the lexical level. Thus, the lexemes *cat*, *dog*, *lion*, *parrot*, and *Casper* may become activated, which would then activate the phonemes /k/, /æ/, /t/, /d/, /ɑ/, /g/, /l/, and so on, at the phonological level. The lexeme *cat* would have received the most activation from the semantic level and would have passed that activation on to the segments /k/, /æ/, /t/; subsequent waves of activation through the network would serve to reinforce the high activation level. When the examinee is ready to say the target word, the word /kæt/ is produced because it is the most highly activated. Conversely, if a person is instructed to point to one of four pictures that matches the spoken word *cat*, then the phonemes /k/, /æ/, /t/ become activated, and pass that activation on to any lexemes that contain those segments: *cat, Casper*, *can*, *hat*, etc. Those lexemes pass activation on to the semantic level to activate any nodes associated with those lexemes, such as *furry*, *pet*, *quadruped*, *small*, *friendly, clothing, metal*, etc. The examinee would then point to the picture that corresponds to the most highly activated nodes.

The IA Model Account of Aphasia. In the IA framework, aphasia occurs when the strength of the connections between semantics and lexemes (lexical-semantic, or simply *semantic*, weights), the connections between lexemes and phonemes (lexical-phonological, or

phonological, weights), the connections between input and output phonology (input-output weights) are degraded, or activation decays too quickly, such that activation no longer transmits between levels easily (Foygel & Dell, 2000; Nozari & Dell, 2013; Nozari, Kittredge, Dell, & Schwartz, 2010). Using this framework as a guide, researchers have sought to describe the types of errors committed by PWA, and the model has done well in accounting for most aphasic errors (Dell et al., 1997; Dell, Schwartz, Nozari, Faseyitan, & Coslett, 2013; Foygel & Dell, 2000; N. Martin & Saffran, 1997; Nozari et al., 2010).

In word production tasks, activation transmission issues arise when (1) target nodes do not achieve sufficient activation levels for selection, (2) non-target nodes achieve higher activation levels than the target nodes, or (3) no nodes reach sufficient levels to be selected. Thus, in the case of semantic weight lesions, PWA may produce semantically related (e.g., *cat* → *dog*), or semantically unrelated (*cat* → *sofa*) word errors. These types of errors are the result of poor activation spread from semantics to lexemes because decreasing activation transmission increases the likelihood of random lexical selection. On the other hand, phonological weight lesions result in phonologically related (*cat* → *cap*) and nonword errors (*cat* → *flod*), since phonological encoding of the target lexeme becomes increasingly random as the lesion severity increases. Disturbances in both semantic and phonological weights result in mixed semantic-phonologic errors (*cat* → *cod*). Finally, at the extreme end of the continuum, omissions (or non-responses) result from severe transmission impairments in semantic, phonological, or both sets of connection weights.

In single-word comprehension tasks, a phonological error may occur where a subject hears the word *cat* but points to a picture of a cab, and this type of error is attributable to phonologic lesions, wherein the target phonemes become activated but cannot signal the correct

lexeme. Thus, any competing item has the same likelihood of being selected as the target. The same activation transmission problem may be true for semantic errors, except that the erroneously selected item shares features or attributes with the target (e.g., *cat* → *dog*). Mixed errors (e.g., *cat* → *rat*) arise from lesions to semantic and phonological weights, although it is unclear whether one set of weights or the other must be more impaired. Finally, non-responses result from severe lesions to either or both sets of weights.

Spoken word repetition is thought to rely on input and output phonological processing, with varying degrees of support from lexical-semantics (N. Martin & Saffran, 2002; Nozari et al., 2010). At a basic level, an auditorily-presented stimulus activates input phonology, which then sends activation up to lexical-semantics, as well as to output phonology. In this case, lexical-semantic activation serves to verify and/or support spoken reproduction of the stimulus. The influence of semantic processing is evidenced through lexicalization of nonwords (*cag* → *cat*) or substitutions of low frequency words with high frequency words (*agonist* → *agony*). Lexicalization and high frequency word substitutions suggest lesions in the connections between input and output phonology, because activation cannot transmit from input to output, and thus activation to output phonology relies on the signal it receives from lexical-semantics. Specifically, nonwords will activate whichever lexical node most closely matches the phonological profile of the nonword stimulus; high frequency lexemes have stronger links to phonemes and semantic features and thus receive more activation than low frequency lexemes, especially in a compromised system (Nozari & Dell, 2013).

In summary, the IA model can be, and has been, deployed to explain a range of phenomena in normal and aphasic language. The bulk of the research concerning the model has revolved around computational simulations and empirical studies of single-word production and

repetition, but studies to understand comprehension impairments in PWA also exist (e.g., Martin, Schwartz, & Kohen, 2006). In essence, many behaviors observed in aphasia are attributable to disruptions in connections between semantics and the lexical level (or semantic weight lesions), the lexical level and phonology (phonological weight lesions), input phonology and output phonology (input-output weight lesions), or some combination of these. The complex mechanisms of aphasic language impairment underscore the need for nuanced analyses of linguistic predictors of treatment response. Understanding the relationship between baseline language characteristics and treatment-induced language change can motivate the selection of appropriate tests of language performance and, thus, maximize patient gains.

The IA Model Account of Therapy-Induced Language Change

Within the context of the IA model, aphasic language behaviors arise from activation transmission impairments. Thus, treatments rooted in this model often focus on repairing or restoring connection weights between levels through Hebbian learning (Hebb, 1949); that is, by repeatedly and simultaneously activating related representations. Using a basic naming paradigm as an example, looking at an image of an item while naming it out loud activates all three levels of the language network: the image activates semantic representations, which send activation to converge on the lexical level, which in turn activates associated phonemes in output phonology. Naming this object over and over strengthens the links between each level *for this item*. In the event that the patient cannot name the item independently, any support that the clinician provides, such as cues or modeling, still serve to strengthen connections over time; however, because the patient receives input in this scenario, input phonology will also receive an extra boost because it is specifically being engaged in the training. Likewise, when a patient must

repeat an item, associations between input and output phonological buffers improve because that processing route is being activated.

The notion that recurrent co-activation is the key mechanism of change in aphasia rehabilitation stems from Hebb's (1949) seminal neuropsychological theories on behavior and learning, and the principle was refreshed in Kleim and Jones' (2008) review of treatment induced neuroplasticity. Ultimately, targeted co-activation of multiple representations, in any modality, strengthens the links between those representations, such that access becomes easier over time, with a caveat being that those with greater network impairment will require more training than those with less network impairment to improve. Nevertheless, this relearning of trained items constitutes treatment acquisition.

Because every word in a language cannot be trained during therapy, clinicians must design their treatments to promote generalization; that is, patients must learn to perform the trained language behavior in novel contexts and with novel stimuli. Across the literature, generalization has been less robust than acquisition (Best et al., 2013; Patterson & Coppens, 2018; Staines, 2008), and the IA model provides insight for this phenomenon, too. As discussed, the links between representations for trained items benefit from active engagement. Generalization occurs when activation from the targeted items spreads to non-target representations through the feedforward and feedback mechanisms described above. Thus, linguistic representations that are adjacent to trained items receive passive activation. For this reason, some researchers have suggested that clinicians use atypical training stimuli: Words that capture the general characteristics of a category, but that also demonstrate the diversity within a category, have been found to result in greater generalization (Kiran & Thompson, 2003; Plaut, 1996; Storkel, Armbruster, & Hogan, 2006; Thompson, Shapiro, Kiran, & Sobecks, 2003).

Although this spreading activation mechanism of generalization appears straightforward, generalization to non-target items has been limited, and that may be due to a few important features of the network. First, the nature of semantic features is such that (1) the "reach" of any given semantic node is limited to a finite set of lexemes (Kendall, Rosenbek, Heilman, Conway, Klenberg, Rothi, et al., 2008), and (2) the semantic associations for any given lexeme vary across individuals (Armstrong & Ferguson, 2010). Thus, a treating clinician would have to somehow determine which semantic features their client associates with which lexemes, and then would have to train a large number of semantic features to see large-scale generalization to untreated items. A second implication of the model's generalization mechanism is that when links are impaired between levels, activation does not spread as far as it does in an intact network, and in severe cases, it may not spread at all. As a consequence, patients can achieve high degrees of acquisition, but they may experience limited generalization if activation cannot reach neighboring nodes.

A final caveat of this generalization mechanism is that if a trained word does not have many neighbors (i.e., it is a "hermit"; Castro, Stella, & Siew, 2020), activation may not be able to spread to many related representations. For example, *platypus* is rare in both its phoneme sequences and its semantic network; thus, it may not be an ideal stimulus if the goal is large-scale generalization to mammals. On its surface, this problem seems to contradict findings on the positive influence of semantic and syntactic complexity on treatment gains (Kiran & Thompson, 2003; Storkel, Armbrüster, & Hogan, 2006; Thompson et al., 2003). Rather than suggest that atypical or complex stimuli should be avoided, the point here is to suggest that a distinction between atypical exemplars and hermits seems to exist, such that these hermit stimuli may be *unhelpfully* distant from other items (Castro, Nadeau, & Kendall, manuscript in review; Plaut,

1996). The point at which atypical items become hermits is yet unknown, although Plaut (1996) recommends that training stimuli "must collectively cover the range of variation in the category while still being balanced around its central tendency" (p. 61). Finally, evidence is emerging to suggest that phonological processing may be more susceptible to hermit effects than semantic processing (Castro et al., 2020).

To recapitulate, treated words are acquired through Hebbian learning (repeated co-activation), and access to untreated words improves through spreading activation to neighboring representations. The IA model predicts that each time representations are co-activated, the links between them are strengthened; strengthening access to trained representations results in acquisition of treatment effects, while improved access to untrained representations reflects generalization. The issue of maintenance is more complicated, however, because according to the model, connection weights should not spontaneously become weaker and decay should not suddenly become overly rapid. Thus, according to the IA framework, once acquisition and generalization occur, such improvements are automatically maintained. However, the vast literature examining acquisition and maintenance of post-treatment effects suggests that decline does indeed occur (Sze et al., 2020), and this phenomenon is consistent with the principles of neuroplasticity (i.e., that representations must be used consistently to prevent degradation; Kleim & Jones, 2008; Raymer et al., 2008). Given that the IA model is agnostic to the undeniable reality of limited post-treatment maintenance, other connectionist models may be instructive in this regard: Hinton and Plaut (1987) and Plaut (1996) discussed the concept of "fast" and "slow" weights in their instantiations of parallel distributed processing (PDP) models. In their model, fast weights enable rapid acquisition of a new skill (e.g., word learning), while slow weights are responsible for transferring this new knowledge to long-term memory. For learning to transfer

from fast to slow weights, training must reach some threshold, and once that threshold is reached, new knowledge may be maintained. The concept of fast and slow weights certainly has some caveats – for example, while researchers can easily identify the transfer threshold in computational instantiations (e.g., 10,000 training epochs), it is much more difficult to define this transfer point in neurobiological terms. Nevertheless, fast and slow weights are the best explanation for observations from behavioral and neurobiological studies of maintenance; thus, this concept is useful in predicting and interpreting results in the present study. For example, because treatment stimuli, by definition, get closer to the transfer threshold than untreated items, treated words are more resistant to attrition than untreated words; more precipitous loss of generalization should occur because the untreated words were only passively stimulated. This is precisely what is observed in the literature (Sze et al., 2020).

The IA model predicts acquisition, generalization, and maintenance in general. However, the specter of outcome variability rears its head again when researchers and clinicians evaluate individual results. Individual variability forces aphasiologists to question precisely how the IA model predicts changes to connection weights, *given a particular arrangement of linguistic deficits*. For example, do those with more severe phonological or semantic impairment benefit more from treatment? Does comprehension or production matter more? Does a certain deficit need to be matched to a certain treatment approach? Dell et al. (1997) and Foygel and Dell (2001) provide simulations that aid in our understanding of how and why certain error types arise and suggest that "recovery entails resolution of abnormal [weights] toward the normal state" (Dell et al., 1997, p. 824). Unfortunately, the optimal approach to resolve abnormal weights is left open to interpretation.

Some researchers have attempted to uncover linguistic modalities (or macro-linguistic skills) as predictors of outcomes for a variety of therapy approaches. As Raymer et al. (1993) noted in their small prospective study of the effect of a phonologically based treatment, two of four participants with strong pre-treatment oral reading abilities demonstrated a large gain in their post-treatment oral reading performance. The authors interpreted this finding as evidence that phonological processing had improved in PWA who already demonstrated strong phonological skills prior to the intervention. In a meta-analysis of response predictors in a sentence production treatment, Dickey and Yoo (2010) found that strong auditory comprehension scores significantly predicted acquisition, but not overall aphasia severity or complex sentence comprehension; no measure predicted generalization. The authors suggested that auditory comprehension may be important for following task instructions in the sentence treatment protocol specifically, and any intervention type in general. In a retrospective analysis of the effects of a cueing hierarchy treatment for anomia, Eoute (2010) found that higher baseline naming ability significantly predicted acquisition of items, but not generalization. Finally, Dignam and her colleagues (2017) found that high auditory and written comprehension scores were significantly related to acquisition and maintenance of treatment effects. This finding is in line with Dickey and Yoo's earlier suggestion that good comprehension ability prior to treatment may be foundational to treatment success. It is noteworthy that all of these macro-linguistic predictors found significant acquisition effects, but not generalization effects. This trend is likely because these measures reflect the integrity of connections between the target items, but do not necessary rely upon the spread of activation beyond target nodes.

If "macro-linguistic" refers to broader skills such as verbal comprehension or naming ability, then "micro-linguistic" should refer to the processing components that enable these skills.

Only two studies were found that specifically probed semantic and phonological contributions to therapy. In a secondary analysis of his data, Eoute (2010) used the number of phonological and semantic errors on the Philadelphia Naming Test (PNT; Roach, Schwartz, Martin, Grewal, & Brecher, 1996) as treatment response predictors in a cueing hierarchy paradigm. He found that baseline phonological skills positively related to increased generalization, while semantic errors were not related to acquisition, generalization, or maintenance. Conversely, Best et al. (2013) found that stronger semantic ability (as measured by auditory comprehension), coupled with poorer phonological ability (as measured by error proportions on a naming task), was significantly related to greater generalization to untreated items following a phonological/orthographic cueing therapy for anomia. The authors of this study concluded that good semantic processing supported strengthening of phonological skills, and that intact lexical-semantic network functioning allowed activation from the phonological level to spread further through the network.

Taken together, these studies seem to imply that good comprehension and semantic abilities are important for achieving acquisition, generalization, and maintenance. However, the evidence is far from conclusive. In fact, Eoute's study seems to suggest that *production and phonological* abilities are most important. That this study stands in opposition to the other findings could simply reflect some methodological differences, but other researchers have found that more severe baseline language is related to greater gains (Kiran, 2016), and still others have found no relationship at all (Babbitt, Worrall, & Cherney, 2016; Harnish & Lundine, 2015; Hunting Pompon et al., 2017). The reason for these discrepancies is unclear, and more than anything, highlights the need for thorough, theoretically based analyses to clarify the contributions of different aspects of baseline language processing to anomia treatment success.

Current Study

The purpose of the current study is to fill a gap in the literature: Namely, aphasiologists have yet to understand the baseline linguistic characteristics that drive treatment response. Using a retrospective analysis of a randomized control trial comparing phonologically and semantically based treatments, the goal of this study was to provide clarity on which macro- and micro-linguistic skills relate to anomia treatment success. This study had several unique features. First, it used data from the largest trial to date of Phonomotor Treatment (PMT), combined with a large sample of participants in the Semantic Feature Analysis (SFA) condition, for a total of 58 PWA. This large of a sample, and for this type of analysis, is very rare in aphasiology. Second, the treatments each focused on different levels of the linguistic network: SFA targeted lexical-semantic processing, while PMT trained lexical-phonological connections. Despite the different approaches, both protocols were administered with the same intensity and frequency, and in the same conditions (i.e., settings and clinicians). Thus, these data were available for analysis both within groups, to determine whether different underlying skills corresponded to response to semantic and phonological treatment approaches respectively, and across groups, to determine whether certain skills related to response to treatment regardless of approach. Third, this study modeled a variety of language behaviors – auditory comprehension, lexical-phonological input, phonological input-output processing, and micro-linguistic skills – as predictors of treatment response. Fourth, the contribution of the predictors was assessed for acquisition, generalization to items related to the treatment stimuli, and generalization to items unrelated to the treatment stimuli, as well as maintenance for each of these measures. Fifth, this study used mixed-effects logistic regression models, which maximized power and were not subject to some of the assumptions of other analysis approaches (e.g., normality).

Research Questions. The specific research questions for this study were as follows.

1. *What is the relative contribution of baseline macro-linguistic input and input-output abilities, and micro-linguistic abilities, to lexical retrieval abilities immediately post-treatment in a group of 58 PWA who participated in an intensive anomia treatment?*
2. *What is the relative contribution of baseline macro-linguistic input and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities 3 months post-treatment in a group of 58 PWA who participated in an intensive anomia treatment?*
3. *What is the relative contribution of baseline macro-linguistic input and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities immediately post-treatment in a group of 28 PWA who were randomized to receive a phonologically based (PMT) treatment?*
4. *What is the relative contribution of baseline macro-linguistic input and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities 3 months post-treatment in a group of 28 PWA who were randomized to receive a phonologically based (PMT) treatment?*
5. *What is the relative contribution of baseline macro-linguistic input and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities immediately post-treatment in a group of 30 PWA who were randomized to receive a semantically based (SFA) treatment?*
6. *What is the relative contribution of baseline macro-linguistic input and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities 3 months post-*

<u>treatment</u> in a group of 30 PWA who were randomized to receive a semantically based (SFA) treatment?

Each of the individual treatment groups was evaluated in this study to determine whether specific baseline skills related to outcomes, as mediated by elements. For example, because PMT entails manipulation of phonemes and phoneme sequences (see *Methods – Treatment* below), input-output phonology – which entails many of the same skills as the treatment – was predicted to significantly relate to outcomes. However, the same types of relationships cannot exist between predictors and outcomes in the whole group, since this group is composed of participants from both groups. The rationale for conducting the analyses for the whole group is manifold. The first justification for the whole group analyses is that the may offer insights into important predictors of post-treatment outcomes *in general*. For example, it seems reasonable to assume that less semantic or phonological processing at baseline would predict better naming ability, regardless of treatment, but this assertion should be verified through direct analysis. Another reason to think that the group analyses could yield different results than the within groups is the pattern of correlations between predictors: More significant relationships exist among the predictors in the whole groups than in the PMT and SFA groups (see *Methods – Analysis*). Related to these different patterns, the final reason for the group analyses, is that the combined group increases power to detect significant relationships, since the distribution of the variances would change with increased observations.

The study outcomes were confrontation naming accuracy of treated words, untreated words that are semantically (SFA) or phonologically related (PMT) to the treatment items (henceforth, untreated-related), and untreated words that are neither semantically nor

phonologically related to the treated items (untreated-unrelated) at two time points: immediately post-treatment and 3 months post-treatment (follow-up). Immediately post-treatment, performance on the treated nouns represented acquisition, untreated-related nouns represented near-generalization, and untreated-unrelated nouns represented far-generalization. At follow-up, performance on treated nouns represented maintenance of acquisition, untreated-related represented maintenance of near-generalization, and untreated-unrelated represented maintenance of far-generalization. Although determining the degree of change from baseline to post-treatment and follow-up is not the purpose of this study, it is worth noting that the main study from which these data were derived indicated differences between all of the outcome stimulus types (Kendall et al., 2019). That is, both PMT and SFA groups demonstrated significant acquisition over baseline with a medium effect size immediately post-treatment, and both groups demonstrated significant near-generalization over baseline immediately post-treatment and at follow-up. The near-generalization effect size was small in the SFA group, but the PMT group's effect size was below the criterion for a small effect. Finally, neither group demonstrated far-generalization and the effect sizes were below the threshold for a small effect size. These findings justify examining the different outcome types in separate models, instead of together. See *Methods: Outcomes* for information on the characteristics of these stimuli.

Predictions. The predictions below were based on a combination of IA model predictions, mechanisms of the treatments in this study (see *Methods: Treatment* below for details), and the current available empirical evidence. Given the skills measured for the predictors and outcomes, predictions will be outlined in detail, first based on macro-linguistic input skills, then micro-linguistic abilities. Please see *Methods: Outcome Variables* and *Linguistic Predictor Variables* for detailed information on these measures. A summary of the

research questions, predictors, predictor relationships to the treatments, and the predictions is available in Table 1.

Macro-linguistic abilities

Input. Whole network auditory comprehension abilities were predicted to significantly, positively relate to naming of treated and untreated-related nouns immediately post-treatment and at follow-up in the PMT group, but not the whole group or the SFA group. Auditory comprehension was not predicted to relate to naming of untreated-unrelated items or maintenance thereof for any group. Likewise, input phonological abilities were predicted to positively relate to naming of treated items in the PMT group immediately post-treatment and at follow-up, but not to any untreated items, and not to the SFA or whole groups.

Whole network input processing was measured in this study by the auditory comprehension T-score of Comprehensive Aphasia Test (CAT; Swinburn, Porter, & Howard, 2004; see Figure 3-A). The auditory comprehension T-score was a standardized score (Mean = 50, *SD* = 10) derived from scores on the spoken word, sentence, and paragraph comprehension subtests. Phonological input processing was measured by a composite score of auditory phonological processing on the Standardized Assessment of Phonology in Aphasia (SAPA; Kendall et al., 2010; see Figure 3-B). The selection of these tasks as indicators of input processing were in line with measures used in previous studies (e.g., Freedman & N. Martin, 2001; Gupta et al., 2006; N. Martin & Saffran, 1999). Overall, input tasks required some auditory stimulus and either a verbal or gestural response. The gestural response involved pointing to one of four line drawings. Referring back to Figure 2, the linguistic network consists of unitary semantic and lexical levels, and connected input phonology and output phonology buffers. Input tasks probed semantics, the lexical level, and the input phonology buffer.

The auditory comprehension T-score represented the integrity of the spread of activation between input phonology and lexemes, and lexemes and semantics, or "whole network" processing. Because sentence and paragraph comprehension entailed maintaining information over time to respond correctly, the T-score also represented verbal short-term memory (STM). The role of verbal STM in linguistic processing has been well established in the literature. Verbal STM was conceptualized in this study as activation maintenance within the linguistic network, and an inability to maintain activation results in a verbal STM deficit (Cowan, 1996; N. Martin & Saffran, 1997). Thus, low comprehension scores may represent a transmission deficit (i.e., activation has difficulty moving from one level to the next) or a maintenance deficit (i.e., activation decays too quickly, verbal STM deficit), but because both are essential for adequate linguistic processing, it is not necessary to tease these apart within the context of this study.

Returning to the role of auditory comprehension in treatment-induced change, good comprehension at the word and sentence level is critical for understanding task instructions during PMT and providing adequate verbal responses to prompts. The PMT protocol will be explained in more depth in *Methods – Treatment*, and a summary of the protocol is available in Appendix B; however, for the predictions to make sense, it is important to note here that participants received complex instructions throughout the protocol. For example, while exploring a single phoneme, the clinician might have asked the participant, "What are your lips doing when you make that sound?" Likewise, in training a phoneme sequence, the clinician might have said to the participant, "This [sequence] is [riʃ]. Make it say [ri]." These types of tasks clearly require a certain level of auditory processing abilities for PWA to fully participate in the treatment. This increased participation also means more training opportunities, which enhances maintenance.

The SFA treatment, on the other hand, may not be as strongly tied to comprehension abilities, since the task itself does not change throughout the protocol. PWA had to name features that they associated with a given target for every trial, so once a participant understood the task, then the real work of the treatment shifted to identifying and verbalizing those features. Likewise, because auditory comprehension is not expected to play a role in participants' ability to complete the treatment tasks, it would have no effect on the amount of training for items and, thus, would not relate to maintenance. Importantly, the point that the work of treatment was in production tasks could be applied more generally to explain why the auditory comprehension T-score was not expected to predict far-generalization in any group: Comprehension ability facilitated participation but attempting or completing the tasks was what strengthened network connections.

On the phonological end of the network, the Auditory Phonological Processing portion of the SAPA (henceforth, SAPA 2) assessed the integrity of the links between the lexical level and input phonology through real word and nonword rhyming, lexical decision, and minimal pair judgment tasks. Participants responded with yes or no on each task, either verbally or gesturally. This subtest was predicted to significantly relate to acquisition and maintenance thereof in the PMT group. This prediction was based on the assumption that better input phonological processing facilitates error detection, at a minimum, which allows participants to self-correct and, thus, practice the correct response, which strengthens network connections. Input phonological processing was also necessary for completing some of the tasks in PMT, such as the perceptual tasks ("Which [mouth picture] matches this sound: [p]?") and Socratic questioning ("Is your [p] the same as my [p]?"). Although input phonological processing is important for completing task elements, it is not likely that this skill will lead to near- or far-generalization

because error monitoring can only lead to acquisition of new linguistic rules or patterns as long as output phonology is engaged (Gambi & Pickering, 2017; Kittredge & Dell, 2016).

A final component of this prediction was that strong input abilities were predicted to relate to improved naming for the PMT group only, but not the SFA or whole group. Why was input expected to *not* relate to the whole group? This, unfortunately, is due to differences in the number of items between SFA and PMT (see Table 5). When tallying responses, the SFA group had nearly twice as many as the PMT group for treated and untreated-related items. (The number of responses on the untreated-unrelated items is the same for both groups after accounting for missingness.) Thus, if an effect is not predicted for the SFA group, then it would also not be expected for the larger group. This raises the question of whether the groups should be combined for a whole group analysis, but there were three arguments in favor of carrying these analyses out. First, if the SFA group was predicted to demonstrate a relationship between a predictor and the outcome, then the effect would likely also be observed in the whole group, but the addition of the PMT data may change the strength of that relationship (e.g., the combined data might have weaker or stronger effect sizes). Second, the number of untreated-unrelated items was equal for both groups (and, once missingness and outliers were addressed, the number of subjects in each group was equal as well; see *Methods – Analysis*). Thus, in the analyses of these items, the whole group could have results that represent fundamental baseline skills irrespective of treatment group. Third, these analyses were experimental, so there was a possibility that unexpected results could occur, but they had to be tested first.

Table 1

Overview of Study Research Questions, Variables, Variable Relationships to Treatments, and Predictions

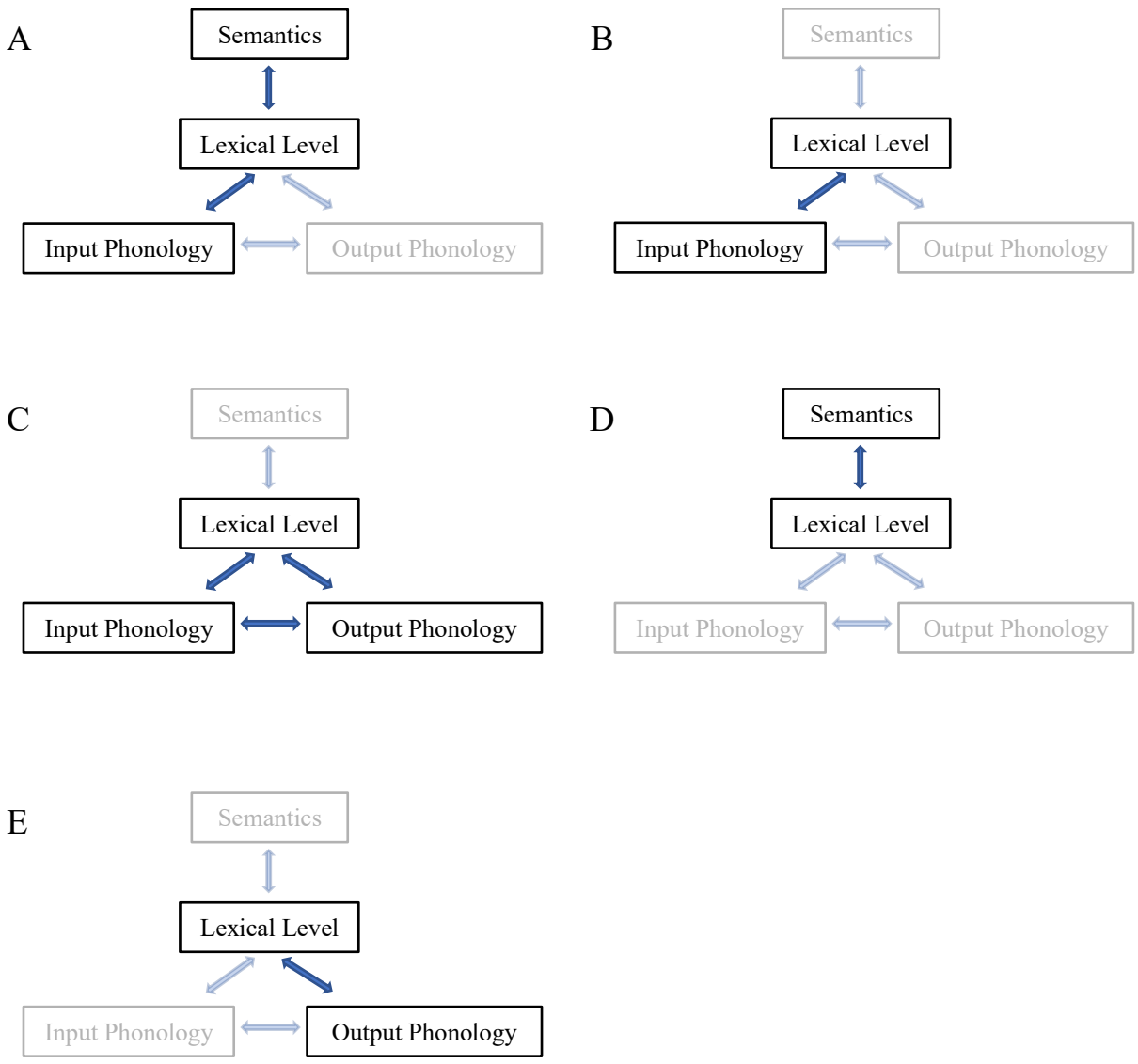
General Research Question Which linguistic skills contribute to naming performance on treated, untreated-related, and untreated-unrelated items immediately post-treatment and at follow-up with the PMT and SFA treatment groups and for both groups combined?			
MACROLINGUISTIC PROCESSING			
Input Processing			
Task Name (Score Type)	***Probes/Relies On…***	***Underlies X Treatment Element (Protocol)***	***Predictions: Stronger skills in this area will result in…***
CAT T-score (T-score)	Comprehension of Spoken Words - Lexical-semantic input processing Comprehension of Spoken Sentences - Lexical-semantic input processing (single words) - Syntactic decoding - Verbal STM (Simple Sentences) - Verbal WM (Complex Sentences) Comprehension of Spoken Paragraphs - Lexical-semantic input processing - Syntactic decoding - Verbal STM - Verbal WM	- Understanding basic prompts to generate features ("*Where?*"; **SFA**) - Understanding basic prompts/cues to complete tasks to aid in metalinguistic processing ("*Look.*", "*Feel*.") and some phonological awareness (PA) tasks ("*Same*?"; **PMT**) - Understanding prompts to name item ("*What's this*?"), especially at the beginning of treatment (**SFA**) - Understanding prompts to generate features ("*What do you do with this?*"*;* **SFA**) - Understanding prompts to perform PA tasks (e.g., *"What do you feel when you make that first sound?"*) and Socratic questioning ("*Was [your production] the same as mine?*"; **PMT**)	- Better naming of treated and untreated-related nouns in the PMT group immediately post-treatment - Maintenance of treated and untreated-related nouns in the PMT group at follow-up
SAPA – Subtest 2: Auditory Phonological Processing (Raw score out of 59)	Real Word Rhyme* - Lexical-phonological processing - Verbal STM Nonword Rhyme - Phonological processing - Verbal STM Lexical Decision* - Lexical-phonological processing (integrity of lexical-phonological links)	- Accurate detection of differences between individual phonemes, phoneme sequences, and errors during PA tasks and to facilitate participation in Socratic questioning (**PMT**)	- Better naming of treated and untreated-related nouns in the PMT group immediately post-treatment - Maintenance of treated and untreated-related nouns in the PMT group at follow-up

	Minimal Pairs (Nonwords) - Phonological processing - Verbal STM *Abilities in this test could partially rely on semantic processing abilities, but skills fundamentally depend on spread of activation from phonology to lexical level in IA network		
Input-Output Processing			
SAPA – Subtest 3: Repetition, Parsing, and Blending (Raw score out of 33)	Real Word Repetition - Input-output phonological processing (integrity of links between input and output phonology, with some influence from lexical-semantics and/or long-term memory) - Verbal STM Nonword Repetition - Input-output phonological processing (integrity of direct links from input to output phonology) - Verbal STM Real Word Parsing - Input-output phonological processing (integrity of links between input and output phonology, with some influence from lexical-semantics and/or long-term memory) - Verbal WM Nonword Parsing - Input-output phonological processing (integrity of direct links from input to output phonology) - Verbal WM Real Word Blending - Input-output phonological processing (integrity of links between input and output phonology, with some influence from lexical-semantics and/or long-term memory) - Verbal WM Nonword Blending - Input-output phonological processing (integrity of direct links from input to output phonology) - Verbal WM	- Accurate repetition of treatment stimuli (both words and nonwords; **PMT**) - Ability to decompose and recompose phoneme sequences into syllables and non-/words during PA tasks (**PMT**) - Ability to repeat target nouns when modelled by clinician (**SFA**)	- Better naming of treated, untreated-related, and untreated-unrelated nouns in the PMT group immediately post-treatment - Better naming of treated nouns in the SFA and whole group immediately post-treatment - Maintenance of treated, untreated-related, and untreated-unrelated nouns in the PMT group at follow-up - Maintenance of treated items in SFA and whole groups at follow-up

MICROLINGUISTIC PROCESSING			
Output Processing			
Task Name (Score Type)	***Indexes…***	***Relates to X Treatment Element (Protocol)***	***Predictions: Stronger skills in this area will result in…***
Semantic Impairment score $\frac{\text{Semantic Errors}}{\text{Total Errors}}$	- Integrity of activation transmission and maintenance among lexical-semantic links - Semantic errors: words that have some superordinate, subordinate, associative, or categorical relationship to the target word	- Accurate naming at the beginning/end of a training trial (**SFA**) - Ability to select related semantic features to target (**SFA**) - Accurate naming of related semantic features (**SFA**) - Supports manipulation of real word stimuli (**PMT**) - Supports spread of activation from lexical-phonological layers to semantics (**PMT**)	- Better naming of treated and untreated-related items in the SFA and whole groups immediately post-treatment - Better naming of treated, untreated-related, and untreated-unrelated items in the PMT group immediately post-treatment - Maintenance of treated and untreated-related items in the SFA and whole groups at follow-up - Maintenance of treated, untreated-related, and untreated-unrelated items in the PMT group at follow-up
Phonological impairment score $\frac{\text{Phonological Errors}}{\text{Total Errors}}$	- Integrity of activation transmission and maintenance among lexical-phonological links - Phonological errors: only phonologically related nonword errors (i.e., at least one segment in the same word/syllable position or a phoneme sequence in any part of the word that overlaps with the target word)	- Accurate naming at the beginning/end of a training trial (**SFA**) - Accurate naming of related semantic features (**SFA**) - Accurate completion of phonemic and phoneme sequencing tasks, including repetition and PA tasks (**PMT**)	- Better naming of treated items in the SFA and whole groups immediately post-treatment - Better naming of treated, untreated-related, and untreated-unrelated items in the PMT group immediately post-treatment - Maintenance of treated items in the SFA and whole groups at follow-up - Maintenance of treated, untreated-related, and untreated-unrelated items in the PMT group at follow-up

Figure 3

Theoretical Network Activation for Each Predictor Measure



Note. Shaded areas represent levels or buffers that are less active during task completion. Network structures as represented by the (A) CAT T-score, (B) SAPA 2, (C) SAPA 3, (D) semantic impairment score, (E) phonological impairment score.

Input-Output. Repetition abilities were predicted to significantly, positively relate to naming of treated, untreated-related, and untreated-unrelated nouns immediately post-treatment and at follow-up in the PMT group. Input-output is predicted to significantly, positively relate to naming of treated nouns immediately post-treatment and at follow-up in the whole group and SFA group.

This study will measure the relationship between input-output phonological processing using the Repetition, Parsing, and Blending subtest of the SAPA (henceforth, SAPA 3). The measure is described in more detail in *Methods*, but briefly this subtest consisted of phonological awareness tasks, which engaged both input-output phonology and the lexical level (see Figure 3-C), as well as verbal working memory (WM), and which some researchers define as attention plus verbal STM (Cowan, 1996, 2008; Engle, 2002; Minkina, Salis, & N. Martin, 2018). Specifically, the parsing and blending tasks required perceiving, maintaining, and manipulating real word and nonword stimuli (e.g., omitting a phoneme from a string or concatenating phonemes to produce a string), and then producing the response of this manipulation. Active manipulation was the aspect of the task that required attentional processes. The repetition task entailed perceiving and immediately reproducing real word and nonword stimuli without any manipulation, and therefore relied on verbal STM processing, rather than WM. In either case, representations must be maintained and transmitted from input to output phonology. Real word stimuli also received activation from lexical-semantics. In the case of an impaired link between the input and output buffers, the activation route for repetition was input-lexeme-output, because activation could not transmit directly from input to output. Ultimately, performance on the SAPA 3 represented the integrity of the links between input and output phonology and the lexical level, and verbal short-term and working memory.

In the PMT group, the skills measured on this subtest closely mirror many of the treatment tasks. Repetition is a key activity for both single phonemes and phoneme sequences, while parsing and blending are core activities in tasks involving phoneme sequences. Being able to complete these activities facilitated progression through the protocol, so that the focus of treatment was strengthening network connections by completing the tasks – not on learning how to complete the tasks in the first place. A growing body of evidence of the relationship between increased "patient acts" and greater gains supports this hypothesis (Baker, 2012; Cherney, 2012; Evans et al., 2020; Gravier et al., 2018). Throughout the course of completing these treatment acts in the PMT protocol, participants acquired the treated items, but the improved functioning of lexical-phonological processing should mean that their ability to name untreated items also improved. Evidence for participants' improvement in their *processing capabilities*, and not just their facility with the trained sequences, would be demonstrated by improvement on untreated-unrelated items, as the sequences in those nouns were not trained at all. Furthermore, maintenance of these gains would be expected because these increased processing abilities would aid in everyday communication after treatment ends, which in turn increases the likelihood of transfer to long-term knowledge (Hinton & Plaut, 1987; Plaut, 1996). The prediction that far-generalization and retention will occur is also supported in previous studies of PMT (Kendall, Oelke, Brookshire, & Nadeau, 2015; Kendall, Rosenbek, Heilman, Conway, Klenberg, Gonzalez Rothi, et al., 2008; Minkina et al., 2019).

In the SFA and whole groups, repetition abilities would have been important when the clinician modeled a target for a participant. Therefore, in participants who relied on modeling cues, increased ability to repeat would have boosted their ability to acquire treated items. However, this scenario only really applies to participants who were more severely impaired. In

participants who were less affected, the lexical to output phonological processing abilities that are reflected on the SAPA 3 would have played a role in participants' ability to name the targets and their features with minimal phonological errors. This hypothesis could be tested directly by evaluating predictors in separate models for milder and more severe participants. Accurate naming would have led to increased acquisition immediately post-treatment. Regarding maintenance, the same improved processing abilities that underlie performance on the SAPA 3 would have again facilitated use of acquired words in everyday contexts, thereby increasing the number of training opportunities. However, accurate, phonologically well-formed productions in this group would not have led to any type of generalization. Generalization was not expected because any reductions in phonological errors would be inherently tied to the stimuli themselves and would not come about as a consequence of improved phonological processing overall. Because the SFA was the larger of the two groups, the predictions for SFA were expected in the whole group; although the effect should be stronger if the PMT groups is also expected to improve.

Table 2

Overview of Error Coding Scheme (Adapted from Roach et al., 1996)

Response Type	Description
Correct	Target is produced as indicated in stimulus list
Motor	Inserted schwa (commonly between 2 consonants): *blue* → [bəlu]] Distortions (do not cross phonemic boundary, but sound distorted)
Filler	Produced prior to correct response: "um, uh, the…, that is…, you know this…"
Incorrect	Target is not produced as indicated in stimulus list
Semantic	Must be a real word that is obviously related to the target, but unrelated to the target according to the phonological criteria below: *chair* → *table*
Phonological	Substitution (replacing one sound with another): *popular* → [pɪpjulɚ] Addition (added sound): *eskimo* → [ɛskɪtmoʊ] Transposition (switching sound placement): *animal* → [æmɪnəl] Omission (deletion of a sound): *eskimo* → [ɛsɪmoʊ] Multiple Errors: *popular* → [pɪpjutɑɪ] or [pɪpjulɑɪd]
Other	
Unrelated real word	Must be a real word: *popular* → *sun*
Mixed	Mix of any 2 types of errors: *maze* → *puzzle*
Omission	Omission (said nothing): "I don't know" or no response
Description	Circumlocution or description: "It's wood and has four legs" for "table"

Micro-linguistic abilities

Micro-linguistic abilities are measured here by the proportion of either semantic or phonological errors produced on 50 naming trials of the untreated-unrelated items at baseline, according to a modification of the criteria established by Roach et al. (1996) for the Philadelphia Naming Test. These scores more accurately represent the degree of impairment. Because these measures were obtained from performance on a production task, the phonological impairment score in particular represented output phonology. The untreated-unrelated items were used to derive the impairment scores because (1) both groups named this exact same set of words, which is not true of the treated and untreated-related stimulus sets, so that the scores can easily be compared across groups, and (2) audio-recordings were available for these data, which was not true of the Boston Naming Test data. Errors were coded as semantic if they bore any obvious semantic relationship to the target (e.g., *table* → *chair*, see Figure 3-D), and phonological if they shared initial or final consonants, stressed vowels, phonemes in the same word positions, or phoneme sequences (*animal* → *aminal*, see Figure 3-E); phonological errors included phonologically-related nonword responses. See Table 2 for an overview of the coding scheme and Appendix E for the complete error coding guide. Once semantic and phonological errors were tallied, they were then divided by the total number of errors, so that higher scores represented worse processing ability and lower scores
represented better processing ability.

Semantic impairment. Semantic impairment was predicted to significantly, negatively relate to improved naming of treated items in all groups immediately post-treatment and at follow-up. Semantic impairment was also predicted to negatively relate to naming of untreated-

related items immediately post-treatment for all groups, and to untreated-unrelated items immediately post-treatment for PMT. General semantic processing underlies all of the other linguistic skills involving real word stimuli that have been discussed. Thus, to perform well on any of the other assessment or treatment tasks, the PWA must possess good semantic abilities. Second, evidence from treatment studies suggests that spared semantic processing is critical for word acquisition and retention (Best et al., 2013; Dignam et al., 2017). This finding makes sense, since semantic features are critical in differentiating lexemes from one another (recall that lexemes "encapsulate" activation sent down from the semantic level). On the other hand, lexical-phonological associations are arbitrary (i.e., the phonological instantiations of the concept *cat* vary across languages, but the features used to identify a cat are nonarbitrary). Less semantic impairment would also facilitate maintenance, since consistent, accurate activation transmission between semantic features and lexical nodes would increase opportunities to use newly acquired nouns.

Beyond acquisition, low semantic impairment was predicted to relate to near-generalization because intact lexical-semantic links are integral for further spread of activation to neighboring nodes. That is, the stronger the connections between lexemes and semantics, the further activation could travel unimpeded. This was true whether the network was being stimulated by the SFA or PMT protocol. However, far-generalization was not predicted in the SFA group (or, by extension, the whole group): The untreated-unrelated items were specifically selected because they were not related to the trained items in any way. These items were likely beyond the "reach" of activation from trained items, because they did not share many semantic nodes. Conversely, because PMT trained phonology, a much wider range of lexemes and semantic nodes could receive activation (assuming that lexical-phonological links were adequate

for activation transmission from the phonological to lexical levels), and thus relate to better post-treatment performance on untreated-unrelated items.

Phonological impairment. Phonological impairment was predicted to significantly, negatively predict naming of treated items across all groups, and these gains would be maintained. Low phonological impairment was also predicted to positively relate to near- and far-generalization, and maintenance thereof, in the PMT group. Just as semantic processing undergirds performance on the predictor variables, so too does phonological processing. Furthermore, phonological processing was needed in SFA to produce adequate verbal responses during the treatment tasks, which entailed a great deal of naming. This prediction aligned with the hypothesized relationship between the SAPA 3 and outcomes in the SFA group. Previous work also indicates that good phonological performance plays a supporting role in word acquisition, especially in the case of low frequency or low imageability words (e.g., N. Martin & Saffran, 1999). Although all of the treated words in the SFA protocol are imageable by nature, frequency was varied across participants so that those with more severe impairment received high frequency words during treatment and assessments; participants who were milder received low frequency words to ensure that PWA had room to improve (Fergadiotis, Swiderski, & Hula, 2019). Since baseline severity determined which set of words participants received, *post hoc* analyses of the relationship between phonological impairment (and other baseline skills) and the outcomes by severity can help determine whether there was any difference between these groups.

For the PMT group, good phonological processing would have facilitated participation and allowed activation transmission from phonology to semantics. That is, participants could have successfully complete exercises with more phonemes and phoneme sequences more quickly, meaning that they moved on to more complex stimuli sooner. Likewise, a participant

who successfully completed tasks could have received more difficult metalinguistic questions from the clinician. Also, as mentioned in the discussion of semantic processing, lexical-phonological links had to be sufficient for activation transmission to flow to semantics. Thus, increased naming performance following PMT critically depended on strong lexical-phonological connections. As with the SFA group predictions, this hypothesis was consistent with the hypothesis that SAPA 3 would relate to higher naming ability across stimulus types.

Chapter 2: Methods

Study Design

This was a retrospective analysis of data from a randomized control trial (RCT; Kendall et al., 2019) funded by a Veterans Administration Rehabilitation Research and Development Merit award (VA RR&D C6572R) comparing Phonomotor Treatment (PMT; Kendall et al., 2015; Kendall et al., 2008) and Semantic Feature Analysis (SFA; Boyle, 2010; Boyle & Coelho, 1995) - two anomia treatment approaches in aphasia. The RCT was approved by the University of Washington Internal Review Board and the VA RR&D.

Participants

People with aphasia were recruited through the Puget Sound Veterans Healthcare System (Seattle and American Lake), the Northwest Aphasia Registry and Repository (maintained by the University of Washington Aphasia Research Lab and Portland State University Aging & Adult Language Disorders Research Lab), and local area speech-language pathology clinics. A total of 96 participants were screened for the RCT: 21 failed the screening, 15 passed the screening but declined to participate in the study, and 60 passed the screening and agreed to participate in the study. Once participants enrolled in the RCT, they were prohibited from participating in individual speech-language therapy but were allowed to attend community-based communication groups for the duration of the study (i.e., six months).

Inclusion. All participants demonstrated chronic (i.e., greater than six months post-onset) aphasia due to cerebrovascular accident, primarily in the left cerebral hemisphere, although two participants demonstrated right-hemisphere lesions. Presence and size of lesion were confirmed by a board-certified neurologist through inspection of computed tomography (CT) and magnetic resonance imaging (MRI) scans obtained at the time of hospitalization. All participants met the

criteria for a diagnosis of aphasia, as demonstrated by scores below cut-off on several Comprehensive Aphasia Test (CAT; Swinburn, Porter, & Howard, 2004) subscales (see Table 1 for specific subtests). All participants were evaluated for presence and severity of apraxia of speech (AOS) by trained speech-language pathologists, using behavioral criteria defined by Duffy (2013), such as slow speech rate, prosodic abnormalities, and sound distortions, during a range of expressive language tasks (i.e., single word-repetition to spontaneous conversation). Participants with mild or moderate AOS, according to these features, were included in the study. Finally, all but two subjects were monolingual English speakers, and the two multilingual speakers reported speaking English as their primary language at the time of their strokes.

Table 3

CAT Subtests Used to Determine Presence and Severity of Aphasia

Subtest No.	Subtest Name
7	Comprehension of Spoken Words
9	Comprehension of Spoken Sentences
11	Comprehension of Spoken Paragraphs
12	Repetition of Words
13	Repetition of Complex Words
16	Repetition of Sentences

Exclusion. Participants were excluded from study participation if any of the following characteristics were observed or reported: neurodegenerative disease; untreated depression or other psychiatric disorders; severe AOS, defined as a restricted phonemic inventory, inability to repeat single phonemes, automatic speech that is more impaired than spontaneous speech, and/or limited meaningful verbal expression; chronic medical conditions that could interfere with standard treatment administration; and/or severe, uncorrected visual or hearing impairment.

Following the consent process and in-depth assessment, 60 participants were randomly assigned to receive either PMT or SFA, so that there were 30 in each group. After the treatment phase had begun, two people were withdrawn from the PMT group: One was withdrawn by the

researcher after it was discovered that she sustained a head injury immediately prior to study enrollment, and one withdrew himself due to family obligations. The number of people who completed the study was 58. There were 33 males and 25 females, the average age was 63.3 (*SD* = 11.43), years of education was 14.8 (*SD* = 2.47), and years post-onset of aphasia was 4.15 (*SD* = 4.39). Thirty people were assigned to the SFA group, which consisted of 18 males and 12 females. The SFA group, on average, was 63.4 (*SD* = 12.3) years old, had 15.2 (*SD* = 2.8) years of education, and were 4.1 (*SD* = 4.7) years post-onset. The number of people who completed the PMT protocol was 28, 15 of whom were males and 13 females. The PMT group had an average age of 63.3 (*SD* = 10.6), 14.3 (*SD* = 2.0) years of education, and were 4.3 (*SD* = 4.7) years post-onset. Chi-square tests determined that there were no significant differences between groups on any demographic characteristics (all *p*s > 0.05). See Table 3 for more details.

Assessment

All participants were assessed for baseline linguistic and cognitive skills at three time points: immediately pre-treatment, immediately post-treatment, and three months post-treatment (follow-up). Assessments were carried out over 2-3 days at each time point, with all measures – excepting the outcome tasks – being administered only once per timepoint. The outcome tasks, which measure confrontation naming of (1) treated nouns, (2) untreated-related, and (3) untreated-unrelated nouns (see *Outcome Variables* below), were administered in randomized lists 2-3 times per timepoint to control measure reliability (Kirk, 2013). The Raven's Progressive Colored Matrices (Raven, Court, & Raven, 1998) was administered to measure non-linguistic cognitive skills. Confrontation naming ability was measured with the Boston Naming Test – 2nd Edition (BNT; Kaplan, Goodglass, Segal, & Weintraub, 2001). The Standardized Assessment of Phonology in Aphasia (SAPA; Kendall et al., 2010) measured phonological processing through

reading (SAPA Subtest 1), and auditory processing (SAPA Subtest 2), and repetition, parsing, and blending (SAPA Subtest 3). To assess overall language abilities, the Boston Naming Test (Kaplan et al., 2001) and selected subtests of the CAT were administered, including tests of auditory and written comprehension, word and sentence repetition, and word and sentence writing. To assess discourse, the Story Retell Procedure (Doyle et al., 2000; Doyle et al., 1998; McNeil, Doyle, Fossett, Park, & Goda, 2001) was administered using parallel forms (three different stories at every timepoint). Finally, to measure different aspects of verbal memory, participants were also assessed with the Wechsler Memory Scales – 4th Edition (Wechsler, 2009) and subtests of the Temple Assessment of Language and Short-Term Memory in Aphasia (N. Martin, Minkina, Kohen, & Kalinyak-Fliszar, 2018). There were no significant differences between the SFA and PMT groups on any baseline language or cognitive measure. See Table 3 for detailed baseline scores. To ensure parity on the semantic and phonological impairment scores across groups, a Mann-Whitney U-test was performed on the unstandardized data to compare the medians[1] for each score between groups. The Mann-Whitney tests revealed that no significant differences existed between groups for either the semantic or phonological impairment scores (all p's > 0.05).

[1] The Mann-Whitney *U*-test uses group medians to compare independent groups.

Table 4

Participant Characteristics

Pt.	Tx Group	Age (yrs)	Sex	Education (yrs)	TPO (yrs)	Handed-ness	RCPM (max. 36)	BNT (max. 60)	CAT: Semantic Processing (T-score)	CAT: Auditory Processing (T-score)	CAT: Reading Comp. (T-score)	CAT: Writing (T-score)	SAPA (max. 144)	TALSA: Digit Span (max. 7)	TALSA: Word Span (max. 7)
1	PMT	71	F	15	1.42	R	35	49	62	59	66	69	116	3.20	3.00
2	PMT	46	F	16	1.25	R	28	23	54	58	62	69	79	4.05	4.05
3	PMT	59	F	16	5.25	R	33	42	62	50	54	56	110	4.00	3.10
4	PMT	67	M	12	2.75	R	29	46	62	60	59	59	94	4.15	3.10
5	PMT	70	F	12	6.75	R	23	6	38	39	48	44	61	1.05	1.05
6	PMT	40	F	16	1.83	R	36	5	62	51	48	52	73	3.00	2.00
7	PMT	59	M	18	2.25	R	32	6	54	54	50	55	58	2.20	2.05
8	PMT	71	F	14	0.83	R	31	36	62	58	60	52	111	2.15	2.00
9	PMT	65	M	16	8.42	R	34	4	62	44	49	53	62	2.10	3.05
10	PMT	73	M	16	2.33	R	28	6	54	49	44	43	49	1.10	2.00
11	PMT	73	F	13	2.50	R	34	12	50	50	54	53	100	3.05	2.00
12	PMT	67	M	16	4.17	R	32	20	54	53	56	57	84	3.15	2.20
13	PMT	46	M	13	7.08	L	--	14	50	45	44	41	48	1.10	1.15
14	PMT	59	F	12	3.67	R	26	0	16	35	40	48	30	--	--
15	PMT	71	F	16	3.67	R	34	41	54	52	66	62	109	3.15	3.10
16	PMT	90	F	12	6.42	L	29	26	54	58	59	60	79	4.20	3.05
17	PMT	63	M	13	4.00	R	35	32	62	45	56	50	77	3.00	2.00
18	PMT	60	M	14	2.75	L	19	2	62	46	50	46	50	2.05	2.00
19	PMT	46	M	11	24.83	R	35	3	62	41	37	51	41	--	--
20	PMT	74	F	14	0.92	R	25	3	62	44	48	44	53	--	1.00
21	PMT	63	M	16	2.00	L	32	2	62	50	42	45	35	--	1.15
22	PMT	62	M	16	0.83	R	35	46	54	58	55	59	105	3.15	3.10
23	PMT	50	M	12	2.08	R	29	1	62	52	51	48	79	--	--
24	PMT	67	F	16	1.67	R	32	39	62	45	54	51	93	1.20	2.05
25	PMT	70	M	12	4.42	R	21	32	54	56	49	50	55	2.15	2.00
26	PMT	66	F	12	9.42	R	22	40	62	58	54	--	94	3.10	2.20
27	PMT	65	M	14	1.42	L	22	52	62	55	54	49	88	3.10	2.15
28	PMT	59	M	18	4.17	L/R	33	16	50	47	54	48	60	3.05	1.20
29	SFA	72	M	20	2.00	R	27	11	62	45	46	47	59	1.05	2.00
30	SFA	69	M	19	8.50	L	35	28	62	50	54	51	61	2.20	1.20
31	SFA	38	F	12	3.42	R	36	38	62	48	49	57	100	3.00	3.05
32	SFA	72	F	16	2.25	R	35	46	54	56	62	57	106	4.15	3.10
33	SFA	57	M	10	0.92	R	25	4	54	47	44	46	49	2.00	1.15
34	SFA	44	F	16	0.75	R	36	50	62	57	60	55	112	4.10	3.10
35	SFA	45	M	14	1.00	R	34	7	41	52	56	52	73	3.15	2.05
36	SFA	91	F	18	0.67	R	--	3	62	39	49	44	32	1.05	1.15
37	SFA	69	M	13	10.33	R	32	30	54	56	52	50	74	4.00	3.05
38	SFA	63	M	12	3.17	L	28	17	54	58	56	53	70	3.00	2.20
39	SFA	70	F	13	3.58	R	34	2	39	43	45	47	44	--	1.20
40	SFA	59	F	16	9.25	R	35	14	39	45	49	50	90	2.15	2.10
41	SFA	65	M	13	0.83	R	25	50	54	58	54	59	92	5.10	3.05

Pt.	Tx Group	Age (yrs)	Sex	Education (yrs)	TPO (yrs)	Handed-ness	RCPM (max. 36)	BNT (max. 60)	CAT: Semantic Processing (T-score)	CAT: Auditory Processing (T-score)	CAT: Reading Comp. (T-score)	CAT: Writing (T-score)	SAPA (max. 144)	TALSA: Digit Span (max. 7)	TALSA: Word Span (max. 7)
42	SFA	56	F	12	2.08	R	27	46	50	50	52	48	80	2.00	2.00
43	SFA	77	F	16	14.75	R	35	11	62	50	57	59	69	2.05	1.20
44	SFA	74	M	16	2.75	R	24	1	41	38	38	44	36	2.00	1.05
45	SFA	64	M	19	2.17	R	28	38	50	52	58	50	84	2.10	1.20
46	SFA	55	M	15	1.58	R	35	54	50	63	63	60	118	6.05	4.20
47	SFA	75	F	12	14.08	R	22	1	50	52	56	52	44	3.00	2.00
48	SFA	77	M	18	11.00	R	34	1	54	47	51	49	53	--	2.10
49	SFA	55	M	16	1.50	R	32	47	62	57	60	62	111	--	--
50	SFA	66	M	16	1.00	R	33	52	62	56	57	55	99	5.05	4.00
51	SFA	68	F	18	1.83	R	22	11	16	41	45	48	58	1.10	1.05
52	SFA	62	M	18	6.17	R	25	17	16	39	50	49	100	1.05	1.00
53	SFA	58	M	12	1.83	R	--	28	54	56	50	58	87	--	--
54	SFA	44	F	16	1.58	L	31	0	47	39	45	47	44	--	1.05
55	SFA	79	F	14	9.25	R	19	47	35	55	59	52	96	4.10	3.15
56	SFA	75	M	20	0.50	R	31	36	54	46	52	62	102	3.20	2.15
57	SFA	47	M	11	1.83	R	36	7	62	48	54	54	69	4.10	2.15
58	SFA	55	M	16	1.17	R	34	52	62	65	65	69	117	6.20	5.00
All Grp. Mean		63.33		14.79	4.15		30.07	23.33	53.31	50.34	52.60	52.63	76.24	2.90	2.22
All Grp. *SD*		11.43		2.47	4.39		5.00	18.71	11.45	6.91	6.70	6.56	25.12	1.29	0.95
PMT Mean		63.29		14.32	4.25		29.78	21.57	55.93	50.43	52.25	52.37	74.75	2.72	2.23
PMT *SD*		10.61		2.02	4.66		4.99	17.84	9.82	6.64	7.17	7.26	24.95	0.98	0.79
SFA Mean		63.37		15.23	4.06		30.36	24.97	50.87	50.27	52.93	52.87	77.63	3.08	2.20
SFA *SD*		12.32		2.79	4.21		5.09	19.65	12.45	7.26	6.34	5.99	25.63	1.52	1.08

Note. Pt. = Participant; Tx = Treatment; TPO = time post-onset; RCPM: Raven's Colored Progressive Matrices; BNT = Boston Naming Test; CAT = Comprehensive Aphasia Test; SAPA = Standardized Assessment of Phonology in Aphasia; TALSA = Temple Assessment of Short-Term Memory and Language in Aphasia; PMT = Phonomotor Treatment group; SFA = Semantic Feature Analysis group

Treatment

For both the PMT and SFA groups, treatment occurred for two 1-hour sessions a day, 4-5 days a week, for 6 weeks. All participants completed 56-60 hours of treatment altogether. The clinician met participants in the university clinic, the participants' homes, or another quiet, well-lit location that was convenient for participants.

Phonomotor Treatment. PMT is a phonologically based anomia treatment motivated by the IA model and informed by the parallel distributed processing (PDP) model of phonology (Nadeau, 2001), which is a neurally motivated model that is compatible with the IA model, to provide specific guidance for intervention techniques. Essentially, PMT operates on the principle that every word contains phonemes and phoneme sequences – the implication being that improved phonological processing will result in improved word-finding. To improve phonological processing, PMT strengthens the connections between the phonological input and output buffers and between the phonological and lexical levels through repeated, multimodal activation of phonemes in isolation, nonword phoneme sequences, and real word phoneme sequences. Visual, auditory, tactile-kinesthetic, and motor modalities are engaged to enhance phonological activation during tasks that include repetition, parsing, and blending.

Stimuli. In this study, treated and untreated-related stimuli consisted of all English phonemes, 72 phonotactically-legal nonword phoneme strings of two to five segments, and 60 nouns selected from the Medical Research Council Psycholinguistic Database (Coltheart, 1981; see Appendix D for a complete list of stimuli). Of the nouns, 39 were selected for treatment, and the remaining 21 were set aside to measure generalization to untrained words that possessed the same phonological characteristics as the trained words. Specifically, all nonword and real word stimuli had low phonotactic probability (PP) and high neighborhood density (ND) to facilitate

acquisition, according to the work of Storkel, Armbrüster, and Hogan (2006), which found that healthy young adults learned more nonwords that featured low PP and ND than those that did not. PP is a measure of phoneme position and biphone frequency, while ND is a tally of all the words in the dictionary that differ from the stimulus item by the addition, substitution, or deletion of a phoneme (Vitevitch & Luce, 2016). Both measures were calculated for every stimulus item using the Irvine Phonotactic Online Dictionary (Vaden, Halpin, & Hickok, 2009).

Protocol. As stated above, PMT trained phonemes in isolation and phoneme sequences. Once the clinician explained the conceptual aspects of the protocol to the participants, they began training every phoneme in isolation. For each phoneme, the clinician provided a color photograph of a mouth in an exaggerated posture for each sound. These images served to help the participants describe the articulatory features of the phonemes during metalinguistic processing tasks, for example "to make /f/, you put your top teeth on your bottom lip." Throughout the single phoneme training, participants saw the mouth pictures, colored blocks, or grapheme tiles (thereby stimulating visual representations), described how the sounds were made (metalinguistic processing), observed the clinician producing the phonemes (visual/auditory), observed oneself producing the phonemes in a mirror (visual/auditory/tactile-kinesthetic), and made phonemic similarity judgments either verbally or by manipulating the visual aids when prompted by the clinician (auditory/visual/tactile/metalinguistic).

As participants mastered individual phonemes (i.e., produced them with 85% accuracy), they began practicing with the phonemes in sequences. When the clinician introduced phoneme sequences, he or she began with CV and VC combinations, and progressed to multisyllabic strings. Initially, sequences focused on phonotactically-legal nonword sequences to emphasize phonological processing but, as the participant progressed through the protocol, real word

sequences were incorporated and treated exactly like nonwords. Participants still practiced phoneme sequences using the same skills and materials as with the individual phonemes, but new target behaviors in this phase included phonological awareness tasks, including but not limited to, parsing (*"What sounds are in 'sog'?"*), blending (*"Put the sounds /p/, /a/, /m/ together. What word is that?"*), and phonemic substitutions (*"This says 'pog.' Replace /a/ with /æ/. What does that say?"*). See Appendix E for a complete protocol description.

Semantic Feature Analysis. SFA has been used as a treatment approach for aphasic anomia for the past 25 years and is based on the connectionist work by Collins and Loftus (1975) and later the IA model by Dell and colleagues (Dell, 1986; Dell & O'Seaghdha, 1992). The basic premise of SFA is that word retrieval is facilitated through repeated, targeted activation of features associated with a stimulus item to strengthen the links between the features and the stimulus (Hashimoto, 2012; Hashimoto & Frome, 2011)[2]. In particular, clients self-generating or self-selecting the features that are associated with the target item is considered to be a key component of the treatment that enhances acquisition and maintenance of treatment gains (Boyle & Coelho, 1995; Gravier et al., 2018) (Boyle & Coelho, 1995; Efstratiadou, Papathanasiou, Holland, Archonti, & Hilari, 2018; Gravier et al., 2018).

Stimuli. The stimuli for the SFA protocol in this study consisted of 240 nouns selected from the English noun imageability set (Reilly & Kean, 2007), and were then divided into eight different categories (see Appendix F). Noun frequency was determined using the SUBTLEX-US database, and nouns were divided into high or low frequency groups by a median split. The high frequency word group occurred in spoken American English 845-995 times per million words, while the low frequency words occurred 111-229 times; this difference was statistically

[2] Although some have argued that SFA simply serves to *organize* the word retrieval process (Anders & Bos, 1986; Boyle, 2017).

significant ($p < 0.01$). Each category contained 15 high and 15 low frequency nouns. For each high and low frequency set, 10 nouns were used for treatment, while the remaining five were set aside to measure generalization to untrained but semantically related stimuli. Finally, a decontextualized color photograph was selected to represent every noun, and then tested with five healthy controls to ensure that the images elicited target responses in 100 percent of naming opportunities. Each participant received either a high frequency or low frequency word set: Standard scores > 50 on the CAT auditory comprehension and repetition subtests were categorized as "less impaired," while standard scores < 50 were tagged as "more impaired." Severity for those whose scores were at 50 was decided through careful examination of their individual scores. Those in the "more severe" category received the high frequency word set, while those who were "less severe" received the low frequency word set. This decision was made to ensure an adequate level of difficulty for more and less impaired participants.

Protocol. For all 80 trained nouns, the clinician presented a color photo of the target and asked the participant to name the item. Regardless of whether the participant correctly named the target, the clinician then prompted the participant to generate features associated with the item. Namely, the participant was asked to name the category to which the item belonged, describe physical features of the target, state its function or an action associated with the target, provide some contextual information (e.g., typical location of the item), and share a personal association with the target (see Appendix G for overview of treatment protocol). A group of 13-14 words from every category was trained for ten hours before the clinician and participant moved on to the next group, regardless of accuracy. Once a group was completed, those words were not trained again.

Outcome Variables

This study contains three outcome types: confrontation naming accuracy of treated nouns, untreated-related nouns, and untreated-unrelated nouns. Treated and untreated-related nouns were presented in one list, while the untreated-unrelated nouns were presented in a separate list; the word order of each list was randomized for each day, at every timepoint. Participants were seated at a comfortable distance (roughly 30 inches) from a computer monitor and instructed to name each picture they saw using a single word. Each stimulus item was presented for 10 seconds, with an interstimulus interval of 2-10 seconds (as determined by participant needs). The stimuli were presented using Microsoft PowerPoint 2013 on either a Dell Latitude 7370 laptop with HP EliteDisplay S140u 14-inch screen, or a Dell Optiplex 9020 desktop with Dell 24-inch monitor. Verbal responses were audio-recorded using a head-mounted Audio-Technica Power Module AT8531 microphone connected to a Tascam US-125M USB audio mixing interface.

Treated nouns. The treated nouns assessed at pre-treatment, immediately post-treatment, and follow-up were the same as those described in *Treatment – Stimuli*: The PMT group trained on 39 words of low PP and high ND, while the SFA group trained on 80 words from across eight semantic categories. The difference in the number of nouns reflects differences in the standardized protocols, and the comparison of these protocols was the purpose of the main study. The statistical implications of these differences were noted in *Current Study – Predictions*. Confrontation naming accuracy on treated nouns immediately post-treatment served as a measure of acquisition of treatment skills, and naming accuracy at follow-up represented maintenance of those skills.

Untreated-related nouns. The untreated-related nouns were the same as those discussed in *Treatment – Stimuli*. The PMT group were assessed with 21 words of low PP and high ND,

and the SFA group were assessed with 40 high- or low frequency words (five words from each of eight categories) to measure the degree of treatment generalization to words that were either phonologically similar (PMT) or semantically related (SFA) to the trained items immediately post-treatment and at follow-up.

Untreated-unrelated nouns. A corpus of 50 nouns were selected from among the items in the Philadelphia Naming Test (Roach et al., 1996) and the Object and Action Naming Battery (Druks & Masterson, 2002). The words were carefully selected for their semantic and phonological attributes: none of the items were in any of the categories trained in the SFA protocol, and none of the items contained the same 3-segment sequences as PMT stimuli. The 50 words were divided into high- and low frequency sets, so that those with more severe aphasia were tested on the high frequency set, and those with less severe aphasia were tested on the low frequency set. Severity was determined in the same manner as described for the SFA group. The untreated-unrelated nouns served as a measure of treatment generalization beyond the semantic and phonological domains targeted during treatment immediately post-treatment and at follow-up.

Linguistic Predictor Variables

Macro-linguistic abilities. In the current study, the term "macro-linguistic abilities" refers specifically to overall comprehension and production skills, typically as measured by subtests on an aphasia battery. For example, single word comprehension and sentence repetition constitute two separate macro-linguistic skills. At the macro-linguistic level, skills can be further differentiated by modality: "input" skills encompass those that require the parsing and/or processing of items that are presented auditorily, such as a word comprehension task; "output" skills are those that require verbal production, such as a confrontation naming task; "input-

output" skills are those that require both of the previous skills in one go, namely, repetition. The following is a description of each of the input and input-output macro-linguistic variables in this study.

Input. In this study, input macro-linguistic skills were represented by the CAT comprehension T-score and the SAPA Subtest 2: Auditory Processing. The CAT T-score is a combination of the scores on the Comprehension of Spoken Words (Subtest 7), Comprehension of Spoken Sentences (Subtest 9), and Comprehension of Spoken Paragraphs (Subtest 11), that are then standardized for a mean of 50 and standard deviation of ±10. The word stimuli were controlled for imageability, frequency, grammatical regularity, animacy, and morphological complexity; the sentence stimuli were controlled for length, reversibility of nouns (i.e., whether only one or both nouns can serve as syntactic subject), and morphosyntactic complexity (Howard, Swinburn, & Porter, 2010). For Subtests 7 and 9, the examiner said a word or sentence, respectively, and the participant had to point to one of four black and white line drawings that best represented the spoken item. For Subtest 11, the examiner read aloud two, 3-sentence paragraphs. Following the reading of each paragraph, the participant answered yes/no questions to test comprehension.

The SAPA 2 is part of the larger SAPA battery (Kendall et al., 2010), which the test authors standardized on 47 people with chronic aphasia and then applied item response theory to improve validity (Baylor et al., 2011; Hambleton & Swaminathan, 2013). The SAPA 2 measures phonological processing ability for auditorily presented stimuli and the total score was 59, which is the raw, combined score from four tasks (number of items in parentheses): real word rhyming (n = 15), nonword rhyming (n = 22), lexical decision (n = 9), and minimal pair discrimination (n = 13). For each of these tasks, the participant listened to non-/word pairs and single non-/words

and indicated yes or no to questions about whether the stimuli rhymed (rhyming tasks), were words (lexical decision task), or were the same (minimal pair task). Although this test was designed to tap lexical-phonological (input buffer) links exclusively, semantic representations also received activation in the case of the real word stimuli. Importantly, the SAPA 2 also requires adequate verbal STM, as participants must maintain these auditory representations long enough to compare stimuli presented in pairs (Cowan, 1996; N. Martin, Minkina, Kohen, & Kalinyak-Fliszar, 2018; Murray, Salis, Martin, & Dralle, 2018; Salis, Kelly, & Code, 2015).

Input-Output. Input-output processes were measured with the SAPA 3. The SAPA 3 total score (33) is the sum of six tasks (number of items in parentheses): real word repetition (n = 9), nonword repetition (n = 5), real word blending (n = 5), nonword blending (n = 4), real word parsing (n = 5), and nonword parsing (n = 5). For the repetition tasks, participants heard a single non-/word and immediately repeated it. For the blending tasks, participants heard phonemes and/or syllables and had to combine them to produce a single non-/word. For example, participants heard [f]-[æd] to produce *fad* or [d]-[ɛ]-[m] to produce *dem*. In the parsing tasks, participants heard non-/words and separated them into phonemes, or they had to repeat the stimulus without a given phoneme or syllable. For example, participants heard *hat* and had to segment it into [h]-[æ]-[t], and they heard and *boxcar* and had to repeat the word without *car* (i.e., *box*).The SAPA 3 measures the integrity of the input-output phonological links, as well as their connections to the lexical level. As with the SAPA 2, real word stimuli in the SAPA 3 subtest entail semantic activation. While the repetition tasks rely on verbal STM, the intense stimulus maintenance and manipulation in the parsing and blending tasks requires verbal WM (McBride-Chang, 1995; Wagner, Torgesen, & Rashotte, 1994).

Micro-linguistic abilities. Micro-linguistic abilities are a measure of underlying linguistic skills, as described in the outline of the IA model above. In this study, micro-linguistic abilities were represented by the degree of semantic and phonological weight lesions, as determined by each participant's baseline performance on 50 untreated-unrelated items. Naming errors were coded according to a modified version of the Philadelphia Naming Test (PNT; Roach et al., 1996) manual, because the PNT provides a much more thorough guide for classifying errors and attributing them to either semantic or phonological lesions than the BNT guide. See Appendix E for the modified guide. Naming responses were coded as either "Correct" or "Incorrect." When the response was incorrect, it was then coded as a "Semantic," "Phonological," or "Other" error type. Semantic errors where any real words that were obviously related to the target and did not meet the criteria for phonological errors (e.g., *table* → *chair*). Phonological errors were more complex and had three key elements. Namely, the response was coded as a phonological error if it: (1) was not obviously semantically related to the target; (2) was phonologically similar to the target, defined as sharing an initial or final consonant, a stressed vowel, two or more phonemes in any position, or one or more phonemes in the same syllable or word position as the target; and (3) included any substitution, addition, transposition, or omission of one or more target phonemes. Nonwords were counted as phonological errors as long as they met these criteria. All other error types, such as omissions and semantically/phonologically mixed errors, were coded as "Other" to improve coding reliability and derive relatively "pure" measures of underlying semantic or phonological processing abilities.

Coding was completed by nine student volunteers in the University of Washington Speech and Hearing Sciences Department and the first author. The first author oriented the

students to the purpose of the study and the data, after which point, the entire team trained on practice data. The training data were recordings of naming trials from the follow-up period, and training continued until all coders reached 90% agreement across 150 trials. To resolve any coding questions and maintain fidelity, team members attended weekly meetings led by the first author. After the team completed error coding, the first author assessed 20% of the coded data for inter- and intrarater reliability using Cohen's κ: interrater reliability was $\kappa = 0.82$, and intrarater reliability was $\kappa = 0.92$. According to the rather conservative benchmarks set forth by McHugh (2012), these statistics indicate "strong" and "almost perfect" reliability, respectively.

Finally, once all data were coded, the first author calculated the semantic and phonological impairment scores. In previous studies that calculated some type of phonological or semantic error scores, the approach has typically been to either divide the number of errors for each type by the number of opportunities (e.g., Dell et al., 1997; Foygel & Dell, 2000), or to divide the number of errors for each type by the total number of errors (e.g., Minkina et al., 2016). Following this prior work, the impairment scores for the current study were calculated as the number of each error type, divided by the total number of errors:

$$\frac{No.\,Semantic\ or\ Phonological\ Errors}{Total\ No.\,Errors} \quad (1)$$

As an example, Participant 12 correctly named 29 items, had 11 semantic errors, five phonological errors, and five "Other" errors (total errors = 21). His semantic impairment score was calculated as 11/21 = .52, and his phonological score was 5/21 = .24. Note that this calculation resulted in lower scores indicating less impairment and higher scores more impairment. Three participants' processing scores could not be calculated because their responses were almost 100% omissions. (See exact scores in Table 4 below.) Using their

processing scores would not accurately reflect their poor baseline performance in the analyses, and omissions obscure the locus of impairment – whether semantic, phonological, or both.

Table 5

Accuracy, Errors, and Impairment Scores for Participants with Extremely High Omission Rates

Pt.	No. Correct (out of 50)	No. Sem. Errors	No. Phon. Errors	No. Other Errors	No. Total Errors	Sem. Score	Phon. Score
14	0	0	1	49	50	.00	.02
44	2	0	0	48	48	.00	.00
54	2	0	1	47	48	.00	.02

Note. Pt. = Participant; Sem. = semantic; Phon. = Phonological.

Analysis

The aim of this study is to determine which, if any, baseline linguistic characteristics account for naming performance immediately post-treatment and at follow-up. The best statistical technique to achieve this aim is regression, because it calculates each predictors' contribution to the amount of variance in the outcome. However, an ordinary least squares (OLS), or a "basic," linear regression is not appropriate for these analyses for two reasons. First, and most importantly, the sample size is too small. A power analysis using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) revealed that, to detect a small effect size with 5 predictors, the sample would need to be n = 92. Second, the assumptions of normality and independence for linear regression may be untenable. Data from clinical populations are notoriously non-normally distributed (Bono, Blanca, Arnau, & Gómez-Benito, 2017; Petitclerc, 2004), which can affect *p*-value calculations in small (n < 200) sample sizes. Specifically, a non-normal distribution of the residual errors (i.e., the distance between the actual and predicted values in the regression) can lead to skewed parameter estimates, meaning the contribution of the predictor to the outcome is not accurately represented. In the case that the assumption of normality is violated, more conservative *p*-values are recommended to avoid Type I error (McDonald, 2014). A model's *p*-

values will likewise be inflated if the cases are not independent, meaning that some values are systematically related in some way. This study contains two treatment groups, and each group contains different numbers of outcomes. These different groups are problematic for Research Questions 1 and 2, which query the relationships between predictors and outcomes across groups, because the outcomes within the PMT group would be more related to each other than with those in the SFA group and vice versa. Given these issues, it is clear that an OLS linear regression is not the most appropriate approach for answering this study's research questions.

Table 6
Number of Outcomes by Group and Timepoint

Group	Outcomes by Timepoint	Total Outcomes
All PWA (n = 58)	A2: 18,150 A3: 17,155	35,305
SFA (n = 30)	A2: 11, 410 A3: 10, 925	22,335
PMT (n = 28)	A2: 6,740 A3: 6,230	12,970

Fortunately, more sophisticated analysis techniques are available. To avoid some of the issues discussed above, this study relied on mixed-effects logistic regression analyses. Using logistic regression can help improve power, since the outcome is binary performance ("correct" or "incorrect") for individual items, rather than an average score for each participant (Wiley & Rapp, 2018). The basic function of logistic regression is to predict the likelihood that an event (Y) will occur, given some condition (X). In this study, the regressions determined the likelihood of a correct naming response (Y), given the predictors outlined in Table 1 (see equation 4 for model terms). Thus, for each participant, there were several hundred responses on the outcomes, rather than one per outcome type. The outcomes then served as the "cases," instead of the subjects. The effect of this conversion was that the power increased exponentially, so that the

number of outcomes now far outstripped the number of predictors. The numbers of outcomes for each group, after accounting for missing data, are shown in Table 6.

That these regressions had a mixed-effects structure also helped with the assumption violations. Unlike linear regression, logistic regression does *not* assume that the residuals are normally distributed, which resolves the normality violation. However, in an OLS logistic regression, the assumption of independence still applies. Recall that each participant responds on several naming trials, and each participant is in one of two groups. These data will violate the assumption of independence if this nesting is not accounted for. This problem can easily be handled using mixed-effects modeling. Mixed-effects models include terms for both fixed factors, or variables of interest (such as word comprehension), and random factors, or factors that may contribute to variance but are not of interest (such as item; Quene & van den Bergh, 2008). Using crossed mixed-effects logistic regression analyses, the goal of this study was to examine the influence of macro- and micro-linguistic factors on acquisition, near-generalization, far-generalization, and maintenance of treatment skills.

The general formula for logistic regression reflects the fact that probabilities are being modelled:

$$\hat{p} = \frac{1}{1 + e^{-1(b_0 + b_1 X_1 + b_2 X_2 + b_p X_p)}} \quad (2)$$

where $\hat{p}$ is the probability that an event will occur, b_0 is the intercept (the estimated value of $\hat{p}$ when all predictors equal zero), each subsequent b_p is the slope for the p^{th} predictor, X_p. Each slope is calculated, and the natural log e is taken of the value of the change in slope after all predictors have been accounted for. The value of "1" represents the event of interest occurring; thus, the numerator is the occurrence of an event, over the denominator, or total possible number of events. In this study, the event of interest is correct naming of an outcome stimulus, and the predictors are

the baseline linguistic variables discussed previously. Beyond the basic model, the random effects of subject, group, and item must also be accounted for to control for the noise that they introduce into the model (random effects have a mean of zero and an unknown standard deviation). This is done by adding terms to the model that account for these random factors:

$$\hat{p} = \frac{1}{1 + e^{-1(b_0 + b_1 G_i + b_2 T_j + b_3 (G_i * T_j))}} \quad (3)$$

where the new term G_i represents a subject, and T_j represents a given item. Note that an interaction term can also be included in the model for an item named by an individual. This modeling approach will be useful in determining the relative contribution of the predictors to the outcome, given that the outcomes represent performance of an individual, in a given group, on a given item. Specifically, the proposed analyses were well suited to this study.

Prior to analysis, the data were inspected and cleaned to the standards necessary for mixed-effect logistic regression. Because the predictors were on different scales and because unstandardized predictors can be difficult to interpret, all predictor variables were mean centered so that *Mean* = 0 and *SD* ± 1. Once standardized, the data were assessed for outliers, defined here as > 3 *SD* from the mean. Subsequently, data from four participants were capped or floored to minimize data loss; i.e., the one score that was > 3 *SD* above the mean was adjusted to +3 *SD*, and the remaining scores were adjusted to -3 *SD* (Kwak & Kim, 2017). Cook's Distance (Cook, 1977) was also used to identify influential data points within each model.[3] Influential data were removed from the models, so that up to four participants were removed from each of the whole group analyses, one person was removed from each of the PMT analyses, and up to three

[3] Cook's Distance D_i is a scaled measure of fitted values in mixed-effect models that calculates the influence of each level in a grouping (i.e., random effect) variable on the parameters. The cutoff value of D_i for a given model is calculated as $4/n$, in which n is the number of levels withing a grouping factor.

participants were removed from each of the SFA analyses. The removal of these participants resulted in better model fits, as determined by Akaike Information Criterion (AIC).

Because of anticipated or known differences between treated, untreated-related, and untreated-unrelated outcomes (see Methods – Outcomes for details), analyses consisted of 18 separate models – outcome type (3) x group (3) x timepoint (2) – with identical predictors. Given these considerations the initial model for all analyses was:

$$\text{Outcome}_{\hat{p}} = 1 + \text{A1pct} + \text{CAT\_T*Time} + \text{SAPA2*Time} + \text{SAPA3*Time} + \text{Sem*Phon*Time} + (1|\text{SubID}) + (1|\text{Tx}) + (1|\text{Item\_No}) \quad (4)$$

where

$\text{Outcome}_{\hat{p}}$	= the predicted outcome
1	= the model intercept
A1pct	= the individual's percent correct for naming the outcomes at baseline
Time	= assessment point (immediate post-treatment or follow-up)
CAT_T	= CAT comprehension T-score
SAPA2	= SAPA 2 total score
SAPA3	= SAPA 3 total score
Sem	= semantic processing score
Phon	= phonological processing score
(1|SubID)	= random intercept for subject
(1|Tx)	= random intercept to treatment group
(1|Item_No)	= random intercept for a given item

Note that the random intercept of treatment group only appeared in the whole group analyses. The asterisks in the model created an interaction term, so that CAT-T*Time meant "evaluate the effect of CAT_T for every value of Time." The software also evaluates the effects of each term in the interaction, as well. These models were reviewed with a statistician familiar with the data, who suggested that the models were overfitted (K. Cain., personal communication, May 21, 2020). Overfitting models results in convergence failures, which means that the model was unable to converge on a solution. False convergence may also occur, such that the model provides estimates, but these estimates are likely unreliable (i.e., the estimates do not reflect the accurate model predictions). Essentially, these convergence errors arise when there are

insufficient data, or the model is more complicated than what might actually be happening in the data. To avoid these situations, the sample should be as large as possible and the model should contain a minimal number of theoretically-motivated predictors (Fraundorf, 2018). For the sake of completeness, the original model was run for each proposed analysis and it did indeed fail to converge. To avoid convergence failures and to maximize model fit, the general model was modified as follows:

$$\text{Outcome}_{\hat{p}} = 1 + \text{A1pct} + \text{CAT\_T} + \text{SAPA2} + \text{SAPA3} + \text{Sem} + \text{Phon} + (1|\text{SubID}) + (1|\text{Tx}) + (1|\text{Item\_No}) \quad (5)$$

In this model, the interaction terms were removed, and individual models for immediate post-treatment and follow-up naming performance were conducted. In exploratory analyses of the data, all predictors had significant relationships with the data when entered singly. However, any relationships between the linguistic predictors and the outcomes were neutralized when baseline performance was entered into the model. Thus, to fulfill the study aims of determining the contribution of macro- and micro-linguistic abilities to naming abilities following treatment, the baseline naming variable was omitted from the analyses. The consequence of this change is that no inferences about predictor effects beyond baseline performance can be gleaned. This issue will be revisited in the limitations section. Thus, the final model for the analyses is represented by Equation 6.

$$\text{Outcome}_{\hat{p}} = 1 + \text{CAT\_T} + \text{SAPA2} + \text{SAPA3} + \text{Sem} + \text{Phon} + (1|\text{SubID}) + (1|\text{Tx}) + (1|\text{Item\_No}) \quad (6)$$

Given the large number of models, the Benjamini-Hochberg (Benjamini & Hochberg, 1995) correction procedure was used to control Type I error rate. The B-H procedure is a method for controlling false discovery rate by comparing each *p*-value from smallest to largest to its B-H critical value, calculated as *(i/m)Q*, where *i* = the rank of the obtained *p*-value, *m* = the total number of tests in the analysis, and *Q* = the false discovery rate. The false discovery rate is the

proportion of false positives that would be acceptable in the analysis results. This study was experimental in nature, and thus the false discovery rate was set at .10; that is, the "acceptable" proportion of false positive results was set at 10% (McDonald, 2014). A higher false discovery rate is recommended for exploratory studies, because the assumption is that future studies will confirm or reject the relationships between variables. In essence, any false positives identified using the B-H method become candidates for further exploration in subsequent analyses.

To ensure that the analyses are not compromised by collinearity before the model was evaluated, the predictors were run through correlational analyses (Figure 4). Collinearity occurs when two or more variables are too highly correlated with each other, meaning that they both seem to measure the same construct. This situation is problematic for regression analyses because multiple measures of the same skills reduce the likelihood that any of those predictors will uniquely contribute to an outcome. In essence, collinearity reduces the chances of finding significant relationships, even when they do exist. A general rule of thumb for determining whether a set of predictors demonstrate collinearity is if correlation coefficients are ±.90 or greater (Tabachnick & Fidell, 2013). As seen in the figures below, none of the coefficients exceed $r = .58$. Collinearity within models was measured by variance inflation factor (VIF) using the vif() command in the `car` package 3.0-8 for R (Weisberg & Fox, 2019). None of the predictors exhibited multicollinearity (all VIFs < 4.0) (Echambadi & Hess, 2007; Francoeur, 2013). [4]

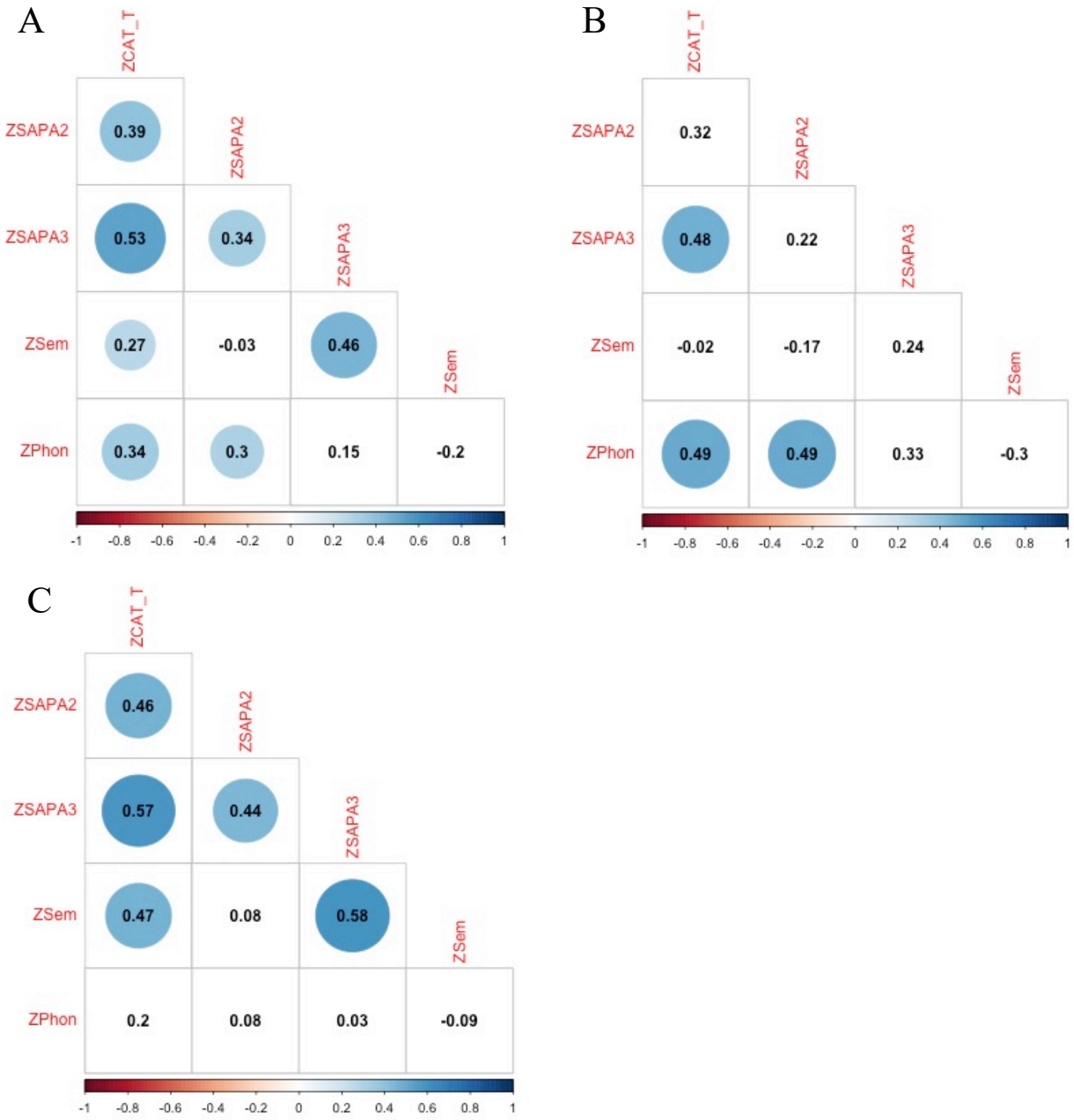
The regression analyses relied on the lme4 package in R Studio (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2013), using the glmer() function, specifically designed for

[4] Variance inflation factor is derived from R^2 and is calculated as $1/(1-R^2)$. A value of 1.0 suggests multicollinearity is not a problem in the model; values between 4.0-9.0 suggest that multicollinearity may pose a problem to the model; values >10.0 suggest that multicollinearity is a problem (Echambadi & Hess, 2007).

mixed-effects logistic regression. Other packages used in these analyses included the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017) to calculate *p*-values, the InformationValue package (Prabhakaran, 2016) to determine model sensitivity and specificity, and the car package (Fox et al., 2020) to estimate Cook's *D*. Packages used to visualize the data in table and figure formats included corrplot (Wei & Simko, 2017), sjPlot (Lüdecke, 2020), ggeffects (Lüdecke, 2018), and ggpubr (Kassambara, 2020). For the sake of completeness, mixed-effect linear regressions were also conducted, and these results were congruent with the findings discussed in the Results (see Appendix F for summaries). The R code for the logistic mixed-effect models is also available in Appendix G.

Figure 4

Predictor Correlations for Each Group Analysis



Note. Shaded correlations represent significant relationship at the α = .05 level. A = Correlations between predictors for the whole group analyses. B = Correlations between predictors for the PMT group analyses. C = Correlations between predictors for the SFA group analyses. ZCAT_T = CAT comprehension t-score; ZSAPA2 = SAPA 2; ZSAPA3 = SAPA 3; ZSem = semantic impairment score; Zphon = phonological impairment score.

Chapter 3: Results

The aim of this study was to determine the relative contribution of macro- and micro-linguistic variables at baseline to naming performance for treated, untreated-related, and untreated-unrelated items immediately following treatment and at follow-up. To answer questions related to these aims, mixed-effect logistic regressions with simultaneous predictor entry were used to predict naming outcomes among a sample of participants who completed either PMT or SFA treatment protocols. See Table 7 for a brief overview of the study findings. Second order Akaike Information Criterion (AIC_c)[5] analyses determined model selection, and the best model was compared to the null model (i.e., null hypothesis) using χ^2 tests. For brevity, only significant results will be interpreted in this section.

Table 7

Overview of Study Results

	Treated		Untreated-Related		Untreated-Unrelated	
	Imm. Post-Tx	Follow-Up	Imm. Post-Tx	Follow-Up	Imm. Post-Tx	Follow-Up
All	SAPA 3 Phon. Imp.	SAPA 3 Sem. Imp. Phon. Imp.	SAPA 3 Sem. Imp. Phon. Imp.	SAPA 3 Sem. Imp. Phon. Imp.	SAPA 3 Sem. Imp. Phon. Imp.	SAPA 3 Sem. Imp.
PMT	SAPA 3 Sem. Imp. Phon. Imp.	SAPA 3 Sem. Imp. Phon. Imp.	SAPA 3 Sem. Imp. Phon. Imp.	SAPA 3 Sem. Imp. Phon. Imp.	SAPA 3 Sem. Imp.	Sem. Imp.
SFA	SAPA 3	SAPA 3	Phon. Imp.	SAPA 3	SAPA 3 Phon. Imp.	SAPA 3

Note. SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA – Repetition, Parsing, and Blending Subtest; Sem. Imp. = semantic impairment score; Phon. Imp. = phonological impairment score.

[5] AIC_c, the criterion for small samples, is calculated as 2(log-likelihood) + 2K + (2K(K+1)/(n-K-1)), where n = the sample size, and K = the number of model parameters (Glen, 2015).

Whole Group Results

Research Question 1: *What is the relative contribution of baseline macro-linguistic input, output, and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities immediately post-treatment in a group of 58 PWA who received anomia treatment?*

The best-fit model for treated items had a significantly better model fit than the null model with no predictors, $\chi^2(5) = 56.78$, $p < 0.001$. Model sensitivity (the ability to correctly predict correct naming responses) was 81% and specificity (the ability to correctly predict incorrect naming responses) was 67%. In the model for untreated-related items, sensitivity was 81% and specificity was 66%; the model also was a significantly better fit that the null model, $\chi^2(5) = 47.89$, $p < 0.001$. The model for untreated-unrelated items was a significantly better fit to the data than the null model, $\chi^2(5) = 50.63$, $p < 0.001$. Finally, the untreated-unrelated model had 72% sensitivity and 76% specificity. None of the model intercepts were significantly different from zero (all B-H corrected *p*-values > .05). See Figure 5 for plots of the predicted probabilities of correct naming for each predictor and Tables 8-10 for detailed results.

In the model for treated items immediately post-treatment, both SAPA 3 (β = 0.79, (*SE* = 0.16), *Wald*(1) = 5.05, B-H corrected $p < .001$, *OR* = 2.20) and the phonological impairment score (β = 0.37, (*SE* = 0.13), *Wald*(1) = 2.78, B-H corrected $p = .023$, *OR* = 1.44) significantly predicted naming performance immediately post-treatment. Another way to interpret these results is that participants who were 1 *SD* above the mean on the SAPA 3 at baseline were 2.20 times more likely to correctly name a treated item than those at the mean, and had a mean 68% predicted probability of doing so (compared to 49% for those at the mean). For those whose phonological impairment scores were 1 *SD* above the mean at baseline, they had a 61% mean predicted probability of correctly naming treated items immediately post-treatment, and they

were 1.44 times more likely to correctly name items than those with average phonological impairment scores. Recall that this means that those who were more phonologically impaired at baseline were better at naming treated items following treatment than those who were less impaired. No other predictors had significant relationships with the outcome (all *p*s > .05).

In the model for untreated-related items, SAPA 3 (β = 0.64, (*SE* = 0.16), *Wald*(1) = 3.88, B-H corrected $p < .001$, *OR* = 1.89) and phonological impairment (β = 0.38, (*SE* = 0.14), *Wald*(1) = 2.70, B-H corrected $p = .026$, *OR* = 1.46) again significantly predicted probability of correctly naming items immediately post-treatment: a 1 *SD* increase in baseline scores equated to 63% and 60% predicted probability, respectively, of correctly naming items immediately post-treatment. Furthermore, semantic impairment at baseline also significantly predicted immediate post-treatment naming ability, β = 0.31, (*SE* = 0.15), *Wald*(1) = 2.09, B-H corrected $p < .094$, *OR* = 1.37, such that those with a 1 *SD* increase in impairment level had a 57% mean predicted probability of correct naming ability. No other predictors had significant relationships with the outcome (all *p*s > .05).

The results for the untreated-unrelated items mirrored the model for the untreated-related items: Baseline SAPA 3 performance continued to predict naming performance, β = 0.64, (*SE* = 0.16), *Wald*(1) = 3.96, B-H corrected $p < .001$, *OR* = 1.91, the semantic impairment score was a significant predictor, β = 0.46, (*SE* = 0.15), *Wald*(1) = 3.04, B-H corrected $p = .012$, *OR* = 1.58, and phonological impairment was also significant, β = 0.29, (*SE* = 0.14), *Wald*(1) = 2.07, B-H corrected $p = .094$, *OR* = 1.34. Said another way, those whose SAPA 3 scores were 1 *SD* above the mean at baseline had a 62% mean predicted probability of naming these items post-treatment. Likewise, those with baseline semantic impairment scores 1 *SD* above the mean had a 58% mean predicted probability of correctly naming untreated-unrelated items immediately post-treatment

and those with phonological impairment 1 *SD* above average had a mean predicted probability of 55% for correct naming. No other predictors had significant relationships with the outcome (all *p*s > .05).

Figure 5

Whole Group Predictors of Naming Performance Across Item Types Immediately Post-Treatment

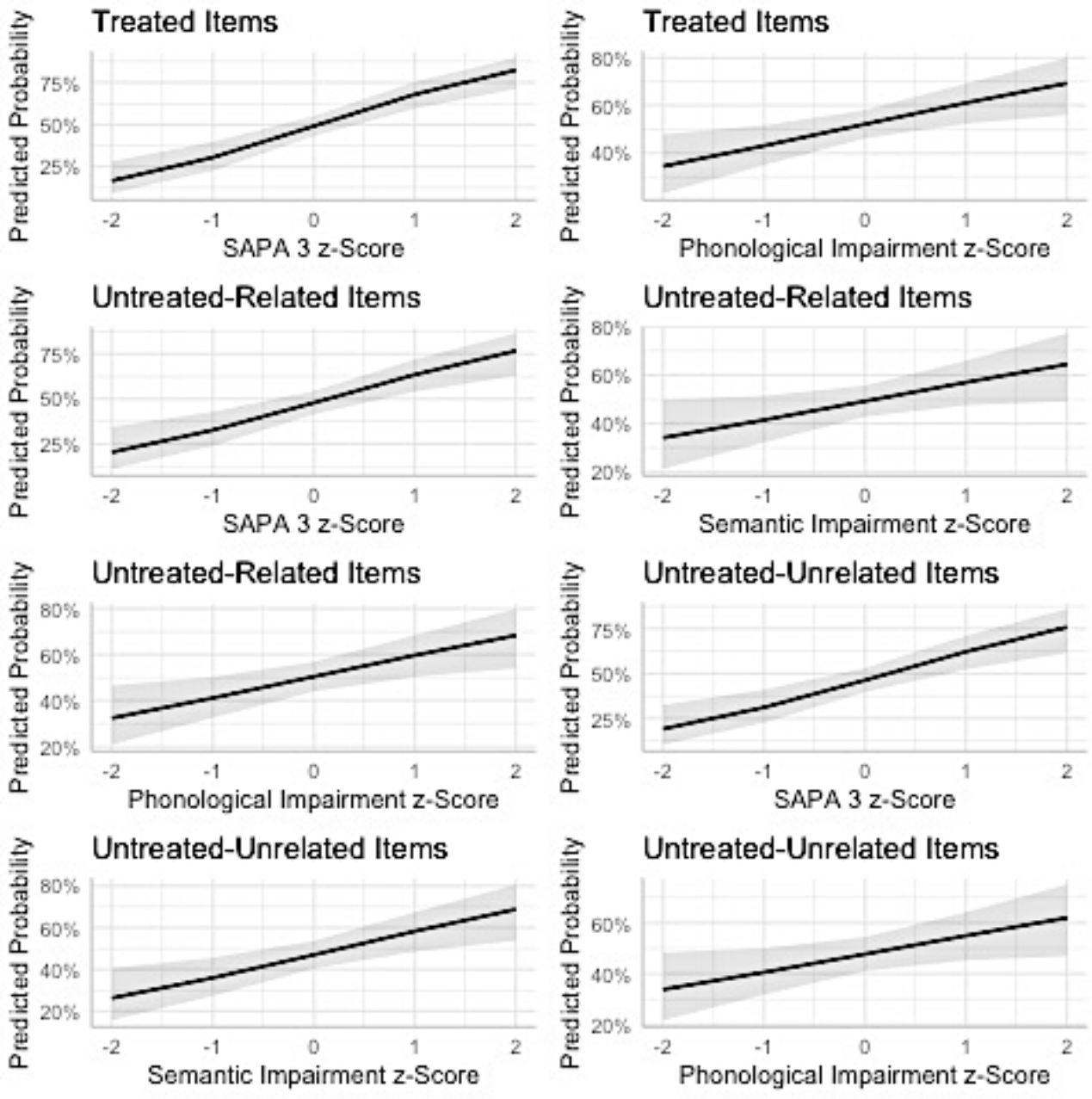


Table 8

Logistic Regression Results: Whole Group – Treated Items Immediately Post-Treatment

	All Treated Immed. Post-Tx Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.95	0.75 – 1.21	0.751	0.695	1.03	0.70 – 1.51	0.893
CAT T-score	1.13	0.83 – 1.55	0.534	0.430			
SAPA 2	1.17	0.89 – 1.53	0.391	0.268			
SAPA 3	2.20	1.62 – 2.98	**<0.001**	<0.001			
Semantic Impairment	1.28	0.97 – 1.69	0.177	0.085			
Phonological Impairment	1.44	1.11 – 1.86	**0.023**	0.006			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.16 $_{\text{Item\_No}}$				0.16 $_{\text{Item\_No}}$		
	0.67 $_{\text{SubID}}$				1.97 $_{\text{SubID}}$		
ICC	0.20				0.39		
N	55 $_{\text{SubID}}$				55 $_{\text{SubID}}$		
	95 $_{\text{Item\_No}}$				95 $_{\text{Item\_No}}$		
Observations	8684				8684		
Marginal R^2 / Conditional R^2	0.239 / 0.392				0.000 / 0.393		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 9

Logistic Regression Results: Whole Group – Untreated-Related Items Immediately Post-Treatment

	All Untreated-Related Immed. Post-Tx Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.90	0.69 – 1.17	0.534	0.428	0.96	0.65 – 1.41	0.835
CAT T-score	1.16	0.83 – 1.61	0.498	0.387			
SAPA 2	1.15	0.87 – 1.54	0.440	0.325			
SAPA 3	1.89	1.37 – 2.61	**<0.001**	<0.001			
Semantic Impairment	1.37	1.02 – 1.84	**0.094**	0.036			
Phonological Impairment	1.46	1.11 – 1.91	**0.026**	0.007			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	$0.15_{\text{Item\_No}}$				$0.15_{\text{Item\_No}}$		
	0.71_{SubID}				1.85_{SubID}		
ICC	0.21				0.38		
N	55_{SubID}				55_{SubID}		
	$56_{\text{Item\_No}}$				$56_{\text{Item\_No}}$		
Observations	4455				4455		
Marginal R^2 / Conditional R^2	0.212 / 0.376				0.000 / 0.378		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 10

Logistic Regression Results: Whole Group – Untreated-Unrelated Items Immediately Post-Treatment

	All Untreated-Unrelated Immed. Post-Tx Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.84	0.64 – 1.11	0.350	0.216	0.89	0.59 – 1.32	0.552
CAT T-score	1.16	0.84 – 1.61	0.485	0.373			
SAPA 2	1.14	0.86 – 1.51	0.485	0.372			
SAPA 3	1.91	1.39 – 2.63	**<0.001**	<0.001			
Semantic Impairment	1.58	1.18 – 2.12	**0.012**	0.002			
Phonological Impairment	1.34	1.02 – 1.76	**0.094**	0.038			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.69 $_{SubID}$				1.96 $_{SubID}$		
	0.12 $_{Item\_No}$				0.12 $_{Item\_No}$		
ICC	0.20				0.39		
N	55 $_{SubID}$				55 $_{SubID}$		
	25 $_{Item\_No}$				25 $_{Item\_No}$		
Observations	3762				3762		
Marginal R^2 / Conditional R^2	0.226 / 0.378				0.000 / 0.388		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Research Question 2: *What is the relative contribution of baseline macro-linguistic input, output, and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities 3 months post-treatment in a group of 58 PWA who received anomia treatment?*

The best-fit model for the treated items was a significantly better fit than the null model, $\chi^2(5) = 51.94$, $p < 0.001$, model sensitivity was 72%, and model specificity was 73%. The best-fit model for the untreated-related items was also a significantly better fit than the null model, $\chi^2(5) = 51.23$, $p < 0.001$. Sensitivity was 79% and specificity was 70%. The model for untreated-unrelated items was significantly better than the null model, $\chi^2(5) = 46.44$, $p < 0.001$, and had sensitivity of 70% and specificity of 78%. See Figure 6 for plots of the predicted probabilities of the significant independent variables and Tables 11-13 for details of each model.

In the model for naming of treated items at the follow-up timepoint, the model intercept was significantly different from zero, β = -0.37, (*SE* = 0.11), *Wald*(1) = -3.32, B-H corrected *p* = .007, *OR* = 0.69, and participants only had a 41% mean predicted probability of correctly naming, holding all other variables constant. Essentially, in this context, a significant intercept means that participants were correctly naming items significantly below a chance level. The SAPA 3 significantly predicted ability to name treated items at follow-up, β = 0.64, (*SE* = 0.15), *Wald*(1) = 4.39, B-H corrected *p* < .001, *OR* = 1.90, as did semantic impairment, β = 0.28, (*SE* = 0.13), *Wald*(1) = 2.06, B-H corrected *p* = .094, *OR* = 1.32, and phonological impairment, β = 0.32, (*SE* = 0.12), *Wald*(1) = 2.58, B-H corrected *p* = .036, *OR* = 1.37. Another way to interpret these results is that participants who were 1 *SD* above average at baseline for SAPA 3 scores were 1.90 times more likely to correctly name items or had a 57% predicted probability of doing so. Likewise, those with higher semantic and phonological impairment were 1.32 times and 1.37

times more likely to correctly name items and had 50% and 52% predicted probabilities of doing so, respectively. No other predictors had significant relationships with the outcome (all *p*s > .05).

In the model for the untreated-related items, the intercept was significantly different from zero, β = -0.38, (*SE* = 0.14), *Wald*(1) = -2.74, B-H corrected *p* = .023, *OR* = 0.68, such that participants only had a 41% predicted probability of correctly naming items at follow-up. Again, SAPA 3, β = 0.69, (*SE* = 0.17), *Wald*(1) = 4.10, B-H corrected *p* < .001, *OR* = 2.00, semantic impairment, β = 0.35, (*SE* = 0.15), *Wald*(1) = 2.24, B-H corrected *p* = .073, *OR* = 1.41, and phonological impairment (β = 0.42, (*SE* = 0.14), *Wald*(1) = 2.95, B-H corrected *p* = .015, *OR* = 1.53) were all significant predictors of correct naming. In other words, participants with SAPA 3 scores that were 1 *SD* above average at baseline had a 58% mean predicted probability of correctly naming the untreated-related items, while those with greater semantic impairment had a 52% predicted probability and those with greater phonological impairment had a 54% predicted probability of correctly naming items. No other predictors had significant relationships with the outcome (all *p*s > .05).

The intercept for the model of untreated-unrelated items was not significant (*p* < .05). Significant predictors of correct naming at follow-up were the SAPA 3 (β = 0.55, (*SE* = 0.16), *Wald*(1) = 3.35, B-H corrected *p* = .007, *OR* = 1.73) and semantic impairment, β = 0.56, (*SE* = 0.15), *Wald*(1) = 3.59, B-H corrected *p* < .001, *OR* = 1.74. Participants who had SAPA 3 scores 1 *SD* above the mean at baseline were 1.73 times more likely to correctly name untreated-unrelated items at follow-up and they had a 61% predicted probability of doing so. Similarly, those with greater semantic impairment at baseline had a 62% predicted probability of correctly naming items and were 1.74 times more likely to do so than those with average semantic impairment levels. Unlike the other models discussed thus far, phonological impairment levels at

baseline did not have a significant relationship with the outcome ($p > .05$), and neither did any other predictors (all $ps > .05$).

Figure 6

Whole Group Predictors of Naming Performance Across Item Types At Follow-Up

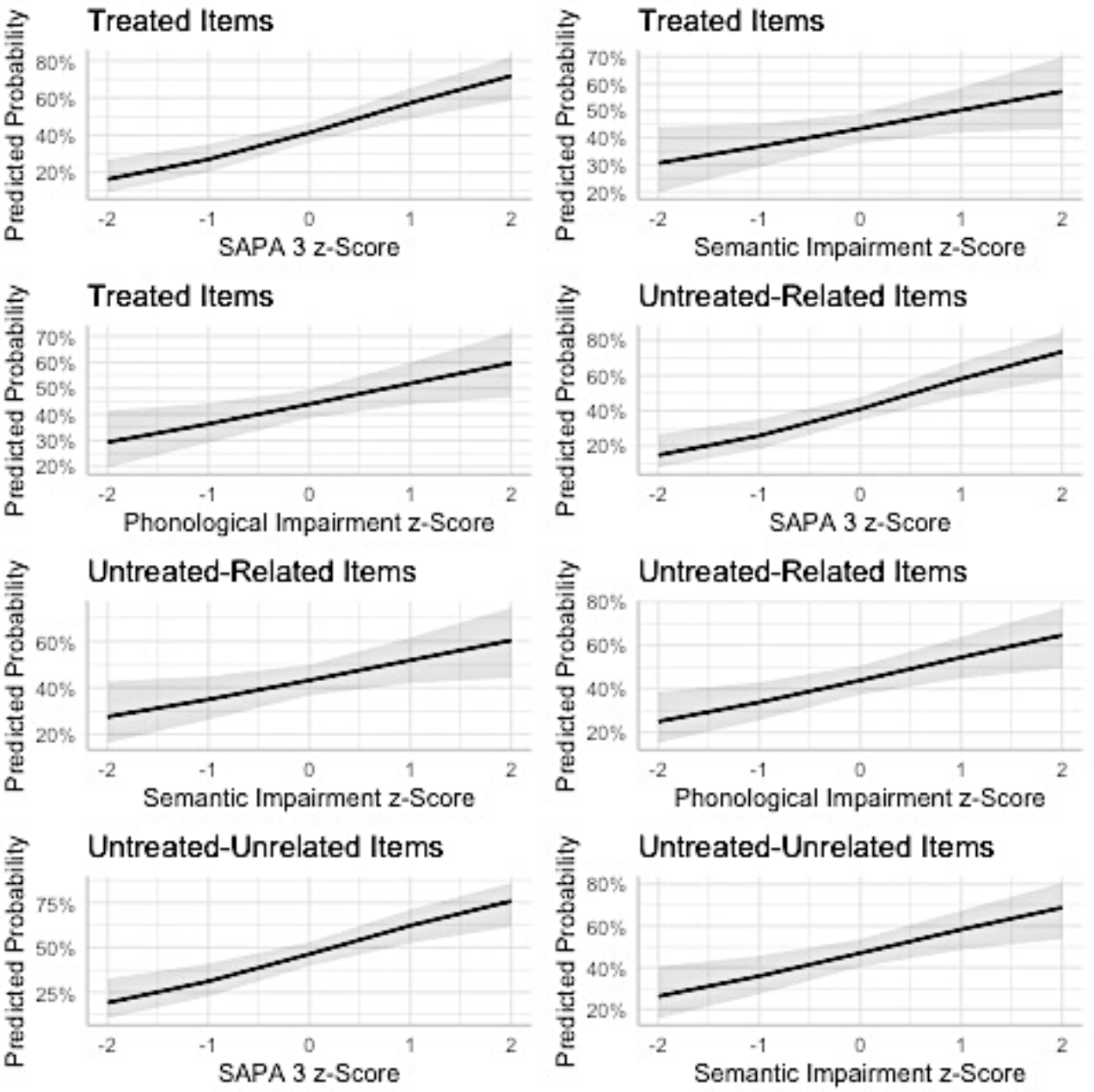


Table 11

Logistic Regression Results: Whole Group – Treated Items at Follow-Up

	All Treated Follow-Up Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.69	0.55 – 0.86	**0.007**	0.001	0.73	0.52 – 1.03	0.075
CAT T-score	1.10	0.82 – 1.48	0.606	0.511			
SAPA 2	1.20	0.93 – 1.55	0.306	0.167			
SAPA 3	1.90	1.43 – 2.53	**<0.001**	<0.001			
Semantic Impairment	1.32	1.01 – 1.72	**0.094**	0.039			
Phonological Impairment	1.37	1.08 – 1.75	**0.036**	0.010			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.09 $_{Item\_No}$				0.09 $_{Item\_No}$		
	0.59 $_{SubID}$				1.60 $_{SubID}$		
ICC	0.17				0.34		
N	55 $_{SubID}$				55 $_{SubID}$		
	95 $_{Item\_No}$				95 $_{Item\_No}$		
Observations	8450				8450		
Marginal R^2 / Conditional R^2	0.203 / 0.340				0.000 / 0.339		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 12

Logistic Regression Results: Whole Group – Untreated-Related Items at Follow-Up

	All Untreated-Related Follow-Up Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.68	0.52 – 0.90	**0.023**	0.006	0.73	0.48 – 1.10	0.13
CAT T-score	1.03	0.73 – 1.45	0.903	0.870			
SAPA 2	1.34	0.99 – 1.80	0.121	0.055			
SAPA 3	2.00	1.44 – 2.78	**<0.001**	<0.001			
Semantic Impairment	1.41	1.04 – 1.92	**0.073**	0.025			
Phonological Impairment	1.53	1.15 – 2.03	**0.015**	0.003			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	$0.20_{\text{Item\_No}}$				$0.20_{\text{Item\_No}}$		
	0.74_{SubID}				2.09_{SubID}		
ICC	0.22				0.41		
N	55_{SubID}				55_{SubID}		
	$56_{\text{Item\_No}}$				$56_{\text{Item\_No}}$		
Observations	4331				4331		
Marginal R^2 / Conditional R^2	0.233 / 0.404				0.000 / 0.411		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 13

Logistic Regression Results: Whole Group – Untreated-Unrelated Items at Follow-Up

	All Untreated-Unrelated Follow-Up Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.87	0.65 – 1.15	0.440	0.326	0.92	0.61 – 1.37	0.678
CAT T-score	1.14	0.81 – 1.60	0.564	0.465			
SAPA 2	1.20	0.89 – 1.61	0.356	0.224			
SAPA 3	1.73	1.26 – 2.39	**0.007**	0.001			
Semantic Impairment	1.74	1.29 – 2.36	**<0.001**	<0.001			
Phonological Impairment	1.21	0.92 – 1.59	0.315	0.175			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.71 $_{SubID}$				1.90 $_{SubID}$		
	0.13 $_{Item\_No}$				0.13 $_{Item\_No}$		
ICC	0.2				0.38		
N	54 $_{SubID}$				54 $_{SubID}$		
	25 $_{Item\_No}$				25 $_{Item\_No}$		
Observations	3612				3612		
Marginal R^2 / Conditional R^2	0.205 / 0.367				0.000 / 0.381		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

PMT Group Results

Research Question 3: *What is the relative contribution of baseline macro-linguistic input, output, and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities immediately post-treatment in a group of 28 PWA who were randomized to receive a phonologically based (PMT) treatment?*

The best-fit model for treated items in the PMT group fit the data significantly better than the null model, $\chi^2(5) = 41.98$, $p < 0.001$, and model sensitivity and specificity were 78% and 75%, respectively. The model for untreated-related items was also a significantly better fit to the data than the null model, $\chi^2(5) = 30.50$, $p < 0.001$. The model had 69% sensitivity and 81% specificity. Finally, the untreated-unrelated items model was, unsurprisingly, a significantly better than the null model, $\chi^2(5) = 19.40$, $p = 0.002$. The model was 59% sensitive and 81% specific. See Figure 7 for plots of the predicted probabilities and Tables 14-16 for detailed results.

In the model for treated items, the intercept was not significant ($p > .05$). Those who scored 1 *SD* higher than the mean on the SAPA 3 at baseline ($\beta = 0.81$, (*SE* = 0.18), *Wald*(1) = 4.62, B-H corrected $p < .001$, *OR* = 2.25) had a 69% mean predicted probability of correctly naming items post-treatment. Increased semantic impairment ($\beta = 0.46$, (*SE* = 0.15), *Wald*(1) = 3.01, B-H corrected $p = .015$, *OR* = 1.59) and phonological impairment ($\beta = 0.61$, (*SE* = 0.18), *Wald*(1) = 3.35, B-H corrected $p = .007$, *OR* = 1.85) related to mean predicted probabilities of 62% and 65% of correctly naming treated items immediately post-treatment. No other predictors had significant relationships with the outcome (all *p*s > .05).

The untreated-related model yielded similar results to the treated item model. The intercept was not significant ($p > .05$), and again baseline SAPA 3 performance was significantly

related to naming ability immediately post treatment, β = 0.66, (*SE* = 0.20), *Wald*(1) = 3.31, B-H corrected *p* = .007, *OR* = 1.93, such that participants with higher scores had a 64% predicted probability of correctly naming items. For semantic impairment, β = 0.44, (*SE* = 0.17), *Wald*(1) = 2.52, B-H corrected *p* = .041, *OR* = 1.55, participants who scored 1 *SD* above average at baseline had mean predicted probabilities of 60% (immediately post-treatment). Finally, increasing phonological impairment (β = 0.49, (*SE* = 0.21), *Wald*(1) = 2.37, B-H corrected *p* = .056, *OR* = 1.64) also related to increased naming ability post-treatment: PWA who had average phonological impairment had a 49% probability of correctly naming untreated-related items post-treatment, while those who were 1 *SD* above the mean had 62% predicted probability. No other predictors had significant relationships with the outcome (all *p*s > .05).

In the model of the untreated-unrelated items, the intercept was significant, β = -0.34, (*SE* = 0.16), *Wald*(1) = -2.08, B-H corrected *p* = .094, *OR* = 0.71; holding all other predictors constant, participants only had a 42% mean predicted probability of correctly naming items. The SAPA 3 (β = 0.40, (*SE* = 0.19), *Wald*(1) = 2.08, B-H corrected *p* = .094, *OR* = 1.49) and semantic impairment score (β = 0.42, (*SE* = 0.17), *Wald*(1) = 2.46, B-H corrected *p* = .044, *OR* = 1.53) were the only significant predictors of naming performance on untreated-unrelated items immediately post-treatment. Specifically, participants whose SAPA 3 scores were 1 *SD* above the mean were 1.49 times more likely to correctly name items, and they had a 52% mean predicted probability of doing so. Likewise, those whose impairment scores were 1 *SD* above the mean were 1.53 times more likely to correctly name items – and had a 53% mean predicted probability of doing so – than those who had average semantic impairment. No other predictors had significant relationships with the outcome (all *p*s > .05).

Figure 7

PMT Group Predictors of Naming Performance Across Item Types Immediately Post-Treatment

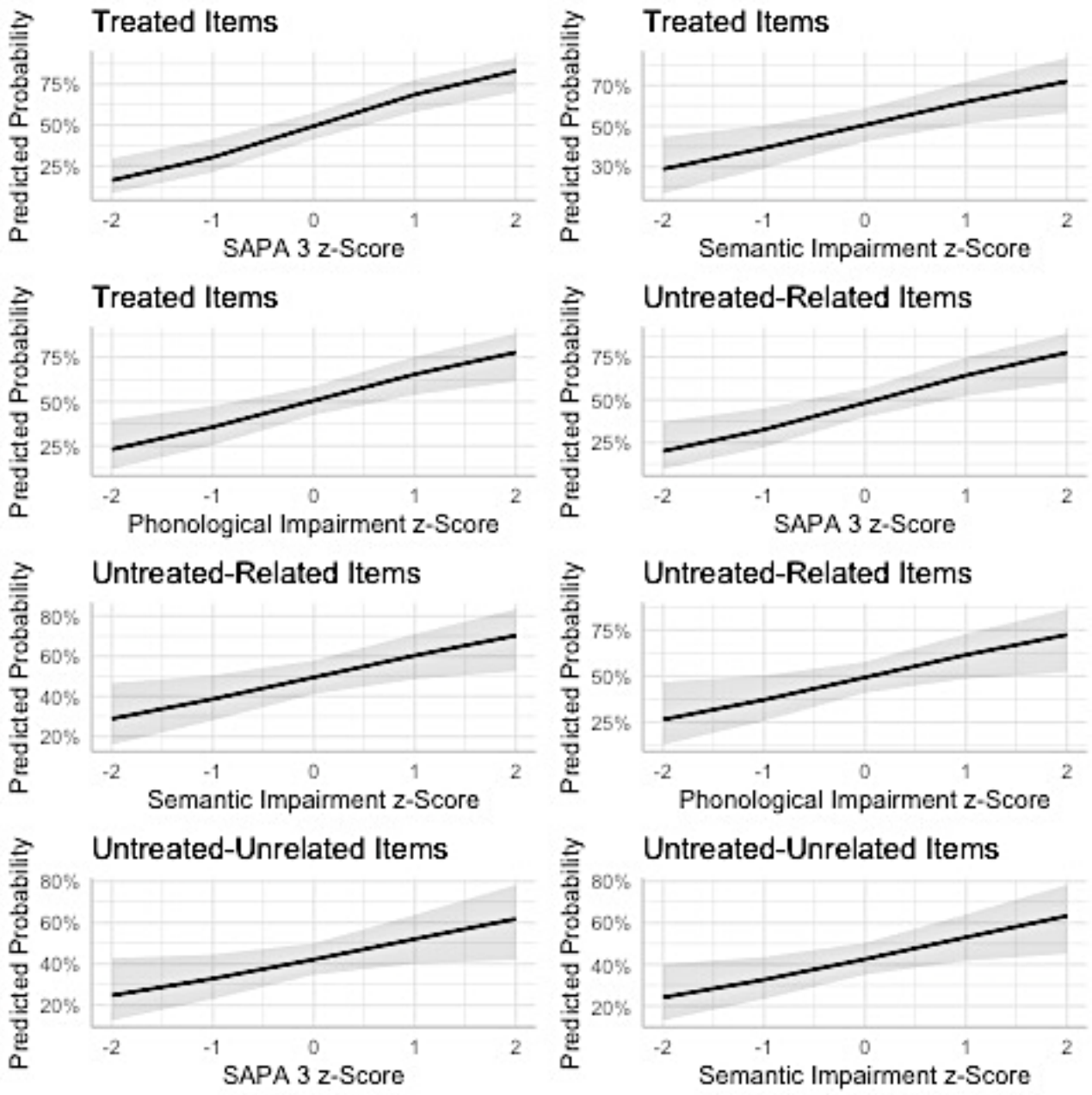


Table 14

Logistic Regression Results: PMT Group – Treated Items Immediately Post-Treatment

	PMT Treated Immed. Post-Tx Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.96	0.69 – 1.35	0.879	0.83	1.04	0.56 – 1.93	0.911
CAT T-score	1.30	0.90 – 1.87	0.294	0.158			
SAPA 2	0.99	0.72 – 1.36	0.946	0.946			
SAPA 3	2.25	1.59 – 3.17	**<0.001**	<0.001			
Semantic Impairment	1.59	1.17 – 2.14	**0.015**	0.003			
Phonological Impairment	1.85	1.29 – 2.65	**0.007**	0.001			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	$0.43_{\text{Item\_No}}$				$0.43_{\text{Item\_No}}$		
	0.41_{SubID}				2.33_{SubID}		
ICC	0.2				0.46		
N	27_{SubID}				27_{SubID}		
	$39_{\text{Item\_No}}$				$39_{\text{Item\_No}}$		
Observations	2886				2886		
Marginal R^2 / Conditional R^2	0.297 / 0.440				0.000 / 0.457		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 15

Logistic Regression Results: PMT Group – Untreated-Related Items Immediately Post-Treatment

	PMT Untreated-Related Immed. Post-Tx Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.92	0.66 – 1.30	0.736	0.647	0.98	0.56 – 1.72	0.953
CAT T-score	1.32	0.87 – 2.00	0.343	0.197			
SAPA 2	0.98	0.68 – 1.41	0.917	0.909			
SAPA 3	1.93	1.31 – 2.85	**0.007**	0.001			
Semantic Impairment	1.55	1.10 – 2.19	**0.041**	0.012			
Phonological Impairment	1.64	1.09 – 2.48	**0.056**	0.018			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.52 $_{SubID}$				1.92 $_{SubID}$		
	0.15 $_{Item\_No}$				0.15 $_{Item\_No}$		
ICC	0.17				0.39		
N	27 $_{SubID}$				27 $_{SubID}$		
	21 $_{Item\_No}$				21 $_{Item\_No}$		
Observations	1533				1533		
Marginal R^2 / Conditional R^2	0.255 / 0.380				0.000 / 0.386		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 16

Logistic Regression Results: PMT Group – Untreated-Unrelated Items Immediately Post-Treatment

	PMT Untreated-Unrelated Immed. Post-Tx Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.71	0.52 – 0.98	**0.094**	0.038	0.76	0.49 – 1.17	0.214
CAT T-score	1.18	0.79 – 1.78	0.527	0.415			
SAPA 2	1.34	0.94 – 1.93	0.214	0.109			
SAPA 3	1.49	1.02 – 2.17	**0.094**	0.038			
Semantic Impairment	1.53	1.09 – 2.14	**0.044**	0.014			
Phonological Impairment	1.11	0.74 – 1.66	0.706	0.608			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.52 $_{SubID}$				1.15 $_{SubID}$		
	0.10 $_{Item\_No}$				0.10 $_{Item\_No}$		
ICC	0.16				0.27		
N	27 $_{SubID}$				27 $_{SubID}$		
	25 $_{Item\_No}$				25 $_{Item\_No}$		
Observations	1850				1850		
Marginal R^2 / Conditional R^2	0.131 / 0.267				0.000 / 0.275		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Research Question 4: *What is the relative contribution of baseline macro-linguistic input, output, and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities 3 months post-treatment in a group of 28 PWA who were randomized to receive a phonologically based (PMT) treatment?*

The best-fit model of treated item naming at follow-up was a significantly better fit to the data than the null model, $\chi^2(5) = 36.57$, $p < 0.001$. Sensitivity for this model was 69% and specificity was 77%. The best-fit model of untreated-related item naming was a significantly better fit than the null model, $\chi^2(5) = 35.53$, $p < 0.001$; sensitivity and specificity were both 77%. Finally, the best-fit model for the untreated-unrelated items was also significantly better than the null model, $\chi^2(5) = 21.19$, $p = 0.001$. Sensitivity was 62% and specificity was 77%. See Figure 8 for a plots of the predicted probabilities of the significant variables and see Tables 17-19 for model details.

The intercept for the treated item naming model was significant, β = -0.39, (*SE* = 0.14), *Wald*(1) = -2.81, B-H corrected *p* = .021, *OR* = 0.68; participants only had a 40% mean predicted probability of correctly naming items at follow-up, holding all other variables constant. The SAPA 3, β = 0.56, (*SE* = 0.16), *Wald*(1) = 3.43, B-H corrected *p* = .007, *OR* = 1.74, semantic impairment, β = 0.42, (*SE* = 0.14), *Wald*(1) = 2.91, B-H corrected *p* = .017, *OR* = 1.52, and phonological impairment (β = 0.43, (*SE* = 0.17), *Wald*(1) = 2.53, B-H corrected *p* = .038, *OR* = 1.54) all significantly predicted naming performance at the follow-up point. For the SAPA 3, this corresponds to a 55% mean predicted probability of correct naming at follow-up for participants whose scores were 1 *SD* above the mean. Those with greater semantic and phonological impairment at baseline both had a 52% predicted probability of correctly naming treated items. No other predictors had significant relationships with the outcome (all *p*s > .05).

After correcting for multiple comparisons, the intercept for the model of untreated-related items at follow-up was not significant (uncorrected $p = .046$, B-H corrected $p = .106$). In either case, participants correctly named items at a below chance level (mean predicted probability was 40%). Again, the SAPA 3, $\beta = 0.70$, ($SE = 0.20$), $Wald(1) = 3.56$, B-H corrected $p < .001$, $OR = 2.01$, semantic impairment, $\beta = 0.49$, ($SE = 0.17$), $Wald(1) = 2.84$, B-H corrected $p = .017$, $OR = 1.64$, and phonological impairment $\beta = 0.50$, ($SE = 0.21$), $Wald(1) = 2.34$, B-H corrected $p = .057$, $OR = 1.63$) significantly predicted correct naming of untreated-related items at follow-up. PWA who scored 1 *SD* above the mean on the SAPA 3 at baseline were twice as likely to correctly name items as those whose scores were at the mean and they had a 59% predicted probability of doing so. Participants whose semantic or phonological impairments were 1 *SD* above the mean had 55% and 54% mean predicted probabilities, respectively, of correctly naming items at follow-up than participants with mean impairment levels at baseline. These more impaired individuals were also 1.64 (for semantic impairment) and 1.63 (for phonological impairment) times more likely to correctly name items than those with less impairment. No other predictors had significant relationships with the outcome (all ps > .05).

The intercept for the untreated-unrelated item model was significant, $\beta = -0.35$, ($SE = 0.17$), $Wald(1) = -2.08$, B-H corrected $p = .097$, $OR = 0.70$, and the participants only had a 41% mean predicted probability of correctly naming these items at the follow-up point. After applying the Benjamini-Hochberg correction, the only significant predictor in the model was semantic impairment level, $\beta = 0.58$, ($SE = 0.18$), $Wald(1) = 3.25$, B-H corrected $p = .007$, $OR = 1.78$. That is, people with semantic impairment 1 *SD* above the mean at baseline were 1.78 times more likely to correctly name untreated-unrelated items at the follow-up point than those with average semantic impairment levels, and these more impaired PWA had a 57% mean predicted

probability of doing so. It is also worth noting that, before the *p*-value correction, the SAPA 2 also significantly predicted follow-up naming performance, β = 0.56, (*SE* = 0.16), *Wald*(1) = 3.43, uncorrected *p* = .045, B-H corrected *p* = .106, *OR* = 1.47, but this relationship was no longer significant after the correction. Although this relationship should be interpreted with caution due to the non-significant relationship, PWA who scored 1 *SD* above the mean on the SAPA 2 at baseline were 1.47 times more likely to correctly name untreated-unrelated items than those with average SAPA 2 scores. No other predictors had significant relationships with the outcome (all *p*s > .05).

Figure 8

PMT Group Predictors of Naming Performance Across Item Types At Follow-Up

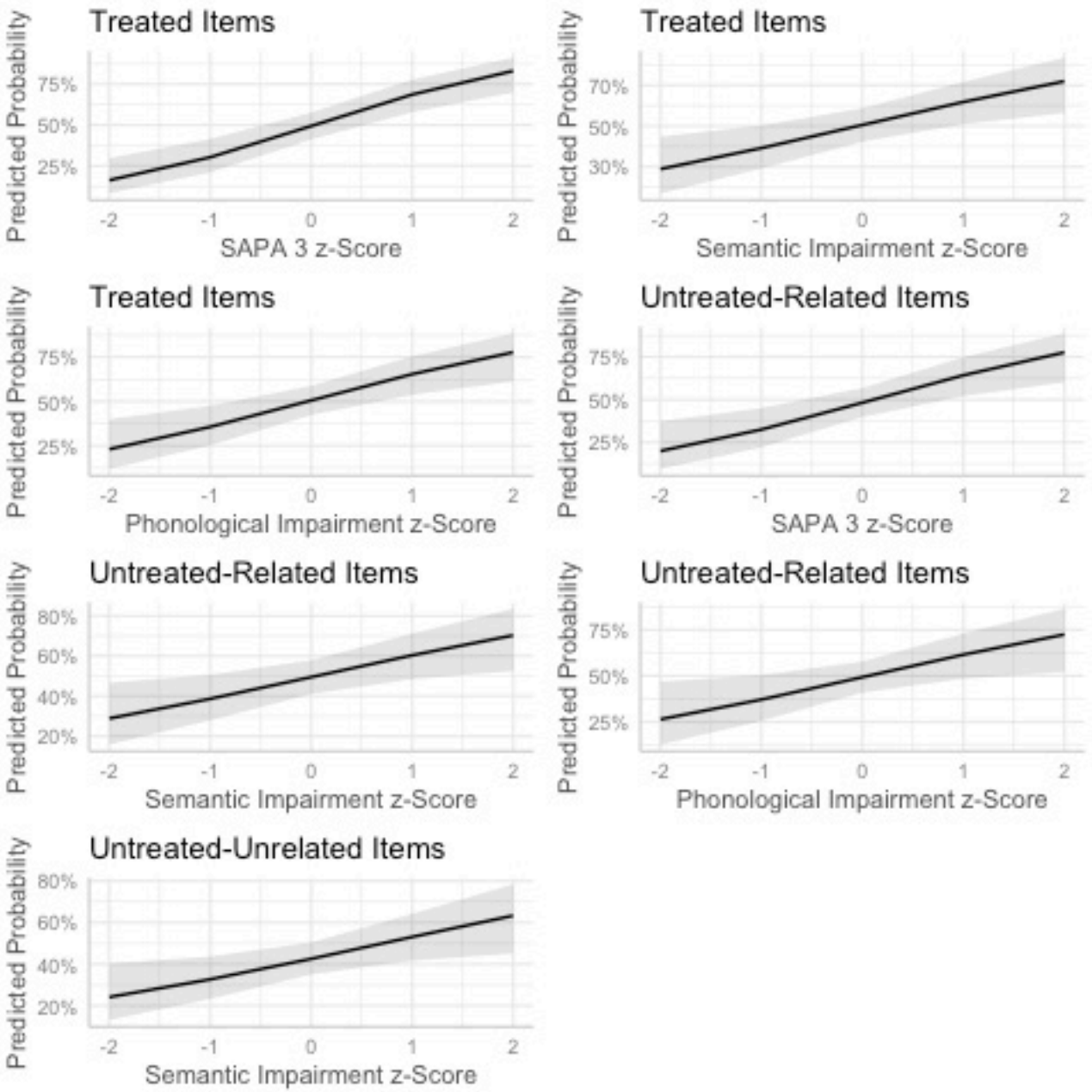


Table 17

Logistic Regression Results: PMT Group – Treated Items at Follow-Up

	PMT Treated Follow-Up Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.68	0.51 – 0.89	**0.021**	0.005	0.72	0.44 – 1.20	0.210
CAT T-score	1.40	0.99 – 1.97	0.121	0.055			
SAPA 2	1.12	0.83 – 1.52	0.555	0.452			
SAPA 3	1.74	1.27 – 2.39	**0.007**	0.001			
Semantic Impairment	1.52	1.15 – 2.02	**0.017**	0.004			
Phonological Impairment	1.54	1.10 – 2.16	**0.038**	0.011			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.14 $_{Item\_No}$				0.15 $_{Item\_No}$		
	0.36 $_{SubID}$				1.62 $_{SubID}$		
ICC	0.13				0.35		
N	27 $_{SubID}$				27 $_{SubID}$		
	39 $_{Item\_No}$				39 $_{Item\_No}$		
Observations	2681				2681		
Marginal R^2 / Conditional R^2	0.243 / 0.343				0.000 / 0.349		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 18

Logistic Regression Results: PMT Group – Untreated-Related Items at Follow-Up

	PMT Untreated-Related Follow-Up Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.67	0.45 – 0.99	0.106	0.046	0.73	0.38 – 1.40	0.343
CAT T-score	1.39	0.92 – 2.10	0.231	0.122			
SAPA 2	1.23	0.85 – 1.78	0.393	0.279			
SAPA 3	2.01	1.37 – 2.95	**<0.001**	<0.001			
Semantic Impairment	1.64	1.17 – 2.31	**0.017**	0.004			
Phonological Impairment	1.63	1.08 – 2.46	**0.057**	0.019			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.48_{SubID}				2.36_{SubID}		
	$0.37_{\text{Item\_No}}$				$0.38_{\text{Item\_No}}$		
ICC	0.21				0.45		
N	27_{SubID}				27_{SubID}		
	$21_{\text{Item\_No}}$				$21_{\text{Item\_No}}$		
Observations	1442				1442		
Marginal R^2 / Conditional R^2	0.294 / 0.440				0.000 / 0.454		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 19

Logistic Regression Results: PMT Group – Untreated-Unrelated Items at Follow-Up

	PMT Untreated-Unrelated Follow-Up Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.70	0.51 – 0.98	**0.094**	0.037	0.76	0.47 – 1.21	0.243
CAT T-score	1.09	0.72 – 1.67	0.745	0.676			
SAPA 2	1.47	1.01 – 2.13	0.106	0.045			
SAPA 3	1.37	0.93 – 2.02	0.218	0.113			
Semantic Impairment	1.78	1.26 – 2.52	**0.007**	0.001			
Phonological Impairment	1.26	0.83 – 1.91	0.393	0.279			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.55_{SubID}				1.35_{SubID}		
	$0.11_{\text{Item\_No}}$				$0.11_{\text{Item\_No}}$		
ICC	0.17				0.31		
N	27_{SubID}				27_{SubID}		
	$25_{\text{Item\_No}}$				$25_{\text{Item\_No}}$		
Observations	1712				1712		
Marginal R^2 / Conditional R^2	0.156 / 0.295				0.000 / 0.307		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

SFA Group Results

Research Question 5: *What is the relative contribution of baseline macro-linguistic input, output, and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities immediately post-treatment in a group of 30 PWA who were randomized to receive a semantically based (SFA) treatment?*

It is worth noting at the outset that the residuals for fitted models were more variable in the SFA group than in the whole or PMT groups. This observation resulted in more participants – or influential data points – being removed from the analyses than in the PMT group and, while this issue did not seem to impact the models for the main analysis, it will become more relevant in the presentation of the *post hoc* analysis results. This issue will be taken up in the limitations section. In any event, the models for the SFA group were still good fits for the data, as measured by the AIC and χ^2 analyses. The best-fit models for treated ($\chi^2(5) = 26.26$, $p < 0.001$), untreated-related ($\chi^2(5) = 24.92$, $p = 0.001$), and untreated-unrelated ($\chi^2(5) = 36.63$, $p < 0.001$) items were all significantly better fits than the null models. Furthermore, the sensitivity and specificity of each model was consistent with that of the models above: 82% and 63% (treated items), 80% and 64% (untreated-related items), and 80% and 75% (untreated-unrelated items), respectively. None of the model intercepts were significant (all *ps* > .05). See Figure 9 for plots of significant mean predicted probabilities across item types and Tables 20-22 for detailed results.

The model for naming of treated items immediately post-treatment revealed significant effects of the SAPA 3 ($\beta = 0.69$, (*SE* = 0.24), *Wald*(1) = 2.95, B-H corrected $p = .015$, *OR* = 2.00). That is, those participants whose SAPA 3 scores at baseline were 1 *SD* above the mean had a 66% mean predicted probabilities of correctly naming treated items immediately post-treatment. In the model for untreated-related items, only phonological impairment significantly

predicted naming accuracy, β = 0.47, (*SE* = 0.19), *Wald*(1) = 2.46, B-H corrected *p* = .044, *OR* = 1.61, which corresponds to PWA with greater impairment having a 63% predicted probability of correctly naming item post-treatment. Finally, in the model for untreated-unrelated items, SAPA 3 (β = 1.02, (*SE* = 0.24), *Wald*(1) = 4.20, B-H corrected *p* < .001, *OR* = 2.78) and phonological impairment (β = 0.51, (*SE* = 0.17), *Wald*(1) = 2.94, B-H corrected *p* = .015, *OR* = 1.66 significantly predicted post-treatment naming accuracy. Said another way, for participants whose SAPA 3 scores were 1 *SD* above the mean, they had a 73% mean predicted probability of correctly naming items, while those with greater phonological impairment had a 65% probability of doing so. No other predictors in these models demonstrated significant relationships with the outcomes (all *p*s > .05).

Figure 9

SFA Group Predictors of Naming Performance Across Item Types Immediately Post-Treatment

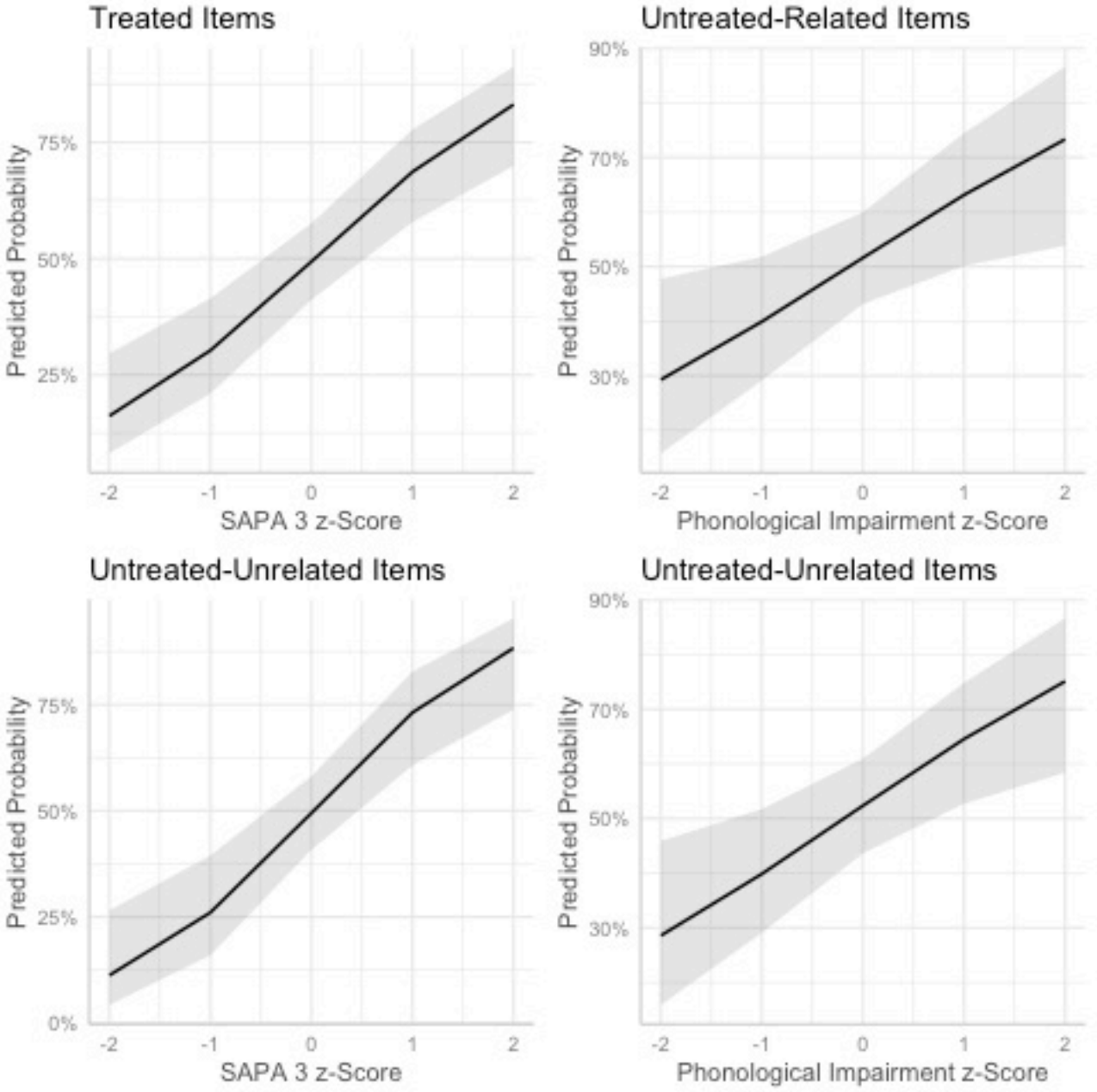


Table 20

Logistic Regression Results: SFA Group – Treated Items Immediately Post-Treatment

	SFA Treated Immed. Post-Tx Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.92	0.66 – 1.26	0.696	0.593	0.99	0.60 – 1.63	0.961
CAT T-score	0.95	0.60 – 1.52	0.885	0.844			
SAPA 2	1.48	0.99 – 2.23	0.123	0.058			
SAPA 3	2.00	1.26 – 3.18	**0.015**	0.003			
Semantic Impairment	1.27	0.83 – 1.95	0.391	0.265			
Phonological Impairment	1.24	0.89 – 1.72	0.349	0.21			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	$0.11_{\text{Item\_No}}$				$0.11_{\text{Item\_No}}$		
	0.66_{SubID}				1.77_{SubID}		
ICC	0.19				0.36		
N	28_{SubID}				28_{SubID}		
	$81_{\text{Item\_No}}$				$81_{\text{Item\_No}}$		
Observations	5720				5720		
Marginal R^2 / Conditional R^2	0.217 / 0.366				0.000 / 0.363		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 21

Logistic Regression Results: SFA Group – Untreated-Related Items Immediately Post-Treatment

	SFA Untreated-Related Immed. Post-Tx Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.93	0.66 – 1.32	0.751	0.69	0.99	0.59 – 1.65	0.956
CAT T-score	1.12	0.71 – 1.77	0.709	0.617			
SAPA 2	1.48	0.99 – 2.21	0.123	0.057			
SAPA 3	1.38	0.84 – 2.28	0.344	0.202			
Semantic Impairment	1.43	0.94 – 2.18	0.192	0.096			
Phonological Impairment	1.61	1.10 – 2.34	**0.044**	0.014			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.19 $_{Item\_No}$				0.19 $_{Item\_No}$		
	0.61 $_{SubID}$				1.66 $_{SubID}$		
ICC	0.2				0.36		
N	27 $_{SubID}$				27 $_{SubID}$		
	41 $_{Item\_No}$				41 $_{Item\_No}$		
Observations	2880				2880		
Marginal R^2 / Conditional R^2	0.198 / 0.355				0.000 / 0.360		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 22

Logistic Regression Results: SFA Group – Untreated-Unrelated Items Immediately Post-Treatment

	SFA Untreated-Unrelated Immed. Post-Tx Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.97	0.69 – 1.38	0.906	0.886	1.03	0.55 – 1.93	0.921
CAT T-score	1.23	0.78 – 1.96	0.485	0.373			
SAPA 2	0.95	0.63 – 1.43	0.868	0.812			
SAPA 3	2.78	1.73 – 4.48	**<0.001**	<0.001			
Semantic Impairment	1.30	0.84 – 2.01	0.364	0.234			
Phonological Impairment	1.66	1.18 – 2.32	**0.015**	0.003			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.59 $_{SubID}$				2.56 $_{SubID}$		
	0.16 $_{Item\_No}$				0.17 $_{Item\_No}$		
ICC	0.19				0.45		
N	28 $_{SubID}$				28 $_{SubID}$		
	25 $_{Item\_No}$				25 $_{Item\_No}$		
Observations	1839				1839		
Marginal R^2 / Conditional R^2	0.318 / 0.445				0.000 / 0.453		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Research Question 6: *What is the relative contribution of baseline macro-linguistic input, output, and input-output abilities, and micro-linguistic abilities, on lexical retrieval abilities 3 months post-treatment in a group of 30 PWA who were randomized to receive a semantically based (SFA) treatment?*

The best-fit for the treated item, $\chi^2(5) = 23.03$, $p < 0.001$, untreated-related item, $\chi^2(5) = 21.21$, $p = 0.001$, and untreated-unrelated item ($\chi^2(5) = 31.70$, $p < 0.001$) models were significantly better fits for the data than the null models. The models for the treated items demonstrated 70% sensitivity and 74% specificity, the untreated-related items had 75% sensitivity and 70% specificity, and both sensitivity and specificity were 75% in the model for untreated-unrelated items. See Figure 10 for plots of predicted probabilities for the significant predictors and Tables 23-25 for detailed results.

In the model for naming of treated items at follow-up, the intercept was significant, $\beta = -0.35$, (*SE* = 0.17), *Wald*(1) = -2.07, B-H corrected $p = .094$, *OR* = 0.71, such that participants had a 41% mean predicted probability of correctly naming items, holding all other variables constant. This finding is also true for the model of untreated-related items: the intercept was significant, $\beta = -0.40$, (*SE* = 0.19), *Wald*(1) = -2.07, B-H corrected $p = .094$, *OR* = 0.67, and participants only had a 40% probability of correctly naming items. Across all models, the SAPA 3 was the only significant predictor of follow-up naming accuracy for treated items, $\beta = 0.72$, (*SE* = 0.24), *Wald*(1) = 3.05, B-H corrected $p = .012$, *OR* = 2.06, untreated-related items, $\beta = 0.77$, (*SE* = 0.27), *Wald*(1) = 2.85, B-H corrected $p = .017$, *OR* = 2.15, and untreated-unrelated items, $\beta = 0.80$, (*SE* = 0.23), *Wald*(1) = 3.57, B-H corrected $p < .001$, *OR* = 2.24. That is, for PWA whose baseline SAPA 3 scores were 1 *SD* above the mean, they had a 60% probability of

naming treated and untreated-related items, and a 70% probability of naming untreated-unrelated items at follow-up. No other predictors were significant (all *p*s > .05).

Figure 10

SFA Group Predictors of Naming Performance Across Item Types At Follow-Up

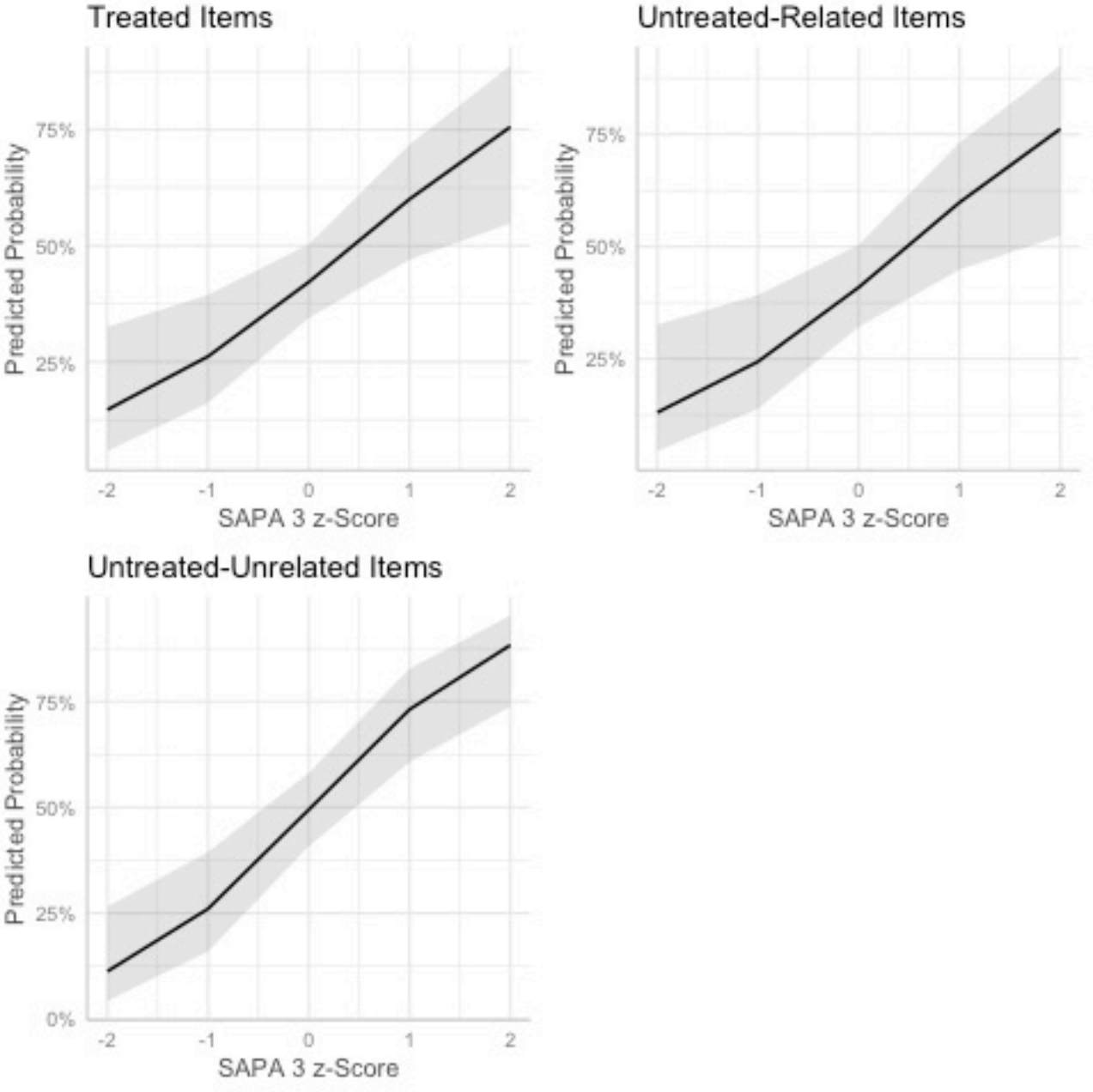


Table 23

Logistic Regression Results: SFA Group – Treated Items at Follow-Up

	SFA Treated Follow-Up Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.71	0.51 – 0.98	**0.094**	0.038	0.75	0.47 – 1.21	0.244
CAT T-score	0.90	0.57 – 1.44	0.745	0.669			
SAPA 2	1.30	0.86 – 1.95	0.350	0.217			
SAPA 3	2.06	1.29 – 3.28	**0.012**	0.002			
Semantic Impairment	1.27	0.82 – 1.94	0.393	0.28			
Phonological Impairment	1.22	0.87 – 1.71	0.364	0.239			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	$0.15_{\text{Item\_No}}$				$0.15_{\text{Item\_No}}$		
	0.68_{SubID}				1.59_{SubID}		
ICC	0.2				0.35		
N	28_{SubID}				28_{SubID}		
	$81_{\text{Item\_No}}$				$81_{\text{Item\_No}}$		
Observations	5627				5627		
Marginal R^2 / Conditional R^2	0.178 / 0.343				0.000 / 0.346		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 24

Logistic Regression Results: SFA Group – Untreated-Related Items at Follow-Up

	SFA Untreated-Related Follow-Up Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	0.67	0.46 – 0.98	**0.094**	0.039	0.71	0.41 – 1.21	0.210
CAT T-score	0.84	0.49 – 1.41	0.605	0.504			
SAPA 2	1.37	0.86 – 2.18	0.328	0.185			
SAPA 3	2.15	1.27 – 3.64	**0.017**	0.004			
Semantic Impairment	1.32	0.81 – 2.14	0.391	0.262			
Phonological Impairment	1.28	0.87 – 1.88	0.344	0.204			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	$0.17_{\text{Item\_No}}$				$0.17_{\text{Item\_No}}$		
	0.84_{SubID}				1.91_{SubID}		
ICC	0.23				0.39		
N	28_{SubID}				28_{SubID}		
	$41_{\text{Item\_No}}$				$41_{\text{Item\_No}}$		
Observations	2830				2830		
Marginal R^2 / Conditional R^2	0.185 / 0.376				0.000 / 0.388		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

Table 25

Logistic Regression Results: SFA Group – Untreated-Unrelated Items at Follow-Up

	SFA Untreated-Unrelated Follow-Up Model				**Null Model**		
Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*Uncorrected p-Value*	*Odds Ratios*	*CI (95%)*	*p-Value*
Intercept	1.02	0.73 – 1.45	0.906	0.889	1.08	0.61 – 1.92	0.782
CAT T-score	1.27	0.80 – 2.04	0.432	0.312			
SAPA 2	0.92	0.62 – 1.35	0.744	0.661			
SAPA 3	2.24	1.44 – 3.48	**<0.001**	<0.001			
Semantic Impairment	1.45	0.95 – 2.24	0.179	0.088			
Phonological Impairment	1.21	0.88 – 1.65	0.364	0.236			
Random Effects							
σ^2	3.29				3.29		
τ_{00}	0.53_{SubID}				2.04_{SubID}		
	$0.15_{\text{Item\_No}}$				$0.15_{\text{Item\_No}}$		
ICC	0.17				0.4		
N	27_{SubID}				27_{SubID}		
	$25_{\text{Item\_No}}$				$25_{\text{Item\_No}}$		
Observations	1800				1800		
Marginal R^2 / Conditional R^2	0.250 / 0.378				0.000 / 0.400		

Note. CI = confidence interval; B-H = Benjamini-Hochberg; CAT Comp. = Comprehensive Aphasia Test Comprehension Subscale; SAPA 2 = Standardized Assessment of Phonology in Aphasia (SAPA) – Auditory Phonological Processing Subtest; SAPA 3 = SAPA Repetition, Parsing, and Blending Subtest.

***Post Hoc* Analyses**

The nature of the data and the main analysis results brought up several questions that were not the focus of this study, but that certainly deserve further inspection. For all of the analyses discussed below, the methods above still applied: All models were assessed for multicollinearity, outliers, and influential data. Except where indicated, the best-fit models were significantly better fits for the data than the null model (all χ^2 *p*-values > .05). Furthermore, all *p*-values were corrected using the Benjamini-Hochberg procedure with a 10% false discovery rate. Again, for brevity, only significant results are presented in this section. For *Post Hoc* 1, please see Appendix H, and for *Post Hoc* 2, see Appendix I for full model results.

***Post Hoc* 1: Severity.** First, as was discussed in the description of the outcome stimuli, participants received either high frequency or low frequency word sets, depending on their CAT comprehension and CAT repetition T-scores. This difference in the stimuli, coupled with established differences between more and less severe participants in previous investigations of PMT (Minkina et al., 2016; Minkina et al., 2019; Zimmerman, Silkes, Kendall, & Minkina, 2019), compelled the first author to determine whether results vary between the severity groups. Thus, the first *post hoc* research question was: *Do more and less severe groups demonstrate different predictors of correct naming performance of treated, untreated-related, and untreated-unrelated items immediately post-treatment and at the three month follow-up, across and between groups?*

Table 26 is a summary the number of mild and severe participants included in the analyses before and after influential data were removed. Table 27 is an overview of the results, which are also visualized in Figures 11-20, and full model results tables are in Appendix H. In this analysis, participants' severity was determined by the word sets they received in the main

study, as discussed in *Methods*: Those with standard scores < 50 on the CAT auditory comprehension and repetition subtests were categorized as more impaired (in this analysis, "severe"), while those whose scores were > 50 were categorized as less impaired ("mild"). Those whose scores were at 50 were determined on a case-by-case basis.

Overall, SAPA 3 remained a robust predictor of performance across groups and items. Interestingly, the mild groups demonstrated significant effects for composite measures, such as the CAT comprehension T-score, the SAPA 2, and the SAPA 3, while the more severe groups demonstrated stronger relationships between micro-linguistic impairment and naming performance across item types immediately post-treatment. Specifically, in the more severe groups, increased semantic and/phonological impairment was associated with greater ability to name items following treatment. Possible reasons for this will be explored in the discussion. It should also be noted that the high degree of variability in the SFA group residuals, coupled with the small sample size, impeded many attempts to model performance on the outcomes – especially among the severe participants – such that the null models were the best fits to the data. Thus, for the mild SFA group, only 3/6 analyses could be completed, while 0/6 models were good fits for the severe SFA data.

Table 26
Sample Sizes for Post Hoc *1: Severity*

	Sample before data removed		*Sample after data removed*	
Group	Mild Group	Severe Group	Mild Group	Severe Group
All	33	25	31	23
PMT	13	15	12	14
SFA	20	10	18	NA

Table 27

Overview of Results from Severity Post Hoc *Analyses*

	Treated Items				Untreated-Related Items				Untreated-Unrelated Items			
	Mild		Severe		Mild		Severe		Mild		Severe	
	Imm. Post-Tx	Follow -Up	Imm. Post-Tx	Follow -Up	Imm. Post-Tx	Follow -Up	Imm. Post-Tx	Follow -Up	Imm. Post-Tx	Follow -Up	Imm. Post-Tx	Follow -Up
All	SAPA 3	SAPA 3	SAPA 3 Sem. Phon.	n.s.	CAT SAPA 3 Sem. Phon.	SAPA 3	SAPA 3 Phon.	SAPA 2 Sem.	SAPA 3	SAPA 3	Sem. Phon.	Sem. Phon.
PMT	SAPA 2		SAPA 3 Sem. Phon.	SAPA 2 Sem. Phon.			Sem. Phon.	SAPA 2 Sem.	CAT SAPA 3 Phon.	SAPA 3 Phon.	SAPA 3 Sem. Phon.	Sem. Phon.
SFA	n.s.				SAPA 2				SAPA 3	SAPA 3		

Note. Imm. Post-Tx = Immediate post-treatment; PMT = Phonomotor treatment group; SFA = Semantic Feature Analysis group; CAT = Comprehensive Aphasia Test – Auditory Comprehension T-score; SAPA = Standardized Assessment of Phonology in Aphasia; SAPA 2 = SAPA Subtest 2 – Auditory Processing; SAPA 3 = SAPA Subtest 3 – Repetition, Parsing, and Blending; Sem. = semantic impairment score; Phon. = phonological impairment score.

Whole Group Analyses

Mild Aphasia. In the mild whole group analyses, the SAPA 3 was significantly associated with accurate immediate post-treatment naming performance on treated items, β = 0.45, (*SE* = 0.17), *Wald*(1) = 2.57, B-H corrected *p* = .059, *OR* = 1.56, untreated-related items, β = 0.29, (*SE* = 0.15), *Wald*(1) = 1.99, B-H corrected *p* = .092, *OR* = 1.34, and untreated-unrelated items (β = 0.78, (*SE* = 0.22), *Wald*(1) = 3.54, B-H corrected *p* < .001, *OR* = 2.10). In the model for untreated-related item naming post-treatment only, the CAT comprehension T-score, β = 0.51, (*SE* = 0.14), *Wald*(1) = 3.99, B-H corrected *p* < .001, *OR* = 1.64, semantic impairment score, β = -0.24, (*SE* = 0.13), *Wald*(1) = -1.91, B-H corrected *p* = .091, *OR* = 0.78, and the phonological impairment score (β = -0.31, (*SE* = 0.14), *Wald*(1) = -2.20, B-H corrected *p* = .050, *OR* = 0.71) significantly predicted correct naming. Note that the relationship between impairment level and naming accuracy is actually negative in this group, such that having semantic and phonological impairment scores 1 *SD below* the mean corresponds to a 77% mean predicted probability for both. No other predictors were significant in these models (all *p*s > .05). See Figure 11 for plots of the predicted probabilities of all of the significant independent variables immediately post-treatment.

At the follow-up period, SAPA 3 again emerged as a significant predictor of naming accuracy for the treated items, β = 0.43, (*SE* = 0.18), *Wald*(1) = 2.44, B-H corrected *p* = .071, *OR* = 1.53, untreated-related items, β = 0.49, (*SE* = 0.19), *Wald*(1) = 2.63, B-H corrected *p* = .050, *OR* = 1.64, and untreated-unrelated items, β = 0.74, (*SE* = 0.21), *Wald*(1) = 3.43, B-H corrected *p* = .010, *OR* = 2.09. No other predictors were significant at the follow-up (all *p*s > .05). See Figure 12 for plots of the follow-up probabilities in the mild group.

Figure 11

Mild Whole Group Predictors of Naming Performance Across Item Types Immediately Post-Treatment

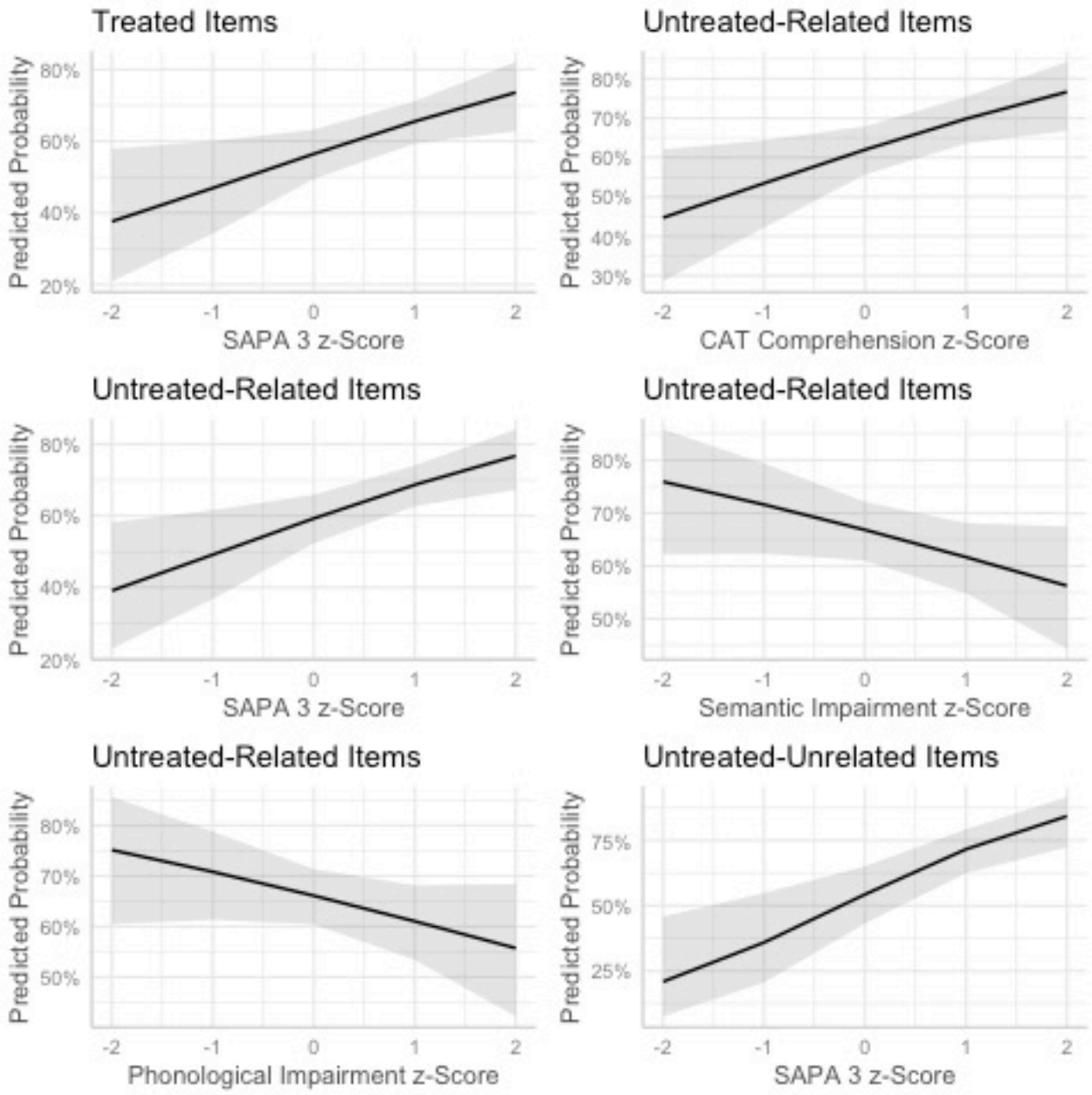
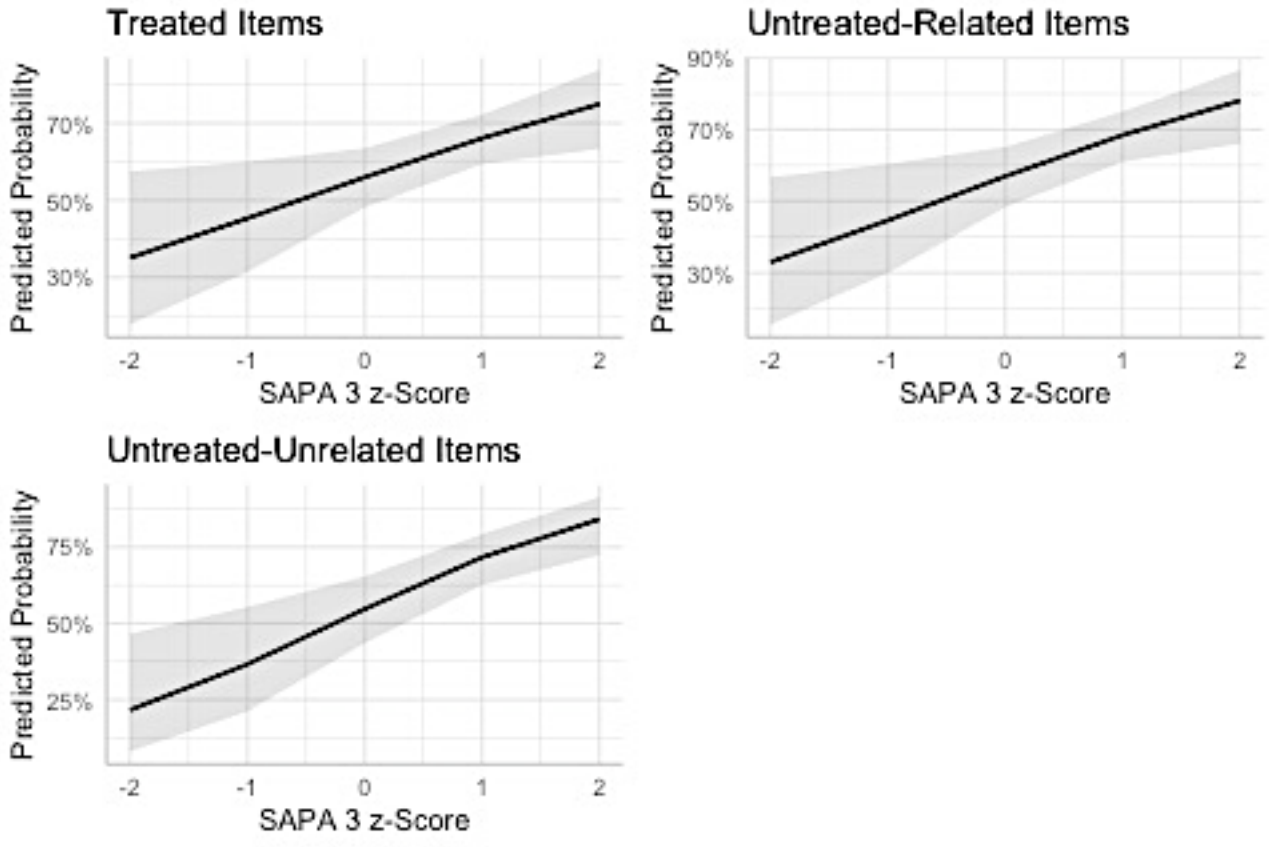


Figure 12

Mild Whole Group Predictors of Naming Performance Across Item Types at Follow-Up

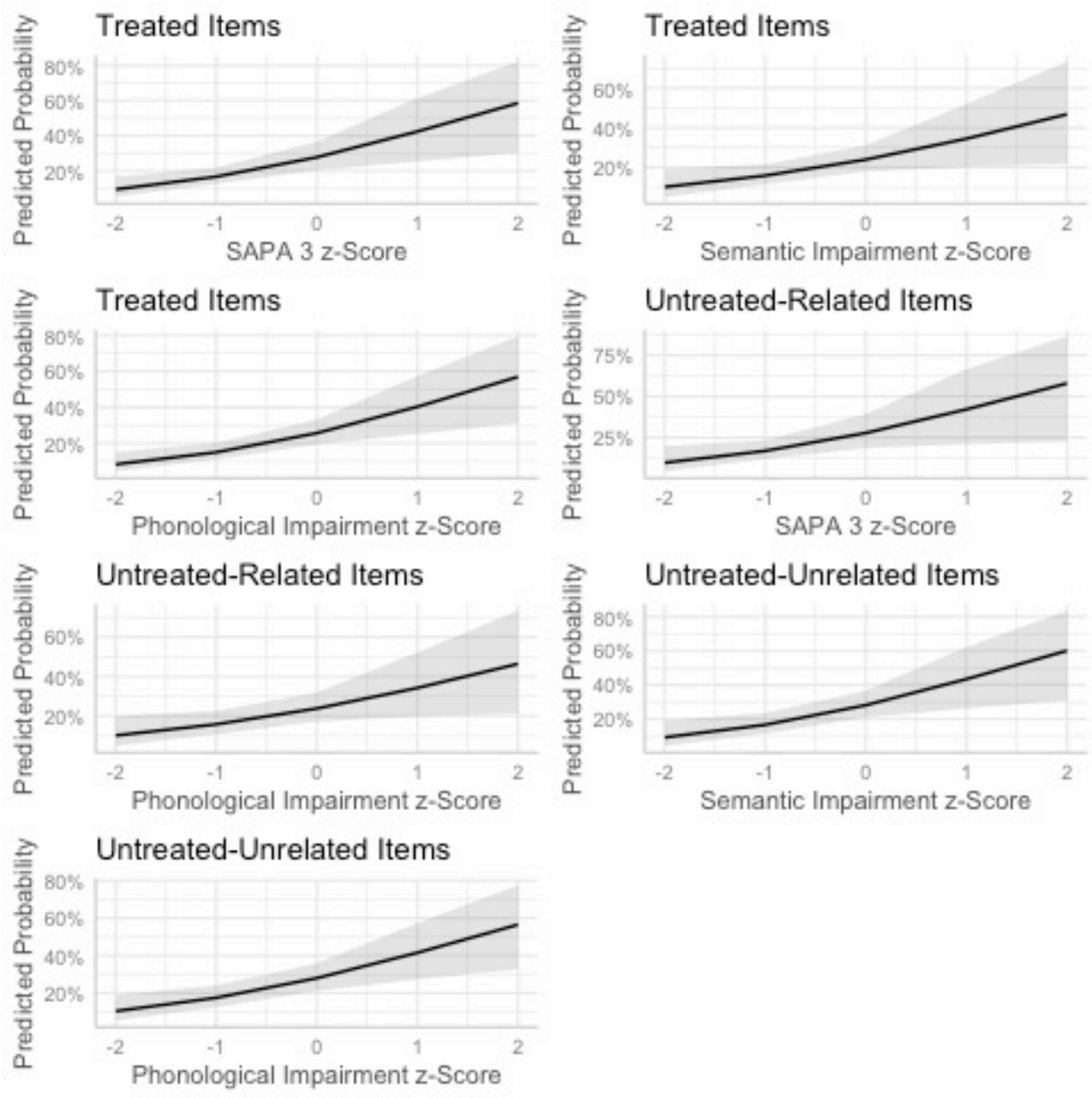


Severe Aphasia: In the severe whole group, immediate post-treatment naming performance on treated items was significantly predicted by baseline SAPA 3 scores, β = 0.63, (*SE* = 0.22), *Wald*(1) = 2.97, B-H corrected *p* = .010, *OR* = 1.95, semantic impairment, β = 0.51, (*SE* = 0.23), *Wald*(1) = 2.26, B-H corrected *p* = .071, *OR* = 1.65, and phonological impairment, β = 0.67, (*SE* = 0.21), *Wald*(1) = 3.18, B-H corrected *p* = .010, *OR* = 1.94. For untreated-related items, both the SAPA 3 (β = 0.67, (*SE* = 0.20), *Wald*(1) = 3.27, B-H corrected *p* = .071, *OR* = 1.95) and phonological impairment (β = 0.67, (*SE* = 0.20), *Wald*(1) = 3.27, B-H corrected *p* = .034, *OR* = 1.82) predicted post-treatment naming ability. Finally, semantic impairment (β = 0.67, (*SE* = 0.20), *Wald*(1) = 3.27, B-H corrected *p* = .039, *OR* = 2.02) and phonological

impairment (β = 0.67, (*SE* = 0.20), *Wald*(1) = 3.27, B-H corrected *p* = .010, *OR* = 1.89) were significant predictors of correct untreated-unrelated item naming. No other predictors were significant in this group (all *p*s > .05). See Figure 13 for plots of predicted probabilities.

Figure 13

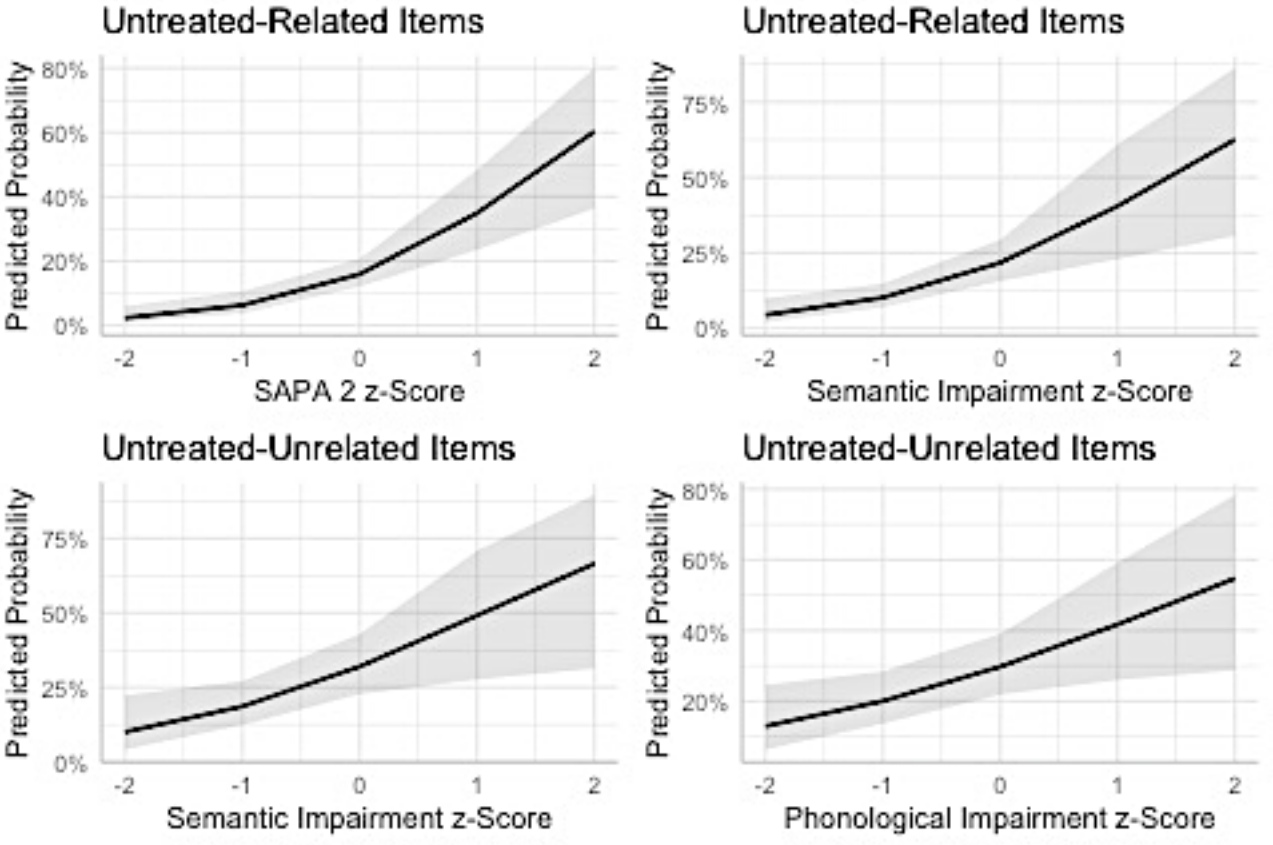
Severe Whole Group Predictors of Naming Performance Across Item Types Immediately Post-Treatment



At the follow-up time point, no variables predicted treated item naming accuracy in the whole severe group (all *p*s > .05). However, for untreated-related items, both the SAPA 2 (β = 1.04, (*SE* = 0.24), *Wald*(1) = 4.39, B-H corrected *p* < .001, *OR* = 2.83) and semantic impairment ability (β = 0.89, (*SE* = 0.27), *Wald*(1) = 3.34, B-H corrected *p* = .010, *OR* = 2.45) significantly predicted naming accuracy, and for the untreated-unrelated items, semantic impairment (β = 0.72, (*SE* = 0.29), *Wald*(1) = 2.47, B-H corrected *p* = .070, *OR* = 2.06) and phonological impairment (β = 0.52, (*SE* = 0.22), *Wald*(1) = 2.39, B-H corrected *p* = .071, *OR* = 1.69) predicted naming accuracy at follow-up. No other predictors in these two models were significant (all *p*s > .05). See Figure 14 for plots of predicted probabilities.

Figure 14

Severe Whole Group Predictors of Naming Performance Across Item Types at Follow-Up



PMT Group Analyses.

Mild Aphasia. In the mild PMT analyses of outcomes immediately post-treatment, the model for untreated-related items did not converge; thus, those results were not available for discussion. In the model of post-treatment performance on treated items, the SAPA 2 had a significant *negative* relationship with naming performance, β = -0.34, (*SE* = 0.10), *Wald*(1) = -3.29, B-H corrected $p = .030$, *OR* = 0.73, while the SAPA 3 had a significant *positive* relationship with the outcome, β = 1.13, (*SE* = 0.18), *Wald*(1) = 6.17, B-H corrected $p < .001$, *OR* = 2.99. In the untreated-unrelated items model, the CAT comprehension T-score (β = 0.63, (*SE* = 0.25), *Wald*(1) = 2.53, B-H corrected $p = .043$, *OR* = 1.94) and the SAPA 3 (β = 0.81, (*SE* = 0.26), *Wald*(1) = 3.13, B-H corrected $p = .010$, *OR* = 2.34) both had significant positive relationships with the outcome; phonological impairment (β = -0.80, (*SE* = 0.22), *Wald*(1) = -3.58, B-H corrected $p < .001$, *OR* = 0.43) had a significant negative relationship with post-treatment naming of untreated-unrelated items. No other variables had significant relationships in these models (all *p*s > .05). See Figure 15 for plots of the results.

At the follow-up timepoint in the mild PMT group, only the model for untreated-unrelated items converged. This model revealed a significant positive relationship between baseline SAPA 3 scores (β = 1.04, (*SE* = 0.27), *Wald*(1) = 3.84, B-H corrected $p < .001$, *OR* = 2.82) and the outcome, and a significant negative relationship between baseline phonological impairment (β = -0.99, (*SE* = 0.23), *Wald*(1) = -4.23, B-H corrected $p < .001$, *OR* = 0.37) and the outcome. No other predictors were significant (all *p*s > .05), and the plotted predicted probabilities for the significant variables can be seen in Figure 16.

Figure 15

Mild PMT Group Predictors of Naming Performance Across Item Types Immediately Post-Treatment

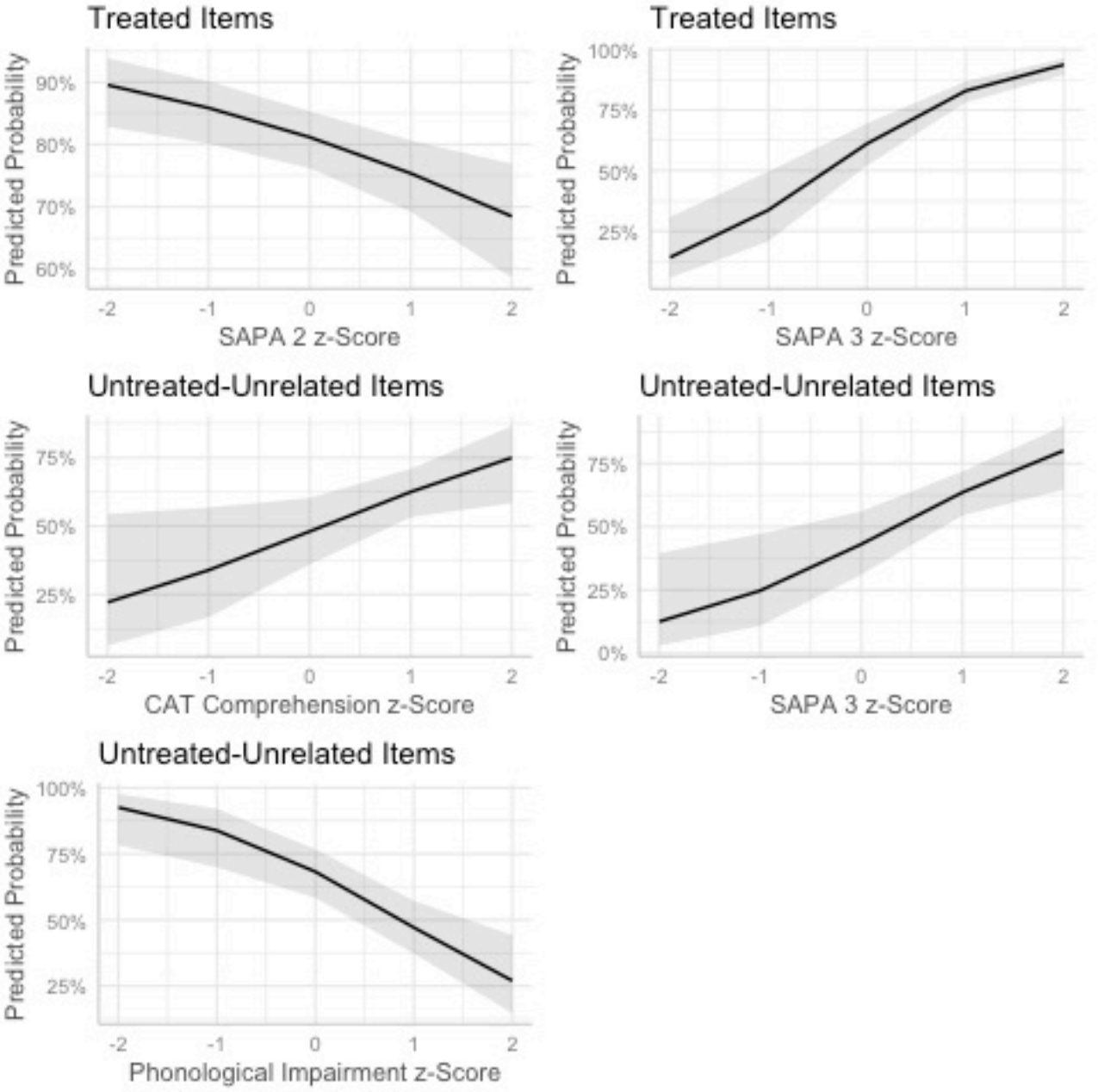
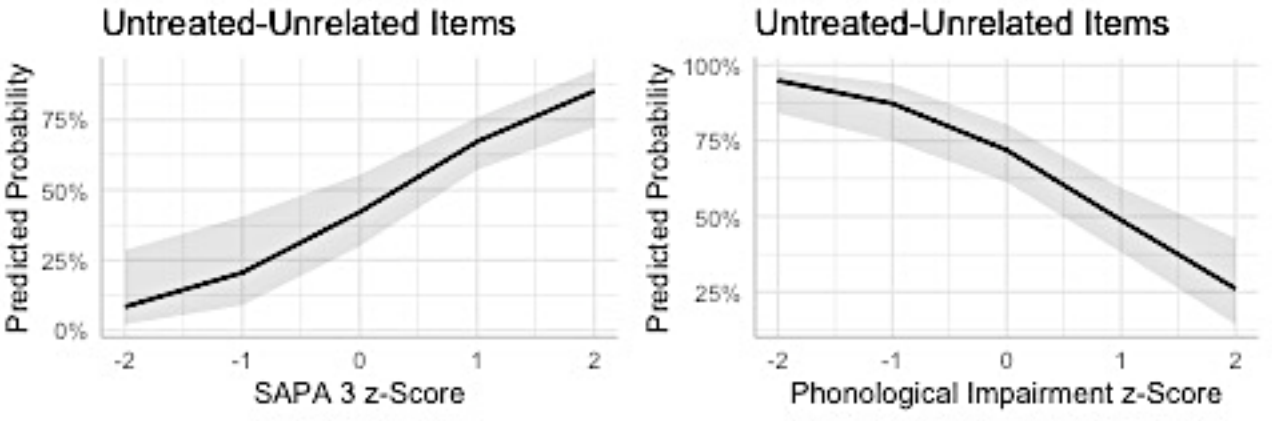


Figure 16

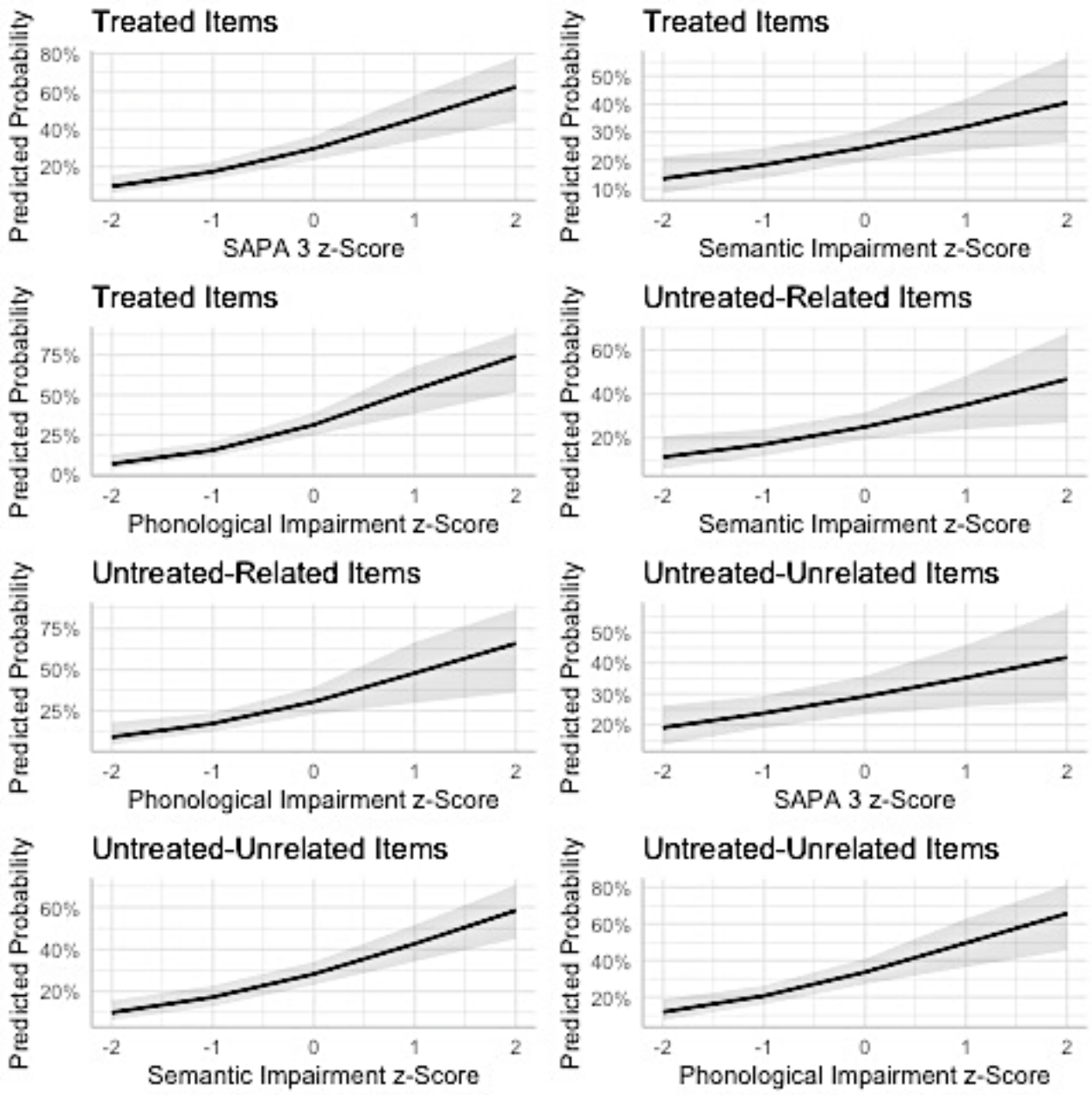
Mild PMT Group Predictors of Naming Performance Across Item Types at Follow-Up



Severe Aphasia. In the severe PMT group immediately post-treatment, naming performance on treated items was predicted by baseline SAPA 3, β = 0.69, (*SE* = 0.15), *Wald*(1) = 4.70, B-H corrected $p < .001$, *OR* = 2.05, semantic impairment, β = 0.37, (*SE* = 0.14), *Wald*(1) = 2.73, B-H corrected $p = .066$, *OR* = 1.43, and phonological impairment scores, β = 0.92, (*SE* = 0.20), *Wald*(1) = 4.58, B-H corrected $p < .001$, *OR* = 2.59. Semantic impairment (β = 0.48, (*SE* = 0.19), *Wald*(1) = 2.60, B-H corrected $p = .050$, *OR* = 1.62) and phonological impairment (β = 0.74, (*SE* = 0.24), *Wald*(1) = 3.05, B-H corrected $p = .017$, *OR* = 2.10) significantly predicted performance on untreated-related items immediately post-treatment. Finally, baseline SAPA 3, β = 0.28, (*SE* = 0.11), *Wald*(1) = 2.49, B-H corrected $p = .092$, *OR* = 1.30, semantic impairment, β = 0.64, (*SE* = 0.12), *Wald*(1) = 5.75, B-H corrected $p < .001$, *OR* = 1.92, and phonological impairment (β = 0.66, (*SE* = 0.16), *Wald*(1) = 4.15, B-H corrected $p < .001$, *OR* = 1.93) all significantly predicted accurate naming of untreated-unrelated items. No other items had significant relationships with the outcomes (all *p*s > .05). See Figure 17 for plots of the predicted probabilities of the significant predictors.

Figure 17

Severe PMT Group Predictors of Naming Performance Across Item Types Immediately Post-Treatment



At follow-up, participants' SAPA 2, β = 0.51, (*SE* = 0.18), *Wald*(1) = 2.79, B-H corrected *p* = .034, *OR* = 1.66, semantic impairment, = 0.55, (*SE* = 0.15), *Wald*(1) = 3.55, B-H corrected *p* < .001, *OR* = 1.73, and phonological impairment (= 0.51, (*SE* = 0.20), *Wald*(1) = 2.61, B-H corrected *p* = .050, *OR* = 1.67) scores significantly predicted treated item naming

accuracy. Performance on the untreated-related items was predicted by SAPA 2 scores (β = 0.98, (*SE* = 0.27), *Wald*(1) = 3.59, B-H corrected $p < .001$, *OR* = 2.67) and semantic impairment level, β = 0.86, (*SE* = 0.20), *Wald*(1) = 4.23, B-H corrected $p < .001$, *OR* = 2.36. For untreated-unrelated items, both semantic (β = 0.89, (*SE* = 0.18), *Wald*(1) = 4.85, B-H corrected $p < .001$, *OR* = 2.44) and phonological (β = 1.04, (*SE* = 0.23), *Wald*(1) = 4.54, B-H corrected $p < .001$, *OR* = 2.83) impairment predicted naming accuracy at follow-up. No other significant predictors emerged in this set of models (all *p*s > .05). See Figure 18 for plotted results.

SFA Group Analyses

Mild Aphasia. As stated earlier, only three of the SFA group analyses were reliable enough to interpret: immediate post-treatment naming accuracy of untreated-related and untreated-unrelated items and follow-up naming accuracy of untreated-unrelated items. All of these analyses come from the mild SFA group, and none of severe group models were significantly different from the null model. Furthermore, fewer predictors appeared to have significant relationships with the outcomes across time points in the mild group. Namely, SAPA 2 scores were associated with post-treatment naming of untreated-related items, β = 0.52, (*SE* = 0.19), *Wald*(1) = 2.71, B-H corrected $p = .071$, *OR* = 1.50, and the SAPA 3 related to correct naming of untreated-unrelated items immediately post-treatment (β = 0.82, (*SE* = 0.25), *Wald*(1) = 3.23, B-H corrected $p = .017$, *OR* = 2.06) and at follow-up, β = 0.68, (*SE* = 0.26), *Wald*(1) = 2.63, B-H corrected $p = .048$, *OR* = 1.97. Figure 19 presents the SFA mild group results immediately post-treatment, and Figure 20 presents the follow-up results.

Figure 18

Severe PMT Group Predictors of Naming Performance at Follow-Up Across Item Types

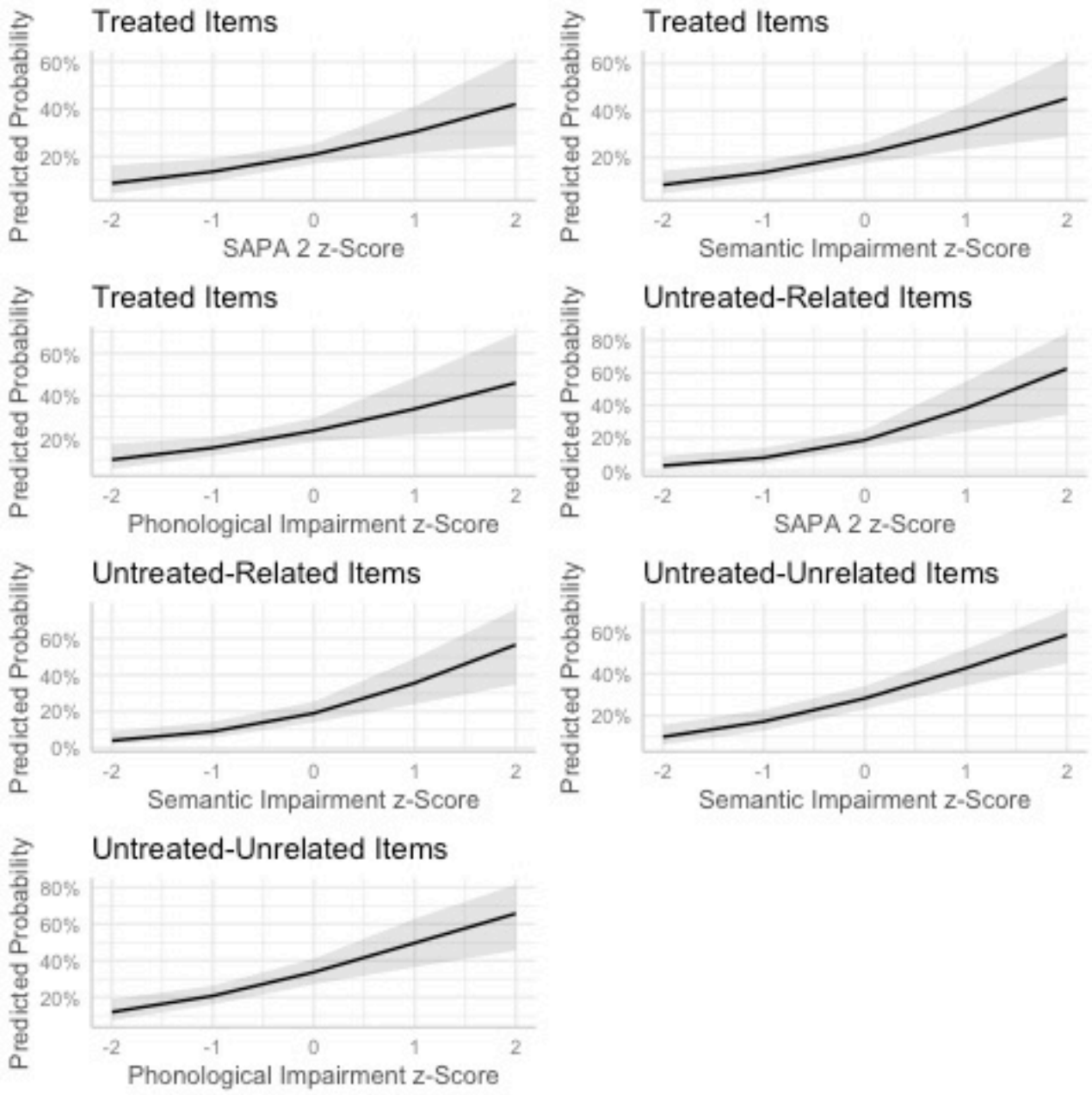


Figure 19

Mild SFA Group Predictors of Naming Performance Across Item Types Immediately Post-Treatment

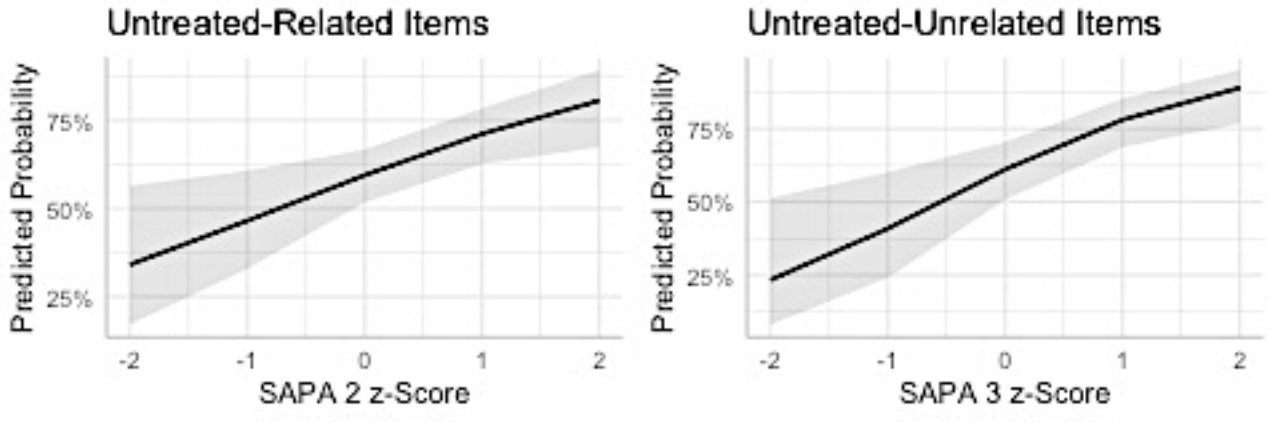
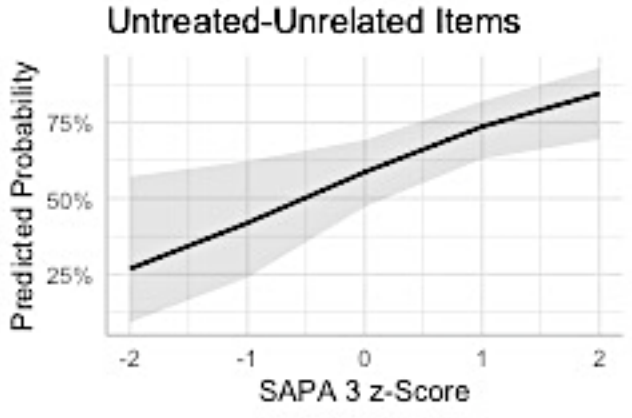


Figure 20

Mild SFA Group Predictors of Naming Performance at Follow-Up



***Post Hoc* 2: SAPA 3.**

As should be clear by now, the SAPA 3 was a particularly robust predictor of outcomes immediately post-treatment and at follow-up, and across groups. Given the results discussed above, it was worth exploring the relationships between the tasks within the SAPA 3 and the outcomes. As a brief reminder, the SAPA 3 score is the sum of scores from the following tasks: Real Word Repetition, Nonword Repetition, Real Word Blending, Nonword Blending, Real Word Parsing, and Nonword Parsing. Thus, this *post hoc* analysis was designed to answer the following question: *Which tasks within the SAPA 3, if any, predict accurate naming of treated,*

untreated-related, and untreated-unrelated items immediately post-treatment and at follow up, across and within groups?

To answer this question, the standardized scores from the SAPA 3 tasks were entered into logistic mixed-effect regression models, along with Time (immediate post-treatment and follow-up), to determine their relationships with performance on treated, untreated-related, and untreated-unrelated items. The general model formula was:

$$\text{Outcome}_{\hat{p}} = 1 + \text{Time} + \text{RWRep} + \text{NWRep} + \text{RWBlend} + \text{NWBlend} + \text{RWParse} + \text{NWParse} + (1|\text{SubID}) + (1|\text{Item\_No}) \quad (7)$$

where
RWRep = Real Word Repetition
NWRep = Nonword Repetition
RWBlend = Real Word Blending
NWBlend = Nonword Blending
RWParse = Real Word Parsing
NWParse = Nonword Parsing

These predictors were assessed for collinearity and, although the correlations between each subscale is much stronger than the relationships between the predictors in the main analyses, all relationships were still well below the benchmark of $r = \pm .90$ (Tabachnick & Fidell, 2013). See Figure 20 for the correlation coefficients and indications of significant findings (indicated by the blue circle). It should also be noted that the number of items in each subtest is rather small: In real word repetition, n = 9; nonword repetition, n = 5; real word blending, n = 5; nonword blending, n = 4; real word parsing, n = 5; nonword parsing, n = 5. Eighteen models in total were evaluated, 3 (Group) x 3 (Item Type) x 2 (Time), so *p*-values were corrected using the B-H correction. An overview of the result findings is in Table 28, the results are visualized in Figures 21-26, and full model results tables are available in Appendix I.

Table 28

Overview of Results from SAPA 3 Post Hoc *Analyses*

	Treated		Untreated-Related		Untreated-Unrelated	
	Imm. Post-Tx	Follow-Up	Imm. Post-Tx	Follow-Up	Imm. Post-Tx	Follow-Up
All	RW Parsing	n.s.	RW Parsing	RW Parsing	RW Parsing NW Parsing	n.s.
PMT	RW Blending RW Parsing	RW Blending RW Parsing	RW Blending RW Parsing	RW Blending RW Parsing	RW Parsing	RW Parsing
SFA	RW Repetition NW Repetition	NW Repetition	RW Repetition	NW Repetition	NW Parsing	n.s.

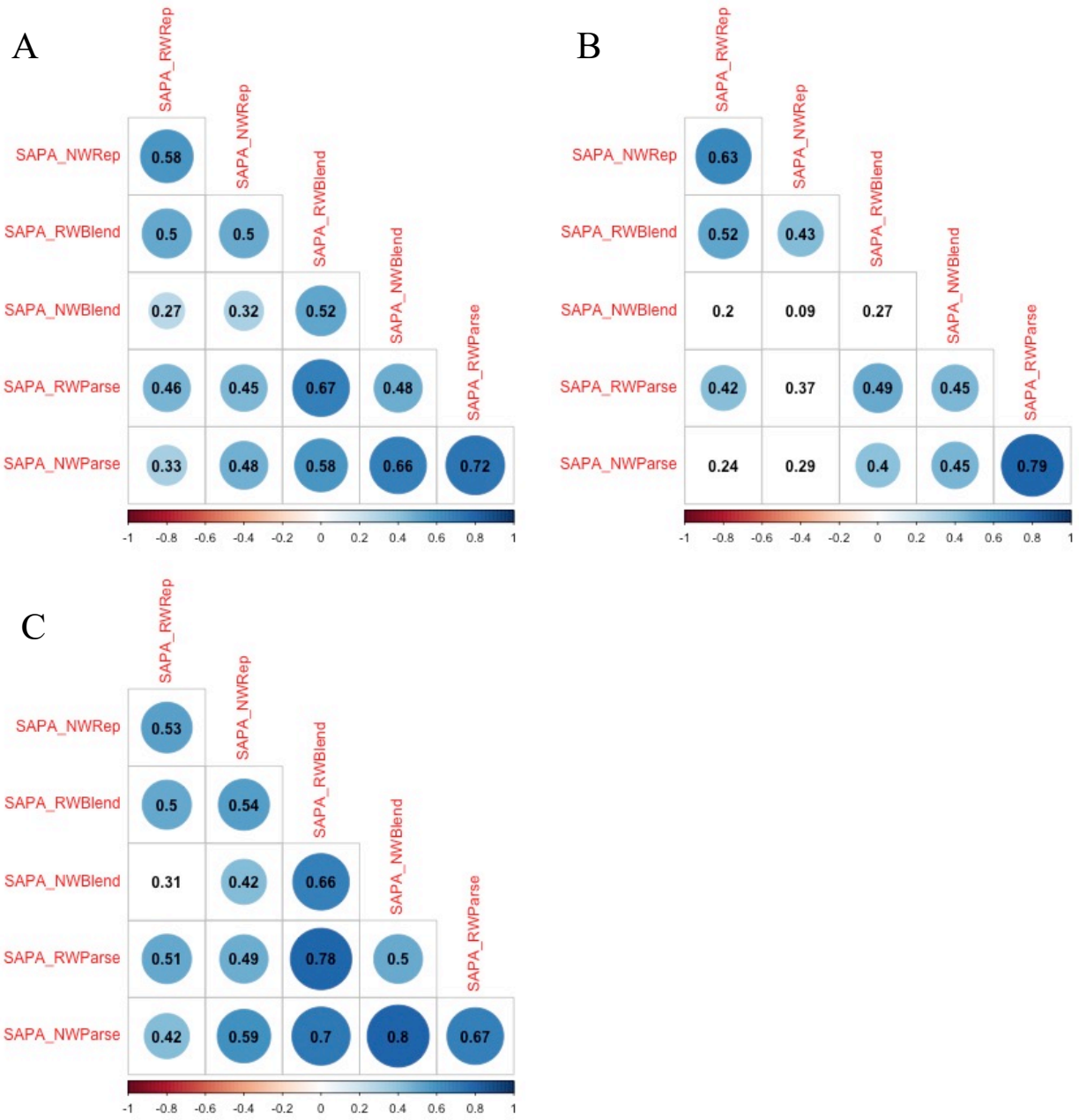
Note. Imm. Post-Tx = immediate post-treatment; PMT = Phonomotor treatment group; SFA = Semantic Feature Analysis group; RW = real word; NW = nonword; n.s. = no significant predictors.

Whole Group Analyses.

In the whole group, naming accuracy immediately post-treatment was significantly predicted by real word (RW) parsing for treated (β = 0.44, (*SE* = 0.20), *Wald*(1) = 2.28, B-H corrected *p* = .086, *OR* = 1.56), untreated-related (β = 0.47, (*SE* = 0.21), *Wald*(1) = 2.31, B-H corrected *p* = .080, *OR* = 1.61), and untreated-unrelated items, β = 0.47, (*SE* = 0.20), *Wald*(1) = 2.38, B-H corrected *p* = .071, *OR* = 1.60. Additionally, for untreated-unrelated items, nonword (NW) parsing ability was also predictive, β = 0.46, (*SE* = 0.21), *Wald*(1) = 2.21, B-H corrected *p* = .092, *OR* = 1.58. At the follow-up period, the only significant relationship found was between untreated-related items and RW parsing, β = 0.51, (*SE* = 0.21), *Wald*(1) = 2.38, B-H corrected *p* = .071, *OR* = 1.67. See Figures 21-22 for plots of the predicted probabilities for the significant predictors.

Figure 21

SAPA 3 Subtest Correlations for Each Group Analysis



Note. Shaded correlations represent significant relationship at the α = .05 level. A = Correlations between predictors for the whole group analyses. B = Correlations between predictors for the PMT group analyses. C = Correlations between predictors for the SFA group analyses. SAPA_RWRep = real word repetition; SAPA_NWRep = nonword repetition; SAPA_RWBlend = real word blending; SAPA_NWBlend = nonword blending; SAPA_RWParse = real word parsing; SAPA_NWParse = nonword parsing.

PMT Group Analyses.

In the PMT group, RW blending significantly predicted naming accuracy for treated (β = 0.62, (*SE* = 0.21), *Wald*(1) = 2.90, B-H corrected *p* = .030, *OR* = 1.85) and untreated-related (β = 0.51, (*SE* = 0.22), *Wald*(1) = 2.33, B-H corrected *p* = .079, *OR* = 1.67) items immediately post-treatment naming. Likewise, RW parsing predicted naming of treated, β = 0.69, (*SE* = 0.30), *Wald*(1) = 2.30, B-H corrected *p* = .080, *OR* = 1.99, untreated-related, β = 0.74, (*SE* = 0.31), *Wald*(1) = 2.39, B-H corrected *p* = .071, *OR* = 2.09, and untreated-unrelated (β = 0.69, (*SE* = 0.28), *Wald*(1) = 2.42, B-H corrected *p* = .071, *OR* = 1.99) items. These results were echoed at the follow-up timepoint, in that RW blending predicted naming of treated (β = 0.61, (*SE* = 0.19), *Wald*(1) = 3.19, B-H corrected *p* = .010, *OR* = 1.84) and untreated-related (β = 0.67, (*SE* = 0.23), *Wald*(1) = 2.95, B-H corrected *p* = .024, *OR* = 2.40) items, while RW parsing predicted accuracy of treated, β = 0.64, (*SE* = 0.27), *Wald*(1) = 2.40, B-H corrected *p* = .071, *OR* = 1.90, untreated-related, β = 0.88, (*SE* = 0.31), *Wald*(1) = 2.81, B-H corrected *p* = .034, *OR* = 2.40, and untreated-unrelated (β = 0.98, (*SE* = 0.33), *Wald*(1) = 2.95, B-H corrected *p* = .024, *OR* = 2.67) items. See Figures 23-24 for plots of the results.

SFA Group Analyses.

Immediately post-treatment, the SFA group demonstrated significant relationships between treated items and RW (β = 0.46, (*SE* = 0.18), *Wald*(1) = 2.50, B-H corrected *p* = .064, *OR* = 1.59) and NW (β = 0.45, (*SE* = 0.20), *Wald*(1) = 2.24, B-H corrected *p* = .091, *OR* = 1.57) repetition, while only RW repetition significantly predicted untreated-related items, β = 0.46, (*SE* = 0.19), *Wald*(1) = 2.43, B-H corrected *p* = .071, *OR* = 1.59. For the untreated-unrelated items, NW parsing (β = 1.12, (*SE* = 0.30), *Wald*(1) = 3.72, B-H corrected *p* < .001, *OR* = 3.05) emerged as the significant predictor of naming accuracy immediately post-treatment. At the follow-up

point, NW repetition was the only significant predictor, and it only predicted naming performance on treated (β = 0.49, (*SE* = 0.21), *Wald*(1) = 2.35, B-H corrected *p* = .078, *OR* = 1.63) and untreated-related items, β = 0.50, (*SE* = 0.23), *Wald*(1) = 2.21, B-H corrected *p* = .092, *OR* = 1.65. There were no significant predictors of untreated-unrelated items at follow-up after the B-H correction (all B-H corrected *p*s > .05), although NW parsing was marginally significant before the correction (*p* = .05). See Figures 25-26 for plotted probabilities of the significant results.

Figure 22

Whole Group SAPA 3 Predictors of Naming Performance Across Item Types Immediately Post-Treatment

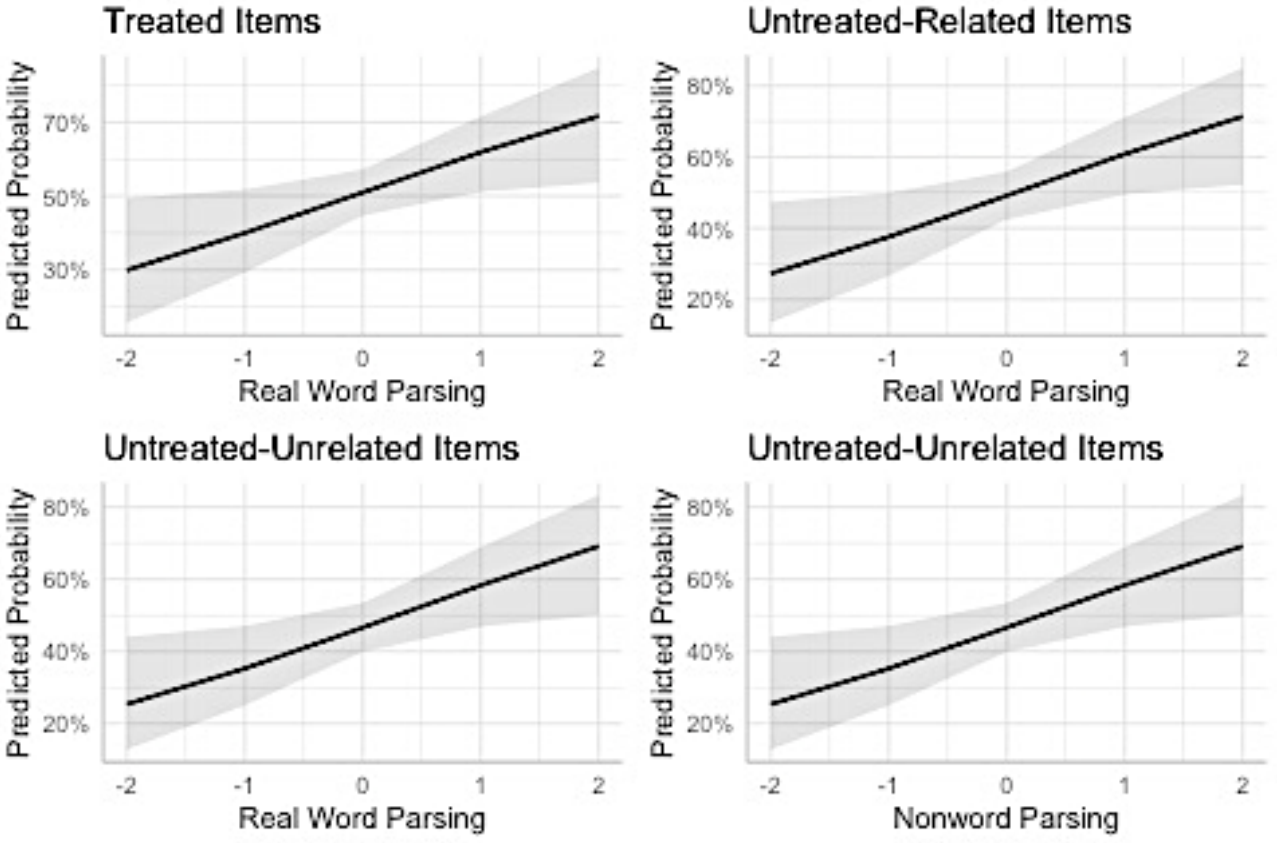


Figure 23

Whole Group SAPA 3 Predictors of Naming Performance Across Item Types at Follow-Up

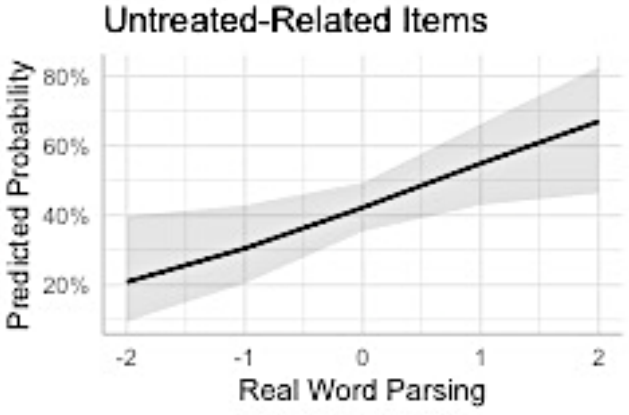


Figure 24

PMT Group SAPA 3 Predictors of Naming Performance Across Item Types Immediately Post-Treatment

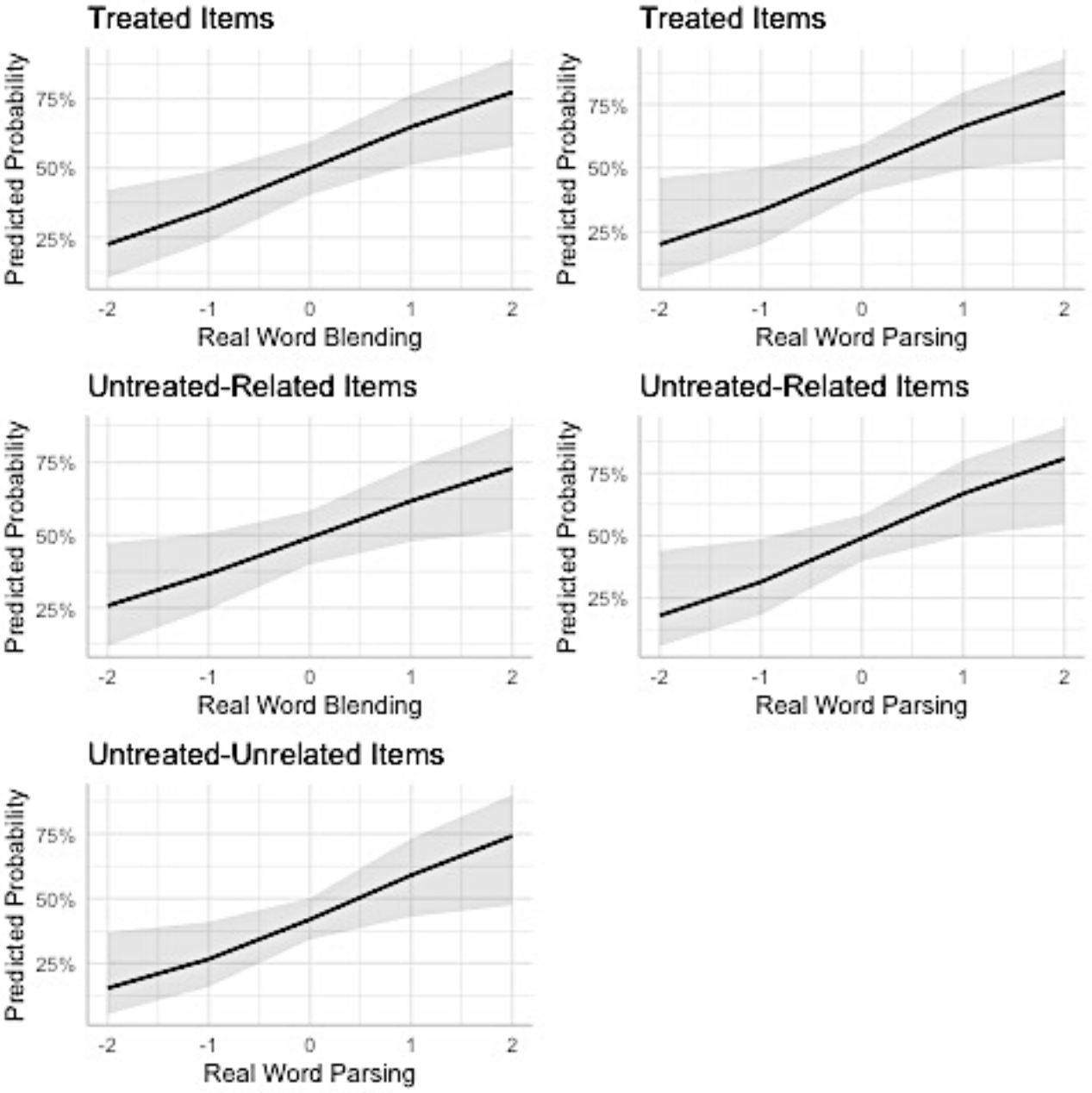


Figure 25

PMT Group SAPA 3 Predictors of Naming Performance Across Item Types at Follow-Up

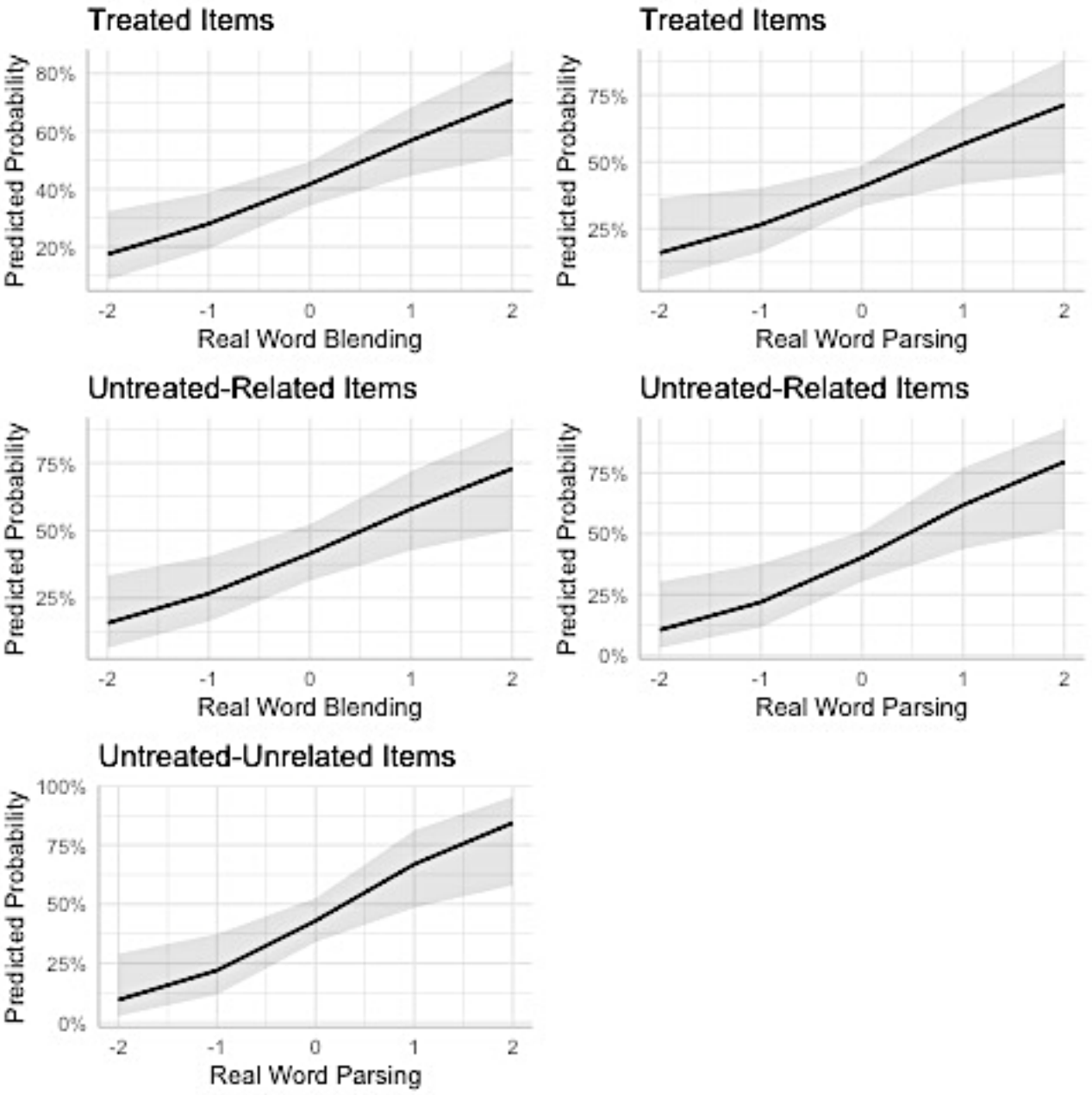


Figure 26

SFA Group SAPA 3 Predictors of Naming Performance Across Item Types Immediately Post-Treatment

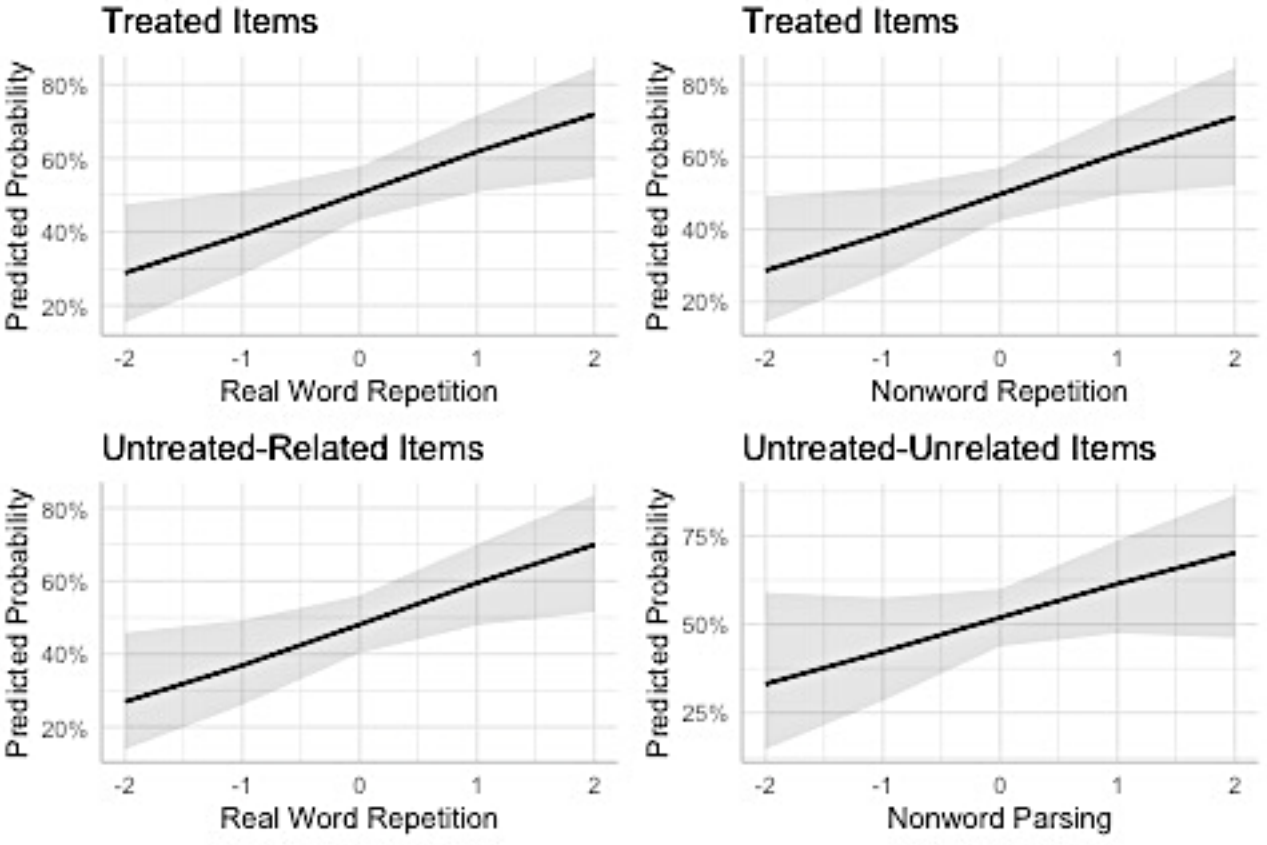
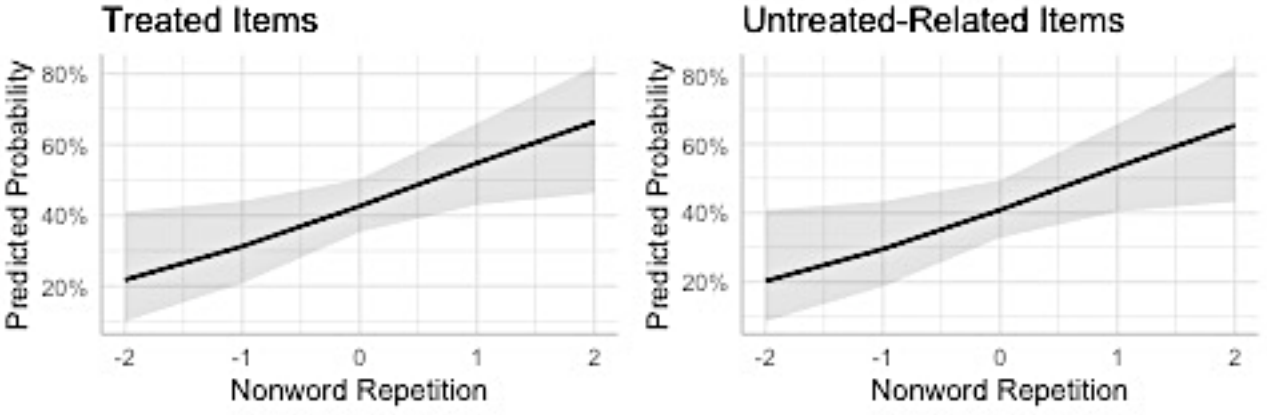


Figure 27

SFA Group SAPA 3 Predictors of Naming Performance at Follow-Up Across Item Types



Chapter 4: Discussion

Many PWA seek treatment to ameliorate their language impairments; however, individual results are highly variable. Understanding the basis of this variability is crucial to selecting and customizing treatments to maximize the benefits that individuals experience as a result of intervention. Although previous research has indicated that aphasia severity is an important factor in treatment response, relatively few studies have explored the relationships between specific aspects of aphasic language processing and treatment response. Thus, the aim of this retrospective study was to investigate the relationship between macro- and micro-linguistic abilities and treatment outcomes immediately post-treatment and at a 3-month follow-up in 58 PWA who were randomly assigned to either a semantically or phonologically based treatment program. To meet this aim, the predictive value of comprehension, input-phonological, input-output phonological, and output semantic and phonological impairment levels were modeled for naming accuracy for treated, untreated-related, and untreated-unrelated items immediately and three months post-treatment using mixed-effect logistic regression.

Overall, the results of this study suggest that, regardless of the type of treatment delivered, input-output phonology is an essential linguistic construct that predict the ability to acquire trained stimuli, generalize to items that were both related (near-generalization) and unrelated (far-generalization) to treated items, and maintain treatment gains. As will be discussed, input-output phonology may be more of a measure of verbal short-term and working memory and, if this is the case, then these results are compatible with prior literature that has identified strong links between verbal STM and treatment acquisition in PWA (Coran et al., 2020; Dignam et al., 2017; Kalinyak-Fliszar et al., 2011; Koenig-Bruhin & Studer-Eichenberger, 2007; N. Martin et al., 2020), as well as significant associations between verbal WM and generalization to untrained items following anomia treatments in PWA (Harnish et al., 2018;

Majerus, 2018). These relationships between STM, WM, and treatment outcomes are likely borne of an increased ability to maintain activation long enough to strengthen targeted and adjacent links, focus attention, synthesize new information with spared representations in long-term memory, and apply strategies during the naming task. Furthermore, greater micro-linguistic impairments predicted acquisition, near- and far-generalization, and maintenance thereof in the whole group, which is also consistent with previous research (Best et al. 2013; Kiran, 2016), and suggests that those with greater impairment are more responsive to treatment.

Another interesting pattern to note was that, in the severity *post hoc* analyses, the composite measures (i.e., CAT auditory comprehension T-score, SAPA 2, and SAPA 3) were typically only significant predictors in the milder groups, but not in the more severe groups; instead, the impairment scores were the most robust predictors in the more severe groups. One possible explanation for these findings could be that when language is less impaired, the contribution of higher level skills that underlie the ability to carry out comprehension or repetition tasks, such as WM, may exert more of an influence on outcomes than some of the more "foundational" skills such as semantic or phonological processing. Conversely, when linguistic deficits are more impaired, the fundamental skills become the the most important factor in determining the likelihood of correct naming. However, this hypothesis is speculative, and is also undermined by the fact that (1) increasing impairment related to better post-treatment naming in the severe group and (2) skills such as verbal STM are important for language learning in general (Cowan, 2008; Gupta & MacWhinney, 1997), so measures such as the SAPA 3 should have consistently demonstrated significant relationships as well. Ultimately, the factors underlying this specific pattern of results prompt some questions about the measures, and this issue will be taken up in *Limitations.*

For the PMT treatment, better STM/WM and more severe phonological and semantic impairments predict success in the acquisition of treated items, near-generalization, and maintenance of these items. Far-generalization in the PMT group is predicted by STM/WM and greater semantic impairment. The strong relationships between verbal STM/WM possibly reflect the fact that the PMT protocol requires intense maintenance and manipulation of phonemes and phoneme strings during phonological awareness tasks (e.g., parsing, blending). Furthermore, the multimodal nature of the tasks entails sensory-motor integration, and the demands of this integration entail working memory processes (Hickok & Poeppel, 2004). Thus, stronger baseline abilities to maintain, manipulate, and integrate sensory-motor information – as demonstrated by the SAPA 3 – supported participants' participation in and completion of the treatment tasks.

Greater phonological impairment predicts acquisition and near-generalization in PMT because damaged lexical-phonological links are being directly targeted during treatment, and the training may be sufficient to make the repairs robust to degredation over time, resulting in maintenance of gains. However, because the untreated-unrelated items do not share any of the same 3-phoneme segments with the treated items, the improvements induced by the PMT protocol may be restricted to sequences of three or more phonemes (Vitevitch & Storkel, 2012), which explains why this result did not extend to far-generalization. Increased semantic impairment, on the other hand, is such a robust predictor of acquisition and near- and far-generalization because a decreased ability to activate semantics during the PMT tasks likely results in decreased interference and/or diversion of resources away from phonological proesses (Dell & Gordon, 2003; Vitevitch & Luce, 1998; Vitevitch & Luce, 1999; Vitevitch & Sommers, 2003). Once the treatment effects arising from reduced interference have been established, they

will persist in the network at least until follow-up, especially in the event that participants attempt to use their improved skills in everyday life.

For the SFA treatment, verbal STM/WM predicts success in the acquisition of treated items, far-generalization, and maintenance of acquisition, near-, and far-generalization. During SFA treatment, participants had to maintain a representation of the target item while activating and verbalizing words that are associated with that target. Such a task relies on strong STM for activation maintenance and WM for attention to task and semantic analysis of the target (Craik & Lockhart, 1972). Greater phonological impairment at baseline also predicts post-treatment generalization to related and unrelated untreated items. Although this finding may be generally inconsistent with other findings in the literature that less impairment is associated with better outcomes (e.g., Quique, Evans, & Dickey, 2018), this discrepancy may be related to the specific measures of baseline impairment – that is, overall pre-treatment naming ability versus proportion of phonological errors. Nevertheless, this finding reflects the effects of broad, but shallow, training of a diverse array of phonological sequences that occur during the production-heavy SFA protocol. That is, since SFA participants are free to choose words that they associate with the target items, they practice producing a variety of words, which in turn boosts phonological processing performance. However, since participants may only produce these idiosyncratic features a handful of times during the treatment, this broad "training" does not produce lasting effects (Gupta & MacWhinney, 1997; Hinton & Plaut, 1987; Plaut, 1996), as indicated by the lack of significant relationships between phonological impairment and generalization and maintenance.

The structure of this discussion is as follows. First, the whole group results will be interpreted, with special attention given to the role of verbal STM and WM in language

processing. Then, the results for the individual treatment groups, PMT and SFA respectively, will be discussed. The clinical implications of this study will follow the discussion of the findings, and then the discussion will conclude with the study limitations and future directions.

Whole Group Findings

In the whole group analyses, baseline SAPA 3 performance was a clear and consistent predictor of naming ability across item types after anomia intervention. Specifically, higher SAPA 3 scores at baseline were associated with greater likelihood of naming treated, untreated-related, and untreated-unrelated items immediately following treatment and at follow-up. As discussed in *Methods*, the SAPA 3 is a measure of phonological processing abilities (e.g. phonological awareness and verbal STM) that includes tasks such as word and nonword repetition, blending, and parsing.

Studies of oral language and reading development in young children have associated the ability to blend and parse phoneme strings with phonological awareness (Mathes & Torgesen, 2000; McBride-Chang, 1995; Schatschneider, Francis, Foorman, Fletcher, & Mehta, 1999; Torgesen, 2000; Wagner, Torgesen, & Rashotte, 1994) – that is, the ability to identify and manipulate phonemes and phoneme sequences, plus the metalinguistic ability to think about these representations (McBride-Chang, 1995; Torgeson & Mathes, 2000). Based on analyses of phonological awareness in children, this ability is thought to rely on oral language skills, typically measured through vocabulary tests, coupled with verbal STM (measured by digit and word spans), WM (e.g., listening spans), and executive functioning skills, such as analysis or inhibition (e.g., block design, cancellation; McBride-Chang, 1995; Wagner et al., 1994; Zifcak, 1981).

The other skill included in the SAPA 3 subtest is the ability to repeat words and nonwords, which depends on both strong input-output phonological processing and verbal STM for single word and nonword repetition, as well as lexical-semantic processing for word repetition (N. Martin, Minkina, Kohen, & Kalinyak-Fliszar, 2018). Nonword repetition in particular has long been a robust predictor of vocabulary acquisition in children (Gathercole, 1995; Service, 1992) and adults with (e.g., Bose et al., 2019; Baddeley et al., 1988) and without brain injury (e.g., Storkel, 2001). Several authors have suggested that the relationship between nonword repetition and vocabulary acquisition is so consistent because all novel words are initially represented as nonwords; the ability to maintain and repeat such nonwords relates to the ability to create more stable, long-term links between the novel phoneme sequence and semantic representations (Cowan, 1996; Gathercole & Baddeley, 1989; Gupta & MacWhinney, 1997; Storkel, 2001).

Taken together, the three tasks included in the SAPA 3 represent the activation, maintenance, and manipulation of phonological information, and each task has a rich history of predicting language performance and/or acquisition in children and adults with typical and disordered language. Thus, it is not surprising that this subtest also emerged as a significant predictor in the current study. But what, exactly, are the mechanisms underlying these findings? There are three possible and related explanations for the results:

- The first possibility is that, because the SAPA 3 subtests rely so heavily on (1) input lexical-phonological, (2) output lexical-phonological, and (3) input-output phonological connections, good performance in the tasks represents good overall integrity of these links within the lexical-phonological level of the language network. Some extant literature suggests that spared lexical-phonological processing abilities facilitate

acquisition and near-generalization following treatment (Eoute, 2010; Jacquemot, Dupoux, Robotham, & Bachoud-Lévi, 2012).

- The second possibility is that, because the SAPA 3 subtests rely so heavily on the ability to activate and maintain phonological representations, good performance on the tasks represents sufficient activation maintenance to allow for adequate spread of activation and, consequently, increased Hebbian restoration of links between representations.
- The third possibility is that, because blending and parsing entail both WM and executive function abilities (e.g., problem-solving, inhibition, attention control), these tasks actually reflect the ability to apply executive function operations to linguistic processing (Cowan, 1996; Engle, 2002; Engle et al., 1999; N. Martin et al., 2012; Purdy, 2002; Unsworth & Engle, 2007). Such higher-order processing has been associated with increased linguistic acquisition and generalization in people with aphasia, because it enables (1) analysis of new representations and synthesis of these representations into existing linguistic schemata (Bose, Höbler, & Saddy, 2018; Brownsett et al., 2014; Craik & Lockhart, 1972; Miyake & Shah, 1999; Yeung & Law, 2010), and (2) problem-solving and communication repair (Fillingham, Sage, & Lambon Ralph, 2006; Hinckley, Patterson, & Carr, 2001; Nicholas, Sinotte, & Helm-Estabrooks, 2005; Purdy, 2002).

Strong links between input phonology, output phonology, and the lexical level would certainly facilitate the ability to participate in the active ingredients of both treatments. In PMT, all of the tasks entail explicit input and output of phonemes and phoneme strings. In SFA, participants with stronger phonological abilities are more likely to correctly name treatment stimuli and produce associated features. In both cases, strong lexical-phonological links would facilitate accurate responses to clinician probes, resulting in acquisition of targeted nodes.

Moreover, this learning would support generalization through spread of activation within the network: Once a node receives activation from the initial, target node, it can then send activation on to other linked nodes, and the effects of co-activation thus strengthen links to neighboring representations. Another consequence of correct responses is that participants move through the protocol more quickly, meaning that they get practice with more (in SFA) or more complex (in PMT) stimuli, which in turn increases the depth of processing. However, while the former explanation seems reasonable, it clearly cannot fully account for the findings, for the simple fact that phonological processing ability, as measured by phonological errors, had a negative association with naming abilities. Therefore, it seems likely that activation maintenance and/or higher order cognitive skills may play a larger role in acquisition and generalization than the integrity of network connections.

Several earlier treatment studies have found significant relationships between treatment outcomes and baseline verbal STM (Coran et al., 2020; Dignam et al., 2017; Kalinyak-Fliszar et al., 2011; Koenig-Bruhin & Studer-Eichenberger, 2007; N. Martin et al., 2020) and verbal and visual WM (Harnish & Lundine, 2015; Harnish et al., 2018; Majerus, 2018). In a treatment study of 34 PWA, Dignam et al. (2017) found that baseline verbal STM, as measured by the Hopkins Verbal Learning Test, significantly predicted acquisition (but not maintenance of acquisition), near-generalization, and maintenance of generalization one month following treatment. Likewise, Harnish et al. (2018) found that baseline measures of visuospatial WM, as measured by the backwards spatial span and a 1-back[6] test, and verbal WM, as measured by a backwards word span, were associated with generalization to untreated items following anomia treatment in

[6] Participants saw a sequence of novel two-dimensional shapes and indicated when the current shape was identical to the previous shape.

seven PWA. Specifically, those who scored higher on baseline cognitive measures also demonstrated generalization, while those who scored lower on these tasks did not.

These studies can be taken as evidence for the role of STM and WM in acquisition, generalization, and maintenance of treatment effects. That is, when people's ability to maintain and manipulate information is strong at baseline, they tend to demonstrate greater gains. These results support the hypothesis that the memory component of the SAPA 3 tasks, rather than the integrity of connections within the phonological level, contributed to the robust relationship between this measure and the outcomes. Recall that repetition is thought to rely on STM, while parsing and blending rely on WM. In the whole group SAPA 3 *post hoc* analyses, parsing ability emerged as the only significant predictor of acquisition, generalization, and maintenance[7]. Furthermore, in the severity *post hoc* analyses, the SAPA 3 remained a robust predictor across items for both immediate post-treatment and follow-up time points, and in the severe group it remained predictive immediately post-treatment. Taken together, these results point to the executive functioning involved in verbal WM being a key skill needed for synthesis of learning with spared linguistic schemata in both mild and severe PWA, and these higher-order processes may also facilitate maintenance of the effects of this integration at least three months after treatment termination in milder PWA.

Another finding of the whole group analyses was that increased phonological impairment was associated with increased likelihood of naming treated, untreated-related, and untreated-unrelated items post-treatment, especially among the severe group, and increased impairment also predicted maintenance of treated and untreated-related items. These results suggest that those who were more phonologically impaired at baseline were better at naming following

[7] The analyses for maintenance of acquisition and far-generalization revealed no significant predictors, but this is likely due to restrictions in variance due to the smaller total scores, relative to the composite SAPA 3 score.

treatment and at follow-up than those who were less impaired. The finding that those with greater impairment improve more is not new. Robey (1998) noted that moderately and severely impaired participants experienced large effects of treatment across the continuum of care. In a more recent analysis of response to treatment among 51 individuals with chronic aphasia, Kiran (2016) noted that those with more linguistic impairment made greater gains on outcome measures than those with milder deficits. While the current results could conceivably be attributed to ceiling effects among the milder participants (Scherr et al., 2016; Šimkovic & Träuble, 2019), this conclusion seems unlikely given that this effect is only seen for phonological impairment, and not the other measures of linguistic ability, in this sample. Furthermore, other treatment predictor studies have also found that decreased phonological processing ability was associated with greater post-treatment gains (Best et al., 2013; Eoute, 2010).

Best et al. (2013) attempted to understand the relationship between an orthographic/phonological cueing therapy, locus of impairment, and generalization to untreated-related items following eight hours of therapy delivered across eight weeks. To do this, the authors measured the baseline semantic and phonological processing abilities of 16 participants with aphasia. Semantic skills were quantified as the number of semantic errors on the better score of two input processing tasks – auditory word-to-picture matching or written word-to-picture matching – on the CAT. Phonological skills were measured using the proportion of phonological errors produced during a confrontation naming task. Although the semantic score is a measure of *input* processing, while the phonological score is a measure of *output* processing, this distinction is not critical, since unitary lexical-semantic representations exist for both input and output, according to the model outlined in the *Introduction*. Once these semantic and phonological proportion scores were derived, they were standardized, and scores above the mean

were considered markers of "better" processing, while scores below the mean were "worse." Next, the authors categorized their participants into one of four groups based on their semantic and phonological processing abilities: good semantics/good phonology; good semantics/poor phonology; poor semantics/good phonology; and poor semantics/poor phonology.

Following the intervention, Best and her colleagues examined the relationship between loci of impairment and degree of generalization to untreated items, and they found that participants in the good semantics/poor phonology group demonstrated the highest degree of generalization, relative to the other groups. Given these results, the authors concluded that strong lexical-semantic abilities were critical in being able to select the correct lexical item during the treatment, which explains why those with good semantic abilities improved. However, the fact that phonology was impaired was also important to generalization, because the orthographic and phonemic cues provided during naming helped strengthen lexical-phonological links not only to the target lexemes, but also to phonologically related lexemes. Participants who had better phonological processing skills would not have needed or benefitted from the cues; thus, connection strengths did not change.

The mechanism for generalization discussed in Best et al.'s study is likely the same as in the current study. That is, when lexical-phonological connections were weak, the act of repeating (in SFA and PMT) and manipulating (in PMT) phonological sequences in the form of words or nonwords strengthened those links. This strengthening not only improved connections for target items, but also spread to related lexemes. Furthermore, repeated activation over the course of treatment likely also improved the ability to maintain activation within the lexical-phonological network, such that the more distant, untreated-unrelated items could also be named immediately post-treatment. Similar to the Best et al. study, the participants in this study who demonstrated

better phonological processing ability at the outset of treatment likely did not receive the same benefit from phonological output tasks during therapy and, consequently, did not demonstrate the same degree of acquisition or generalization.

The foregoing discussion explains the logical relationship between increased phonological impairment and greater naming accuracy for treated, untreated-related, and untreated-unrelated items immediately after treatment. What is more difficult to explain is the significant relationship between greater baseline semantic impairment and better post-treatment naming accuracy for all untreated items immediately post-treatment, as well as the significant relationships between semantic impairment and maintenance across item types. Although several possibilities exist to explain these results, the most straightforward explanations are that, (1) in the case of generalization observed across items, more broken systems that receive *any* activation are more susceptible to change and, (2) in the case of maintenance across items, naming errors borne of semantic impairment take more time to resolve.

In the first instance, treatment-induced activation would course through the language network and, with each subsequent activation, broken lexical-semantic links would be incrementally repaired (while no change would occur in intact links). This repair would be more important for untreated than treated items because the treated items have the benefit of direct practice during treatment, especially in the case of SFA; the ability to name items that were not targeted during treatment, however, relies solely on spread of activation to restore impaired links. For the maintenance effects, it could be that the idiosyncrasies of semantic networks for individual words mean that consolidation of learning takes much longer than for phonological forms, which are relatively invariant across speakers (Castro et al., 2020). Because the time to consolidation is longer, it is possible that those who are more impaired show particularly robust

gains at the follow-up point. The hypotheses about generalization and maintenance presented here are supported by the findings of the severity *post hoc* analyses, which demonstrate that this relationship between semantic impairment is especially strong in the severe group, while this pattern is not observed in the mild group.

Although the arguments for the relationships between the semantic and phonological impairments presented here are entirely plausible, alternative possibilities exist. First, it could be that these findings are driven by the PMT group, whose results closely resemble the results for the whole group. The benefit of increased semantic and phonological impairments arise from a lack of interference during the treatment, and this interaction is discussed below. If these relationships were strong enough, they could have persisted in the whole group analyses. On the other hand, this result could have also been driven by the SFA group, who benefitted from direct treatment of lexical-semantic lesions and the phonological output required in the protocol. Unfortunately, it is difficult to rule out these possibilities, and little insight is gleaned from the results of the *post hoc* analyses. Thus, more research is needed to understand the role of semantic and phonological impairment in response to treatment.

Before moving on to discuss relationships in the PMT group with more depth, the nonsignificant results bear some consideration as well. Neither the CAT comprehension T-score nor the SAPA 2 were predicted to significantly relate to outcomes in the whole group, primarily on the basis that these skills were not seen as essential to the larger SFA group. While the predictions were correct, it must be noted that the relationship between the whole group and the SFA group was overstated: Neither of these measures are significant in the main PMT and SFA group analyses, and the severity *post hoc* tests reveal that, while the CAT T-score predicts performance in near-generalization for the mild whole group, this association is not present in

the mild SFA group. On the other hand, the SAPA 2 predicts maintenance of near-generalization in the severe whole group, and the same relationship is also observed in the severe PMT group.

In any case, comprehension ability was likely not important overall because, in both treatments, the clinicians would have provided cues, models, and other assistance to the participants to facilitate their ability to complete the tasks. Thus, so long as participants had the requisite skills to carry out the tasks, they would have been supported in understanding what they were meant to do. Given this explanation, it is unclear why the CAT T-score is only predictive for the mild whole group's ability to name untreated-related items immediately post-treatment. It is possible that, among those who were less impaired, there was some benefit conferred by better lexical-semantic processing, such that spread of activation to semantically-related neighbors during treatment facilitated naming immediately post-treatment. This possibility is supported by the significant *negative* relationship between semantic impairment and post-treatment near-generalization in this same group (i.e., those with less semantic impairment in the mild whole group had better naming of untreated-related items immediately post-treatment).

As for the SAPA 2, it is possible that, when considering treatment in general, the skills represented by this subtest may be subsumed by the SAPA 3. As a reminder, the key differences between the two subtests are that one of the SAPA 2 tasks explicitly requires verification of a lexical entry (the lexical decision task), while the SAPA 3 focuses on the relationship between input-output processing, and the tasks could theoretically be completed without any influence from lexical-semantics. Incidentally, this difference may explain the significant relationship between the SAPA 2 and maintenance of near-generalization in the severe whole group: If this test somehow represents the nexus between lexical-phonology and lexical-semantics, then being able to link phonological forms with semantic representations may be important for being able to

transfer skills that were trained in the treatment phase with lexical items that were related to treated items after consolidation has occurred in the system. Other than this difference, however, the SAPA 2 and SAPA 3 cover similar territory, in that the former measures input lexical-phonological verbal STM, and the latter measures input-output lexical-phonology and verbal STM and WM. Thus, in the overall group, the SAPA 3 might subsume the SAPA 2. The similarities between these subtests may also explain why the SAPA 2 is not generally predictive within the severity *post hoc* analyses.

PMT Group Findings

The SAPA 3 significantly predicted acquisition and near-generalization in the PMT group. Earlier, three possible explanations for the predictive abilities of the SAPA 3 were offered: (1) The SAPA 3 represents the integrity of links within the phonological level of the language network; (2) the SAPA 3 represents phonological STM (i.e., the ability to maintain phonological representations); or (3) the SAPA 3 represents verbal WM (i.e., the ability to maintain *and* manipulate phonological representations). Given that the outcomes are predicted by both greater phonological impairment in the main analyses and real word parsing and blending in the SAPA 3 *post hoc* analyses, the most likely interpretation is that the SAPA 3 represents verbal WM.

As discussed above, parsing and blending tasks draw on WM and other executive function skills (McBride-Chang, 1995). Hickok and Poeppel (2004) suggest that sublexical processing requires executive functions and that "verbal working memory (and perhaps WM in general) can be viewed as a form of sensory-motor integration" (p. 87). Viewed from this perspective, it makes sense that increased WM would contribute to outcomes, given the intensive input-output processing (i.e., sensory-motor integration) required by PMT.

As if to highlight the need for WM components, the SAPA 3 *post hoc* analysis revealed that real word blending and parsing together significantly related to acquisition, near-generalization, and maintenance of these two outcomes, while only real word parsing significantly related to far-generalization and maintenance thereof. Research has indicated that blending is easier than parsing (Schatschneider, Francis, Foorman, Fletcher, & Mehta, 1999; Stahl & Murray, 1994; Yopp, 1988), so this finding further supports the possibility that more intensive processing abilities may be needed as the goal moves from acquisition to far-generalization. Although this last point requires validation through more research, a tentative explanation for this finding is that increasing WM ability, as measured by blending and parsing tasks, reflects an increasing ability to apply attentional processes to linguistic tasks. The ability to apply attentional processes, in turn, facilitates the participants' ability to engage with the tasks, which results in even greater activation transmission and maintenance within the linguistic network (Hula & McNeil, 2008).

In a theoretical overview and summary of the literature to that point, Hula and McNeil (2008) suggested that language processing occurs in the context of a limited-capacity, central resource. From their perspective, linguistic tasks can never truly occur in isolation, because internal (thoughts) or external (background noise) factors are always in competition for resources. In neurologically healthy individuals, these distractions pose little difficulty for carrying out the language task, but in PWA and other people with brain damage, these distractions may be difficult to overcome, and task performance suffers as a result. The presence or absence of competing processes accounts for the variable language performance seen among PWA (McNeil, 1983; McNeil, Odell, & Tseng, 1991). Hula and McNeil (2008) synthesized the results of attentional studies with healthy individuals, people with closed head injury, PWA, and

the IA model of language processing: They suggested that phonological encoding, in particular, may be the most attention-demanding aspect of language production because it was the most susceptible to interference from competing tasks. This view is supported by a number of studies that have identified attentional impairments in PWA (Carbonell, Valdois, Charnallet, Lyard, & Pellat, 1991; Christensen, Wright, & Ratiu, 2018; Hunting-Pompon, Kendall, & Bacon Moore, 2011; Murray, 2002; Tseng, McNeil, & Milenkovic, 1993; Villard & Kiran, 2017). They furthermore conclude that any treatment that aims to remediate attentional processes in aphasia must pair the attention and language components.

Evidence in the aphasia treatment literature exists to support the relationship between attentional processing and language improvement (Coelho, 2005; Crosson et al., 2007; Geranmayeh, Brownsett, & Wise, 2014; Peach, Nathan, & Beck, 2017). In an interesting study that compared two treatments – an *intention* and an *attention* treatment – Crosson et al. (2007) administered a naming treatment with vanishing attentional cues to 34 people with moderate to profound nonfluent aphasia. In Phase I, visual and auditory cues to pay attention were presented for 4 seconds before each of 50 stimulus pictures. In Phase II, only the auditory cue was presented before the stimulus, again for 4 seconds. In Phase III, the participants heard the auditory cue for only .5 seconds before being prompted to name the stimulus picture. All participants demonstrated significant improvement over baseline for treated and untreated-related items, which the authors did not expect (the intention protocol was the experimental treatment). Given that the treatment essentially involved pairing an increasingly rare cue to attend with a linguistic task, the study results could reflect improved abilities to allocate attention to the linguistic task – a hypothesis that is in line with Hula and McNeil's (2008) suggestions.

Taken together, these considerations point to (1) the attentional component of WM underlying the relationship between the SAPA 3 and the current outcomes, (2) the importance of attention in aphasic language performance, and (3) attention as a mechanism of treatment-induced language improvement. If the PWA in this study indeed had spared abilities to allocate attention during baseline tasks, those same abilities could be applied during the PMT protocol, which in turn would enhance their performance with the treatment stimuli – *despite* their phonological processing impairments. Of course, this conclusion requires validation from other treatment studies that include measures of verbal STM, WM, and attention; thus, this hypothesis is speculative. Nevertheless, this explanation seems the most reasonable in the absence of more information.

Although the attentional/WM processing hypothesis is compelling, the alternative explanations for the findings bear some consideration. First, the SAPA 3 is clearly dependent on activation transmission between input phonology, output phonology, and lexemes; without some connection between these representations, the participant could not complete the tasks. However, as discussed earlier, the integrity of the links cannot be the only factor driving the relationship between the SAPA 3 and outcomes, for the simple fact that those with greater phonological impairment at baseline demonstrated better naming performance post-treatment than those with less phonological impairment. Since the impairment score is a measure of output phonology, SAPA 2 scores could indicate input phonological processing, but no relationship existed in the overall group, either. Ultimately, these results suggest that something beyond simply phonological ability drove this association.

The other possibility is that the SAPA 3 measures phonological STM. If this were the case, then this relationship would make sense, given the theoretical relationship between verbal

STM and language learning (Cowan, 2008; Gupta & MacWhinney, 1997), and the demonstrated relationship between verbal STM and acquisition difficulties in adults with brain damage (Baddeley, Papagno, & Vallar, 1988; Coran et al., 2020; Dignam et al., 2017). However, three pieces of evidence point away from this conclusion. First, the SAPA 2 is certainly a measure of phonological STM, as it entails perception and maintenance of phonological forms sufficient to make judgments about lexicality, rhyme, or similarity. However, this measure is not associated with outcomes in the overall PMT group. Second, the SAPA 3 *post hoc* analyses revealed significant relationships between blending, parsing, and the outcomes, but not repetition – a classic measure of verbal STM in psychological, linguistic, and aphasia research (Allen, R. C. Martin, & N. Martin, 2012; Cowan, 2008; Daneman & Merikle, 1996; Engle, Tuholski, Laughlin, & Conway, 1999; N. Martin & Ayala, 2004; Murray et al., 2018; Salis et al., 2015). Finally, Zimmerman et al. (2019) evaluated the predictive value of verbal STM, as measured by word and digit spans, in this specific cohort of participants, and these measures of STM did not significantly predict acquisition, generalization, or maintenance. Taken together, these findings suggest that the relationship between SAPA 3 and the outcomes relies on something beyond STM, as well.

An important caveat on the foregoing discussion is that the SAPA 2 was not universally insignificant. In the severity *post hoc* analyses, significant relationships were found between the SAPA 2 and maintenance of acquisition and near-generalization in the severe group. These associations, in and of themselves, do not invalidate the argument that WM/attention is likely driving the relationship between the SAPA 3 and the outcomes. Instead, these associations may highlight the specific need for the nexus between phonology and semantics. That is, the ability to link phonological forms to their associated semantic features may be a prerequisite for

maintaining the network links that were reinforced during the treatment phase (which would also explain why this finding does not extend to the untreated-unrelated items). The integrity of this "phonological-semantic junction" may be reflected in the SAPA 2 score.

Returning to the overall results, increased phonological impairment significantly related to acquisition, near-generalization, and maintenance of these two outcomes. As noted earlier, this phonological treatment explicitly targets impaired links within the language network. Best et al. (2013), Jacquemot et al. (2012), and Nickels (2002) all suggest that treatments that target the locus of impairment seem to demonstrate greater gains (but see also work by Wambaugh and colleagues that suggests that targeting the more intact capacity can result in greater benefits; J. L. Wambaugh, 2003; J. L. Wambaugh, Linebaugh, Doyle, & Martinez, 2001). Jacquemot et al. (2012) conducted a case study in which they identified their participant's locus of impairment at the phonological output lexicon, and then conducted 9 hours each of treatments targeting different levels of the language network: semantic, input phonology, output phonology, and phonological output lexicon. They found that the approach that targeted his locus of impairment resulted in the greatest acquisition and generalization (there was no information on maintenance). This study relied on a slightly different language model than the one considered here, and it certainly had some limitations; e.g., the phonological output lexicon treatment was a naming paradigm, which entailed semantic processing. Nevertheless, this study is a first step that suggests that matching the impairment and the treatment is more effective than when the locus of impairment and treatment target are mismatched.

The argument that PMT restored targeted and adjacent links is bolstered by the finding that phonological impairment was not related to far-generalization: None of the untreated-unrelated items had the same 3-phoneme sequences as the treated items in either group, so the

benefit of any lexical-phonological strengthening would be confined to words that are phonologically similar. Such an explanation does bring up the question of whether lexical selection demands a quorum of phoneme segments, such that phonemes must be trained in sequences of three or more to improve access to lexical items. This is not a novel question, as the nature of lexical selection during verbal input and output tasks has been the subject of research since at least the 1970s (Vitevitch & Luce, 2016). Marlsen-Wilson (1987) proposed cohort theory, which suggested that, during word recognition tasks, lexical items are identified sequentially – phoneme by phoneme – until enough segments have been presented to correctly identify the target lexeme. Cohort theory has been refined over the last 30 years to become the foundation of theories of phonotactic probability (PP) and neighborhood density (ND). In studies of the effects of PP and ND on word recognition and production, these variables slow down recognition but speed up production (Vitevitch & Luce, 2016): Words with similar phoneme sequences clustered together creates competition during decoding tasks, but these same clusters boost selection during production tasks via richly enforced lexical-phonological links. In the context of this study, this phenomenon provides further support for the assumption that intervention improved connections – and therefore activation transmission – among treated items and their neighbors, but that the more distant, untreated-unrelated items could not benefit from strengthened phonological links.

Semantic impairment had a significant negative relationship with all outcomes in this group, such that more impairment was associated with greater post-treatment naming accuracy. In this case, it is probable that semantic processing actually impeded phonological processing. This idea is based on the processing demands imposed by the lexical-phonological characteristics of the PMT stimuli. Essentially, studies of healthy and aphasic adults have found that words in

dense neighborhoods were processed more slowly due to competition effects during word recognition and repetition (Dell & Gordon, 2003; Vitevitch & Luce, 1998; Vitevitch & Luce, 1999; Vitevitch & Sommers, 2003). However, when the lexical effect is removed (i.e., by using nonwords), higher PP leads to faster processing during production tasks (Vitevitch & Luce, 1999).

To explore the neighborhood density effect during production and comprehension in PWA, Dell and Gordon (2003) conducted network simulations of word recognition and production in normal and lesioned networks. They also manipulated the words so that some had high ND, some had low ND, and some words were hermits. In the word recognition conditions, ND negatively affected correct target selection across all network configurations. In the word production conditions, higher ND corresponded with fewer errors overall but, interestingly, the normal network demonstrated a high rate of semantic errors for the high ND words and almost no phonological errors. The authors took this finding to mean that competition among words occurred at the lexical-semantic level. Although this study did not extend to repetition, the implication is that high ND creates competition during input processing and, in the context of normal (or unimpaired) semantic processing, high ND also creates competition during output tasks. In the context of the present study, such an assumption would mean that competition was an issue during input-output tasks in PWA with spared semantic processing.

Another possible explanation for these findings is that spared semantic processing abilities interfered with the treatment's focus on phonological processing of the stimuli. In the PMT protocol, real word stimuli were handled exactly the same as nonword stimuli, so that they were subjected to parsing, blending, repetition, elision, etc., and the clinician did not address the word meanings at all. In this case, better semantic abilities might have resulted in implicit

semantic processing, which would have in turn led to resources being diverted away from phonological processing. Tentative evidence for this hypothesis comes from the mild PMT group results: There was a negative association between SAPA 2 scores and acquisition only; i.e., lower SAPA 2 scores significantly related to greater naming accuracy for treated items. If, as suggested earlier, one of the distinctive characteristics of this subtest is lexical verification, it appears that a reduced ability to complete lexical-semantic searches results in better acquisition. The fact that this effect is only seen with the treated items further supports this explanation, since the results imply that only the items that were exposed were affected. The relationship between increased semantic impairment and increased post-treatment naming ability is seen across item types for the severe group, and this is likely because this group is more sensitive to any processing interference.

These competing hypotheses – ND interference versus diverted resources – may not be mutually exclusive, especially since they both point to some type of competition between semantic and phonological processing. A future investigation could tease apart these possible explanations by varying the stimuli characteristics (i.e., high ND and low ND treatment items) and evaluating phonological and semantic processing abilities in one go. If the relationship between semantic impairment disappears in the context of low ND stimuli, then the competition effects among words in dense neighborhoods is the likely culprit. If the semantic impairment persists as a significant predictor regardless of the stimuli ND, then diverted resources would be a more likely explanation.

Finally, the CAT was predicted to be a significant predictor because of the prompts for metalinguistic analysis of the stimuli, but this was not the case in the PMT group as a whole. However, this finding *was* significant in PMT participants with milder impairment. Perhaps then,

comprehension is not important in general, because of the support provided by the clinician, but for milder participants, being able to engage with the stimuli in a more meaningful way actually did facilitate generalization. This finding bolsters the argument for the role of higher order processing in treatment effects, because the metalinguistic tasks require attention and executive function skills (Hernández-Sacristán, Rosell-Clari, Serra-Alegre, & Quiles-Climent, 2012). The SAPA 2 was also predicted to significantly predict naming outcomes, but this relationship was not borne out in the main analyses. As noted before, the distinctive skills involved in the SAPA 2 were likely subsumed by other measures. In this case, the SAPA 3 might capture some of the input and STM components of the SAPA 2, and semantic processing may actually also represent the influence of lexical-semantics.

SFA Group Findings

In the SFA group, the SAPA 3 significantly predicted acquisition, and near- and far-generalization. Given all of the preceding discussion about the relationship between SAPA 3 and higher order cognitive abilities, there are two possible mechanisms underlying these results. The first possibility is that higher scores on the subtest reflected higher WM abilities, which in turn facilitated acquisition and generalization. The second possibility is that higher scores reflected higher attention and/or executive function abilities, which increased use of a word-finding strategy across stimuli. Both explanations have been offered in the literature (Efstratiadou et al., 2018). The first author takes the view that SFA does indeed improve language processing, or at least results in some language improvement and some compensation.

In the scenario in which the linguistic network actually improved as a result of SFA, evidence of this mechanism might come in the form of associations between other linguistic measures and outcomes, effects of linguistic characteristics of the words, and/or changes in

performance on measures beyond naming accuracy. In the current study, higher SAPA 2 scores, which measured input phonological STM with influence from semantics, predicted greater near-generalization in the mild SFA group immediately post-treatment. Likewise, the SAPA 3 *post hoc* analyses demonstrated that higher word and nonword repetition – measures of input-output phonological STM – were associated with higher naming accuracy for treated and untreated-related items in the whole SFA group immediately and three months post-treatment. These results suggest that a true association exists between the ability to maintain linguistic information, which is critical to language acquisition (Gupta & MacWhinney, 1997; Gupta, N. Martin, Abbs, Schwartz, & Lipinski, 2006; N. Martin & Gupta, 2004), and naming ability following the treatment. The SAPA 3 *post hoc* also revealed a significant relationship between nonword parsing and naming accuracy for untreated-unrelated items immediately post-treatment. Although this finding could still point to the requisite skills for applying a word-finding strategy, it likely underlies the skills necessary to effectively maintain a representation while also actively identifying and verbalizing features associated with it. Cumulatively, these findings suggest that SFA actually affected the language network.

The strong relationships between increased baseline phonological impairment and increased near- and far-generalization immediately post-treatment further suggest that change within the linguistic system occurs as a result of SFA. These associations may arise from repaired lexical-phonological links as a result of language production tasks during the treatment: Participants are required to name the targets, as well as words that each participant associates with each target, and these verbalizations must improve connection strengths – especially as impairment level increases. Incidentally, the fact that participants selected features themselves, possibly based on how readily retrievable the words were in a given instance, may explain why

phonological impairment generalizes to untreated-unrelated items in SFA, but not in PMT. Verbalizing a variety of words that were not constrained to the same high ND low PP characteristics as the PMT stimuli meant that more phoneme sequences could be practiced, which may have boosted word-retrieval ability for words that were not trained immediately post-treatment. However, this "convenience-based" word selection has the knock-on effect of any improvements in performance being temporary, since words/phonological sequences were not trained to criterion, as would have been the case in PMT. Said another way, SFA participants with greater baseline phonological impairment may have experienced broad but shallow improvements in phonological processing ability owing to the production-heavy nature of the treatment, and these improvements led to increased naming of untreated items immediately post-treatment. However, because the phonological training was shallow, any network changes did not persist to the follow-up point (Gupta & MacWhinney, 1997; Hinton & Plaut, 1987; Plaut, 1996).

Although it is unclear exactly how many repetitions are needed to induce long-term maintenance of an item, it has been widely acknowledged that more is better. Some of the foundational research that informs this understanding comes from animal models of neurorehabilitation, and the principle that "repetition matters" was codified in Kleim and Jones' influential white paper on neural plasticity (2008). Computational instatiations of connectionist language models have provided additional, language-specific support for the notion that more presentations improve retention of trained items in lesioned networks. For example, Hinton and Plaut implemented a connectionist model that incorporated short-term and long-term relearning weights. The short-term weights relearned information quickly and, importantly, improved the weights of related nodes; however, the relearning that occurred in the short-term weights

decayed relatively quickly. The only method for preventing total loss of relearned information was to train to saturation, at which point more permanent weight changes stabilized in the long-term weights. Gupta and MacWhinney (1997), also proposed a similar mechanism for long-term retention in new learning, but specifically stated that, for phonological sequences to be learned, they must be repeated, either in naturalistic contexts or through rehearsal. In this light, the immediate post-treatment – but not follow-up – improvements in naming of untreated items experienced by those with greater baseline phonological impairment were a direct result of short-term phonological learning. In any case, this short-term learning supports a restorative mechanism of change in the SFA group.

It is worth noting that the significant relationship between greater baseline phonological impairment and better post-treatment naming ability in this study is incongruent with previous findings that suggest that *less* impairment is better in SFA (Efstratiadou et al., 2018). Generally speaking, many of the studies that have found that less impairment relates to better outcomes have used different measures of severity than this study. For example, in many studies of SFA, baseline severity is quatified by overall naming ability (e.g., Quique et al., 2018) or semantic impairment only (Julie L. Wambaugh, Mauszycki, Cameron, Wright, & Nessler, 2013). Thus, the difference may be attributed to differences in measures. Furthermore, the relationship between less impairment and better outcomes is not universal (e.g., Boyle, 2010) and, thus, the finding of the opposite result in this study is not unique. That being said, it is possible that the way the phonological impairment score was calculated may have resulted in paradoxical results; this issue is take up in *Limitations*. Either way, more research into this effect is warranted.

Other evidence for restorative, rather than compensatory, effects comes from typicality studies in SFA. Wambaugh, Kiran, and colleagues (Gilmore, Meier, Johnson, & Kiran, 2020;

Kiran & Thompson, 2003; Wambaugh, et al., 2013) have conducted a number of studies over the past 20 years that explore the influence of typical versus atypical exemplars on acquisition and generalization in SFA. For example, in a recent study, Gilmore et al. (2020) administered typicality-based treatment to 27 PWA. Half of the trained and untrained items were typical of a category (e.g., *robin* for the "bird" category), while the other half of the trained and untrained items were atypical (e.g., *ostrich* for the "bird" category). Following treatment, the authors found that participants named typical untrained items significantly better than atypical untrained items, which suggests that the participants experienced improvements in linguistic processing, rather than simply learning to apply a word-finding strategy. If they had learned a compensatory strategy, they should have been able to name atypical items well, too. Furthermore, participants demonstrated gains on tests of semantic processing and the Western Aphasia Battery, which also suggests improvement of the linguistic network.

On the other hand, a number of studies have asserted that compensation is the mechanism (Anders & Bos, 1986; Boyle, 2010; Boyle & Coelho, 1995), while others have assumed that linguistic processing improves but also questioned whether compensation contributes to treatment gains (Evans et al., 2020; Silkes, Fergadiotis, Graue, & Kendall, 2020). Two recent analyses completed on data from a single large-scale SFA program highlight the tension between these two views. In the first, preliminary study, Gravier and colleagues (2018) explored treatment factors (e.g., dose, number of patient acts) as predictors of treatment response in 17 PWA. They found that the number of participant-generated features significantly related to treatment acquisition and generalization, such that more features generated during the treatment phase predicted greater post-treatment naming. This study was taken as evidence of linguistic restoration. However, in a follow-up study using data from the original 17 participants plus an

additional 27 PWA, Evans et al. (2020) found that only treated items related to the number of features generated. The lack of generalization to untreated items, and a lack of relationships with the types of features and stimuli categories, led the authors to wonder whether the participants had learned to self-cue for the treated items.

Silkes et al. (2020) also analyzed changes on measures of discourse performance in the same cohort of participants as the current study and found that, although the SFA group demonstrated a large increase in their percent of correct information units, this change was not significant. Results from pre-treatment to the 3-month follow-up revealed that participants performed close to baseline on discourse. Like Evans and his colleagues, Silkes and her colleagues questioned whether the immediate post-treatment bump in discourse performance simply reflected application of self-cueing strategies that participants had recently learned during the treatment phase.

Although the results of these two studies certainly call into question the underlying mechanism of improvement, it is likely that the truth lies somewhere in between these extremes. To better understand what might be happening, future analyses should examine more qualitative changes in performance, such as error evolutions or the number of circumlocutions observed during naming and discourse tasks. Some authors have advocated for post-treatment error analyses as an insight into network change (e.g., Kiran & Thompson, 2003; Walker, Hickok, & Fridriksson, 2018). To date, three analyses of error evolution following PMT have demonstrated decreases in omission errors and circumlocutions and increases in target-related semantic and phonological errors, demonstrating that PMT indeed induces network improvements (Kendall et al., 2013; Minkina et al., 2016; Minkina et al., 2019). If similar shifts were revealed in analyses

of post-treatment naming performance among SFA participants, such findings would serve as more tangible evidence of linguistic restoration.

Clinical Implications

This study has several clinical implications, the first being treatment selection for individual PWA. Generally speaking, clients who demonstrate spared verbal STM and WM, as demonstrated by performance on phonological awareness tasks, will likely have the greatest response to treatment, regardless of whether the clinician chooses a semantic or phonological treatment approach. For those clients who demonstrate decreased STM/WM ability, treatment should include tasks that target those specific skills, using linguistic stimuli (Kalinyak-Fliszar et al., 2011; Majerus, 2018). With regard to specific treatment approaches, the results from this study suggest that those with more severe phonological and semantic impairment, as demonstrated by the number of errors on a naming test, are good candidates for PMT. The protocol's specific focus on fundamental phonological processes, coupled with the intense STM/WM aspect of the phonological awarenesss tasks, means that clients who are more impaired will experience greater gains than with SFA. On the other hand, because baseline impairment levels were generally not associated with outcomes in the SFA group, this treatmern seems well-suited for people with milder aphasia, since the treatment requires adequate WM to complete the feature generation task, adequate semantic skills to identify appropriate features for each target item, and adequate phonological skills to verbalize the identified features. Selecting treatments for clients following these criteria is likely to result in increased acquisition, generalization, and maintenance of treatment effects.

The second implication of this study is that PMT is appropriate for clients with phonological and semantic impairment, while SFA may be best for clients with spared semantic/impaired phonological skills. Because PMT directly remediates impaired phonological processing deficits through manipulation of phonemes and phoneme strings, diminished semantic processing ability actually facilitates the restorative effects of the treatment protocol. This finding is in line with previous literature that suggests targeted treatment of phonological processing, especially when phonology is the locus of impairment (Jacquemot et al., 2012). Alternatively, clients who present with spared semantic abilities but phonological impairment benefit most from SFA because of their ability to generate semantic features, as described above. However, impairments in phonological processing are responsive to the phonological stimulation that occurs during the naming and feature generation tasks that occur during treatment. This suggestion echoes findings from Best et al. (2013). Matching baseline linguistic profiles and treatment approaches in this way seems to results in the best outcomes, and these results align with previous findings in the literature (e.g., Best et al., 2013; Eoute, 2010; Jacquemot et al., 2012; Raymer et al., 1993).

Limitations

This study has several limitations. The most important limitation is that these analyses only evaluate the relationship between the predictors at baseline and performance post-treatment, without accounting for change over baseline. That is, the models did not include a variable for pretreatment performance on the outcomes. As noted in *Methods – Analysis*, when baseline performance (quantified as either performance by item or as a change score) was included in the model, it was the only significant predictor of outcomes. Thus, the results discussed in this study do not represent how person-level differences moderate treatment-related change. Of course, this

limits the conclusions that can be drawn from this study. Namely, *this study does not speak to the relationship between the predcitors and magnitude of treatment-induced change*. Future studies must incorporate baseline performance on the outcomes to better understand the question of the role of baseline linguistic characteristics in response to treatment.

Although mixed-effect logistic regression is a useful approach for increasing the sample size, it clearly cannot eliminate problems that arise secondary to small numbers of participants. In the case of this study, some models remained under-powered, despite the logistic approach. The severity *post hoc* analyses were especially prone to convergence failures because of the small numbers of participants, coupled with high variance, in the SFA group. This issue prevented the formation of more complete views of linguistic factors that influence outcomes across severity levels, and limits the interpretability of the findings. Future studies should endeavor to have as many participants as possible to be able to draw conclusions about these relationships.

The high degree of variance in the SFA group raised two key questions. The first question is why the SFA group demonstrated so much more variability than the PMT group. The reason for the large distribution of the residuals remains unclear at the moment, as a visual inspection of the data and descriptive analyses did not reveal the source of the variability. A tentative explanation may come from the severity *post hoc* analyses, which demonstrated that the variance was more extreme in the severe SFA participants: Only one-third of participants in the SFA group received the high-frequency stimulus sets, as compared to 46% in the PMT group. This suggests that there were fewer "severe" participants in the SFA group – yet the group mean did not significantly differ from the PMT group. This suggests that the distribution of at least some of the baseline scores may have been negatively skewed, so that most of the scores were

around or above the mean, but a greater spread of a few scores at the tail of the distribution may have pulled the mean down. This hypothesis requires further exploration, but if it is indeed a factor in the results, future studies should strive for samples with more symmetrical distributions. Furthermore, future studies should define severity based on some criterion other than stimulus set. For example, using baseline BNT scores and/or evaluating the top and bottom thirds of the distribution may have alleviated issues of extreme variance.

The second question that the issues of variance raises is how the exclusion of these data points influences the results. When more extreme values are excluded from the analyses, the models represent expected patterns of performance for the overall group, but outliers may contribute valuable information that influences the interpretation of the results. In this study, all of the models were evaluated before outliers were removed: In every model that converged, exclusion of outliers did not dramatically change the results. That is, predictors that demonstrated significant relationships with the outcomes did so whether outliers were present or not. The coefficients did change, but this is to be expected, but the direction of the relationship and the overall magnitude were relatively consistent.

The measures in this study raise several questions related both to interpretability of the study findings and application in clinical contexts. First, the method for calculating the micro-linguistic scores (i.e., semantic and phonological impairment) may not have been optimal. In studies of computational instantiations of aphasic deficits, Dell and colleauges (Dell & O'Seaghdha, 1991; Dell et al., 1997; Foygel & Dell, 2000; N. Martin & Saffran, 1997) often derived impairment scores in a manner similar to the current study – that is, those studies used proportion scores. However, clinical studies of aphasic language performance have relied on raw error scores to determine degree of phonological or semantic impairment, relative to the study

sample (e.g., Best et al., 2013; Eoute, 2010). In this study, some interesting patterns of results emerged, such as the positive relationship between increasing impairment and greater post-treament naming ability or impairment scores being the strongest predictors in the more severe groups. These patterns call into question the wisdom of using a proportion score rather than a raw score, since using the total number of errors as the denominator to determine the proportion may have skewed this measure somewhat. For example, the possibility exists that introducing all other error types into this calculation meant that each of the impairment scores become more of an index of overall anomia severity, rather than tied solely to semantic or phonological processing. For future clinical studies, raw scores are recommended.

Another limitation of this study is that the link between outcomes and verbal STM/WM are based on phonological – but not semantic – tasks. While STM and WM operate similarly across tasks, one might expect that performance on phonological STM/WM tasks may look different from performance on semantic WM tasks. These differences might arise because the degree of impairment that needs to be overcome in each level of the network may dictate the ability to carry out the memory task. For example, if someone demonstrates severe lexical-semantic deficits, they might perform more poorly on a WM task because more resources are needed to maintain the semantic representations. Conversely, if this same person has relatively spared phonological abilities, they may perform better on the WM task because the basic task of maintaining phonological representations is easier, which frees up processing capacity. Given that there was no semantically-based STM/WM measure in this study, it would be interesting to investigate whether strong associations between STM/WM persist, regardless of the level of the linguistic network being tapped, and how the semantic and phonological STM/WM tasks would interact with or relate to measures of semantic and phonological impairment levels.

From a clinical perspective, the measures selected for this study may or may not correspond to tools that are widely used in clinical contexts, which limits its applicability to those settings. Specifically, this study relied on the comprehension test of the CAT, two subtests of the SAPA, and proportion of speech errors from the main study's primary outcome measure (i.e., the untrained-unrelated items). The CAT is a well-motivated and thorough test of various linguistic domains, but it has only been validated with British English speakers and little evidence exists to suggest its widespread clinical use in the United States. Likewise, although the SAPA is standardized on American English speakers, it is unclear that clinicians use (or would be able to interpret) this tool. Finally, micro-linguistic scores were derived from error types and frequencies on naming of the untreated-unrelated items. This approach to determining semantic and phonological processing in their clients is likely to be infeasible for most clinical SLPs, who work under high productivity constraints. Furthermore, given that the items on which these errors occurred are specific to this study, the micro-linguistic scores may not be readily transferrable to scores derived from more common measures such as the BNT. Finally, regarding the micro-linguistic scores, the criteria for determining semantic and phonological errors were highly constrained, which could have resulted in scores that were not fully representative of participants' impairments.

Finally, although the mixed-effect logistic regression models in this study revealed statistically significant relationships between the linguistic predictors and the outcomes, it is unknown whether these findings are clinically significant. That is, it is unclear how meaningful these relationships would be in real-world, clinical contexts. Some researchers have done work to determine the clinical significance of effect sizes derived from linear regression models, and a good recent example of this work comes from Gilmore et al. (2019), who suggest that a change

of 3.30 or more on the BNT represents clinically significant change. However, these same kinds of conclusions cannot yet be drawn for mixed-effect logistic regression, possibly because the application of this technique in aphasiology is relatively new. Thus, for example, while milder PMT participants with higher baseline SAPA 3 scores were three times more likely to correctly name treated items post-treatment than those with lower baseline scores, it is unclear whether these participants would have reported greater participation in activities of daily living. Future studies must establish benchmarks of the clinical significance of mixed-effect logistic regressions effect sizes, especially as this approach grows in popularity, for our results to be meaningful beyond research contexts.

Future Directions

The results of this study lay the groundwork for several follow-up studies. An initial starting point may be to conduct a similar study with measures of linguistic and non-linguistic cognition to determine whether factors such as executive function or WM predict outcomes. The findings from such a study could either support or refute the current study's conclusion that the verbal WM components of the SAPA 3 tasks were the basis of that measure's relationship with the outcomes. Another avenue of research is to explore linguistic predictors among those who demonstrate significant improvement over baseline (i.e., "responders") and those who do not ("non-responders"). Although the distinction between the current study and a responder/non-responder study is very fine, there are valid reasons for carrying out this slightly different study: The current study asked about predictors of naming accuracy *regardless* of responder status, while a responder/non-responder study would investigate the difference between those who experience gains and those who do not. Both types of studies are needed to help researchers and clinicians alike understand how best to support PWA in their language goals.

This study revealed questions that are less directly tied to predictors of treatment response, but that are nonetheless interesting and potentially useful areas for exploration. Phonological awareness (PA) in healthy and aphasic adults is not well understood, but some evidence (e.g., Kendall et al., 2010; Madden et al., 2018) suggests that PA skills may be compromised in PWA. Future studies should systematically explore the relationships between PA skills, WM and executive function, and linguistic ability in PWA and neurologically healthy adults. Furthermore, more fine-grained analyses of the predictive value of PA skills to treatment outcomes could further inform theory and clinical practice. The other issue that this study was not able to unequivocally resolve was the underlying mechanism of change in SFA. It was noted in the discussion above that analyses of anomic error evolution following SFA may shed more light on whether this protocol is compensatory, restorative, or some combination of the two. Other work that would be useful in this regard is a qualitative analysis of changes in word-finding strategies employed by PWA from pre-to-post-treatment, since this type of study could indicate whether (and how often) participants use techniques learned during SFA.

Conclusion

The aim of this study was to determine which linguistic variables predicted acquisition, generalization, and maintenance among 28 PWA who completed PMT, 30 PWA who completed SFA, and the combined groups. Input-output phonological processing was the most robust predictor across item types, timepoints, and groups, and it is likely that the skills that underlie the measure of input-output processing, verbal STM and WM, drove this finding. Greater micro-linguistic impairment at baseline was associated with greater acquisition, generalization, and maintenance in the whole and PMT groups, suggesting that greater impairment of fundamental linguistic processing skills is more amenable to treatment in general, and PMT in particular, than

less impairment. Furthermore, the results indicate that SFA may be more appropriate for less impaired PWA. Although more work remains to be done in understanding why some PWA respond to treatment and others do not, this study serves as a stepping stone on the path to maximizing outcomes for every person who seeks treatment for their aphasia.

References

Allen, C. M., Martin, R. C., & Martin, N. (2012). Relations between Short-term Memory Deficits, Semantic Processing, and Executive Function. *Aphasiology, 26*(3-4), 428-461. https://doi.org/10.1080/02687038.2011.617436

Anders, P. L., & Bos, C. S. (1986). Semantic Feature Analysis; An Interactive Strategy for Vocabulary Development and Text Comprehension. *Journal of Reading, 29*(7), 610-616.

Armstrong, E., & Ferguson, A. (2010). Language, meaning, context, and functional communication. *Aphasiology, 24*(4), 480-496. https://doi.org/10.1080/02687030902775157

Babbitt, E. M., Worrall, L., & Cherney, L. R. (2016). Who Benefits from an Intensive Comprehensive Aphasia Program? *Topics in Language Disorders, 36*(2), 168-184. https://doi.org/10.1097/TLD.0000000000000089

Baddeley, A., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of memory and language, 27*(5), 586-595. https://doi.org/10.1016/0749-596X(88)90028-9

Baker, E. (2012). Optimal intervention intensity. *International Journal of Speech-Language Pathology, 14*(5), 401-409. https://doi.org/10.3109/17549507.2012.700323

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Baylor, C., Hula, W., Donovan, N. J., Doyle, P. J., Kendall, D., & Yorkston, K. (2011). An introduction to item response theory and Rasch models for speech-language pathologists. *American journal of speech-language pathology, 20*(3), 243-259. https://doi.org/10.1044/1058-0360(2011/10-0079)

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Best, W., Greenwood, A., Grassly, J., Herbert, R., Hickin, J., & Howard, D. (2013). Aphasia rehabilitation: Does generalisation from anomia therapy occur and is it predictable? A case series study. *Cortex, 49*(9), 2345-2357. https://doi.org/10.1016/j.cortex.2013.01.005

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition, 104*(3), 437-458. https://doi.org/10.1016/j.cognition.2006.07.003

Bono, R., Blanca, M. J., Arnau, J., & Gómez-Benito, J. (2017). Non-normal Distributions Commonly Used in Health, Education, and Social Sciences: A Systematic Review. *Frontiers in psychology, 8*, 1602. https://doi.org/10.3389/fpsyg.2017.01602

Bose, A., Höbler, F., & Saddy, D. (2019). Deciphering the Mechanisms of Phonological Therapy in Jargon Aphasia. *International journal of language & communication disorders, 54*(1), 123. https://doi.org/10.1111/1460-6984.12437

Boyle, M. (2010). Semantic Feature Analysis Treatment for Aphasic Word Retrieval Impairments: What's in a Name? *Topics in Stroke Rehabilitation, 17*(6), 411-422. https://doi.org/10.1310/tsr1706-411

Boyle, M. (2017). Semantic Treatments for Word and Sentence Production Deficits in Aphasia. *Semin. Speech Lang., 38*(1), 52-61. https://doi.org/10.1055/s-0036-1597256

Boyle, M., & Coelho, C. A. (1995). Application of Semantic Feature Analysis as a Treatment for Aphasic Dysnomia. *American Journal of Speech-Language Pathology, 4*(4), 94. https://doi.org/10.1044/1058-0360.0404.94

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. https://doi.org/10.1037/0278-7393.22.6.1482

Brookshire, C. E., Conway, T., Pompon, R. H., Oelke, M., & Kendall, D. L. (2014). Effects of Intensive Phonomotor Treatment on Reading in Eight Individuals With Aphasia and Phonological Alexia. *American Journal of Speech-Language Pathology, 23*(2), 300-311. https://doi.org/10.1044/2014_ajslp-13-0083

Brookshire, R. H., & Nicholas, L. E. (1994). Speech sample size and test-retest stability of connected speech measures for adults with aphasia. *Journal of speech and hearing research, 37*(2), 399.

Butterworth, B., Howard, D., & McLoughlin, P. (1984). The semantic deficit in aphasia: The relationship between semantic errors in auditory comprehension and picture naming. *Neuropsychologia, 22*(4), 409-426. https://doi.org/10.1016/0028-3932(84)90036-8

Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition, 21*(2), 117-175. https://doi.org/10.1016/0010-0277(85)90048-4

Caplan, D., & Hildebrandt, N. (1988). Specific deficits in syntactic comprehension. *Aphasiology, 2*(3-4), 255-258. https://doi.org/10.1080/02687038808248920

Caplan, D., Michaud, J., & Hufford, R. (2013). Dissociations and associations of performance in syntactic comprehension in aphasia and their implications for the nature of aphasic deficits. *Brain and Language, 127*(1), 21-33. https://doi.org/10.1016/j.bandl.2013.07.007

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behav. Brain Sci., 22*(1), 77-94. https://doi.org/10.1017/S0140525X99001788

Carbonell, S., Valdois, S., Charnallet, A., Lyard, G., & Pellat, J. (1991). On the Effect of Selective Attention Deficit in Language Processing: A Single Case Study. *Grazer Linguistische Studien, 35*, 115-128.

Castro, N., Nadeau, S. E., & Kendall, D. L. (Manuscript in review). Phonological path distance influences confrontation naming accuracy in some people with aphasia.

Castro, N., Stella, M., & Siew, C. S. Q. (2020). Quantifying the interplay of semantics and phonology during failures of word retrieval by people with aphasia using a multiplex lexical network. *Cognitive Science, 44*(9). https://doi.org/10.1111/cogs.12881

Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. *Cognitive Science, 26*(5), 609-651. https://doi.org/10.1207/s15516709cog2605_3

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming Syntactic. *Psychological Review, 113*(2), 234-272. https://doi.org/10.1037/0033-295X.113.2.234

Cherney, L. R. (2012). Aphasia treatment: Intensity, dose parameters, and script training. *International Journal of Speech-Language Pathology, 14*(5), 424-431. https://doi.org/10.3109/17549507.2012.686629

Christensen, S. C., Wright, H. H., & Ratiu, I. (2018). Working memory in aphasia: Peeling the onion. *Journal of Neurolinguistics*. https://doi.org/10.1016/j.jneuroling.2018.02.001

Coelho, C. (2005). Direct Attention Training as a Treatment for Reading Impairment in Mild Aphasia. *Aphasiology, 19*(3-5), 275-283.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428. https://doi.org/10.1037/0033-295X.82.6.407

Coltheart, M. (1981). MRC Psycholinguistic Database. Retrieved from http://websites.psychology.uwa.edu.au/school/MRCDatabase/Retrieved from mrc2.html

Cook, R. D. (1977). Detection of influential observations in linear regression *Technometrics, 19*(1), 15-18.

Coran, M., Rodriguez-Fornells, A., Ramos-Escobar, N., Laine, M., & Martin, N. (2020). Word Learning in Aphasia: Treatment Implications and Structural Connectivity Analyses. *Topics in language disorders, 40*(1), 81-109. https://doi.org/10.1097/TLD.0000000000000204

Cowan, N. (1996). Short-term memory, working memory, and their importance in language processing. *Topics in Language Disorders, 17*(1), 1-18.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research, 169*(7), 323-338. https://doi.org/10.1016/S0079-6123(07)00020-9

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior, 11*(6), 671-684. https://doi.org/10.1016/s0022-5371(72)80001-x

Crosson, B., Fabrizio, K. S., Singletary, F., Cato, M. A., Wierenga, C. E., Parkinson, R. B., . . . Rothi, L. J. G. (2007). Treatment of naming in nonfluent aphasia through manipulation of intention and attention: A phase 1 comparison of two novel treatments. *J Inter Neuropsych Soc, 13*(4), 582-594. https://doi.org/10.1017/S1355617707070737

Daneman, M., & Merikle, P. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422-433. https://doi.org/10.3758/BF03214546

Davidson, B., Howe, T., Worrall, L., Hickson, L., & Togher, L. (2008). Social Participation for Older People with Aphasia: The Impact of Communication Disability on Friendships. *Topics in Stroke Rehabilitation, 15*(4), 325-340. https://doi.org/10.1310/tsr1504-325

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review, 93*(3), 283-321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., & Chang, F. (2014). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philos. Trans. R. Soc. B-Biol. Sci., 369*(1634). https://doi.org/10.1098/rstb.2012.0394

Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In (Originally published 2003 ed., Vol. 6, pp. 9-38). Berlin, New York: Berlin, New York: DE GRUYTER MOUTON.

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of Errors of Omission in Aphasic Naming. *Cognitive Neuropsychology, 21*(2-4), 125-145. https://doi.org/10.1080/02643290342000320

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production. *Bulletin Of The Psychonomic Society, 28*(6), 502-502.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42*(1), 287-314. https://doi.org/10.1016/0010-0277(92)90046-K

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*(4), 801-838. https://doi.org/10.1037/0033-295x.104.4.801

Dell, G. S., Schwartz, M. F., Nozari, N., Faseyitan, O., & Coslett, H. B. (2013). Voxel-based lesion-parameter mapping: Identifying the neural correlates of a computational model of

word production. *Cognition, 128*(3), 380-396. https://doi.org/10.1016/j.cognition.2013.05.007

Dickey, M. W., & Yoo, H. (2010). Predicting outcomes for linguistically specific sentence treatment protocols. *Aphasiology, 24*(6-8), 787-801. https://doi.org/10.1080/02687030903515354

Dignam, J., Copland, D., Brien, K., Burfein, P., Khan, A., & Rodriguez, A. D. (2017). Influence of Cognitive Ability on Therapy Outcomes for Anomia in Adults with Chronic Poststroke Aphasia. *Journal of Speech, Language, and Hearing Research, 60*(2), 406-421. https://doi.org/10.1044/2016_JSLHR-L-15-0384

Doyle, P. J., McNeil, M. R., Park, G., Goda, A., Rubenstein, E., Spencer, K., . . . Szwarc, L. (2000). Linguistic validation of four parallel forms of a story retelling procedure. *Aphasiology, 14*(5-6), 537-549. https://doi.org/10.1080/026870300401306

Doyle, P. J., McNeil, M. R., Spencer, K. A., Goda, A. J., Cottrell, K., & Lustig, A. P. (1998). The effects of concurrent picture presentations on retelling of orally presented stories by adults with aphasia. *Aphasiology, 12*(7-8), 561-574. https://doi.org/10.1080/02687039808249558

Druks, J., & Masterson, J. (2002). *An object and action naming battery*. Hove, East Sussex, UK: Psychology Press.

Duffy, J. R. (2013). *Motor speech disorders: substrates, differential diagnosis, and management* (3rd ed. ed.). St. Louis, Mo.: St. Louis, Mo.: Elsevier Mosby.

Døli, H., Helland, T., & Andersen Helland, W. (2017). Self-reported symptoms of anxiety and depression in chronic stroke patients with and without aphasia. *Aphasiology, 31*(12), 1392-1409. https://doi.org/10.1080/02687038.2017.1280595

Echambadi, R., & Hess, J. D. (2007). Mean-Centering Does Not Alleviate Collinearity Problems in Moderated Multiple Regression Models. *Marketing science (Providence, R.I.), 26*(3), 438-445. https://doi.org/10.1287/mksc.1060.0263

Edmonds, L. A. (2016). A Review of Verb Network Strengthening Treatment: Theory, Methods, Results, and Clinical Implications. *Topics in Language Disorders, 36*(2), 123-135. https://doi.org/10.1097/TLD.0000000000000088

Efstratiadou, E., Papathanasiou, I., Holland, R., Archonti, A., & Hilari, K. (2018). A Systematic Review of Semantic Feature Analysis Therapy Studies for Aphasia. *J. Speech Lang. Hear. Res., 61*(5), 1261-1278. https://doi.org/10.1044/2018_JSLHR-L-16-0330

Ellis, C., Peach, R., & Rothermich, K. (2019). *Relative Weight Analysis of the Western Aphasia Battery*. Paper presented at the Clinical Aphasiology Conference, Whitefish, MT.

Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science, 11*(1), 19-23. https://doi.org/10.1111/1467-8721.00160

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach. *Journal of Experimental Psychology: General, 128*(3), 309-331. https://doi.org/10.1037/0096-3445.128.3.309

Eoute, D. (2010). Predictors of aphasia treatment outcome. In J. Fridriksson, H. McDade, A. Montgomery, & S. Shinkareva (Eds.): ProQuest Dissertations Publishing.

Evans, W. S., Cavanaugh, R., Gravier, M. L., Autenreith, A. M., Doyle, P. J., Hula, W. D., & Dickey, M. W. (2020). Effects of Semantic Feature Type, Diversity, and Quantity on Semantic Feature Analysis Treatment Outcomes in Aphasia. *American journal of speech-language pathology*, 1-15. https://doi.org/10.1044/2020_ajslp-19-00112

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience, 15*(2), 399-402. https://doi.org/10.1046/j.0953-816x.2001.01874.x

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G * Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Fergadiotis, G., Swiderski, A., & Hula, W. D. (2019). Predicting confrontation naming item difficulty. *Aphasiology, 33*(6), 689-709. https://doi.org/10.1080/02687038.2018.1495310

Ferrill, M., Love, T., Walenski, M., & Shapiro, L. (2012). The Time-Course of Lexical Activation During Sentence Comprehension in People With Aphasia. *American Journal of Speech - Language Pathology (Online), 21*(2), S179-S189A. https://doi.org/10.1044/1058-0360(2012/11-0109)

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2006). The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation, 16*(2), 129-154. doi.org/10.1080/09602010443000254

Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., . . . Heiberger, R. (2020). Companion to applied regression (Version 3.0-10).

Foygel, D., & Dell, G. S. (2000). Models of Impaired Lexical Access in Speech Production. *Journal of Memory and Language, 43*(2), 182-216. https://doi.org/10.1006/jmla.2000.2716

Francis, D. R., Clark, N., & Humphreys, G. W. (2003). The treatment of an auditory working memory deficit and the implications for sentence comprehension abilities in mild

"receptive" aphasia. *Aphasiology, 17*(8), 723-750. https://doi.org/10.1080/02687030344000201

Francoeur, R. B. (2013). Could Sequential Residual Centering Resolve Low Sensitivity in Moderated Regression? Simulations and Cancer Symptom Clusters. *Open journal of statistics, 3*(6), 24-44. https://doi.org/10.4236/ojs.2013.36A004

Fraundorf, S. (2018). Contrast coding for variables with two categories & "Main effects" vs. "simple main effects" [Lecture notes]. http://www.lrdc.pitt.edu/maplelab/statistics.html

Fromm, D., Holland, A., Armstrong, E., Forbes, M., Macwhinney, B., Risko, A., & Mattison, N. (2011). "Better but no cigar": Persons with aphasia speak about their speech. *Aphasiology, 25*(11), 1431-1447. https://doi.org/10.1080/02687038.2011.608839

Gambi, C., & Pickering, M. J. (2017). Models linking production and comprehension. In E. M. Fernandez & H. S. Cairns (Eds.), *The Handbook of Psycholinguistics* (pp. 157-182). Hoboken, NJ: Wiley-Blackwell.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory & Cognition, 23*(1), 83-94. doi.org/10.3758/BF03210559

Geranmayeh, F., Brownsett, S. L. E., & Wise, R. J. S. (2014). Task-induced brain activity in aphasic stroke patients: what is driving recovery? *Brain, 137*, 2632-2648. https://doi.org/10.1093/brain/awu163

Gervain, J., & Mehler, J. (2010). Speech Perception and Language Acquisition in the First Year of Life. *Annual review of psychology, 61*(1), 191-218. https://doi.org/10.1146/annurev.psych.093008.100408

Gilmore, N., Dwyer, M., & Kiran, S. (2019). Benchmarks of Significant Change After Aphasia Rehabilitation. *Archives of Physical Medicine and Rehabilitation, 100*(6), 1131-1139.e1187. https://doi.org/10.1016/j.apmr.2018.08.177

Gilmore, N., Meier, E. L., Johnson, J. P., & Kiran, S. (2020). Typicality-based semantic treatment for anomia results in multiple levels of generalisation. *Neuropsychological Rehabilitation, 30*(5), 802-828. https://doi.org/10.1080/09602011.2018.1499533

Glen, S. (2015). Akaike's Information Criterion: Definition, formulas. Retrieved from https://www.statisticshowto.com/akaikes-information-criterion/

Goodglass, H., & Kaplan, E. (1972). *Boston Diagnostic Aphasia Examination (BDAE)*. Philadelphia: Lea & Febiger.

Goodglass, H., & Wingfield, A. (1997). Word-finding deficits in aphasia: Brain-behaviour relations and clinical symptomatology. In H. Goodglass & A. Wingfield (Eds.), *Anomia: Neuroanatomical and Cognitive Correlates* (pp. 3-30). San Diego, CA: Academic Press.

Gravier, M., Dickey, M., Hula, W., Evans, W., Owens, R., Winans-Mitrik, R., & Doyle, P. J. (2018). What Matters in Semantic Feature Analysis: Practice-Related Predictors of Treatment Response in Aphasia. *Am. J. Speech-Lang. Pathol., 27*(1), 438-453. https://doi.org/10.1044/2017_AJSLP-16-0196

Grawburg, M., Howe, T., Worrall, L., & Scarinci, N. (2013). Third-party disability in family members of people with aphasia: a systematic review. In (Vol. 35, pp. 1324-1341): Taylor & Francis.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*(2), 267-333. https://doi.org/10.1006/brln.1997.1819

Gupta, P., Martin, N., Abbs, B., Schwartz, M., & Lipinski, J. (2006). New word learning in aphasic patients: Dissociating phonological and semantic components. *Brain and Language, 99*(1-2), 8-9. https://doi.org/10.1016/j.bandl.2006.06.015

Gvion, A., & Friedmann, N. (2012). Does phonological working memory impairment affect sentence comprehension? A study of conduction aphasia. *Aphasiology, 26*(3-4), 494-535. https://doi.org/10.1080/02687038.2011.647893

Hadidi, N., Treat-Jacobson, D. J., & Lindquist, R. (2009). Poststroke depression and functional outcome: A critical review of literature. *Heart & Lung - The Journal of Acute and Critical Care, 38*(2), 151-162. https://doi.org/10.1016/j.hrtlng.2008.05.002

Hambleton, R. K., & Swaminathan, H. (2013). *Item Response Theory: Principles and Applications*. Dordrecht: Dordrecht: Springer Netherlands.

Harnish, S. M., & Lundine, J. P. (2015). Nonverbal Working Memory as a Predictor of Anomia Treatment Success. *Am J Speech Lang Pathol, 24*(4), S880-894. https://doi.org/10.1044/2015_ajslp-14-0153

Harnish, S. M., Schwen Blackett, D., Zezinka, A., Lundine, J. P., & Pan, X. (2018). Influence of working memory on stimulus generalization in anomia treatment: A pilot study. *Journal of Neurolinguistics*. https://doi.org/10.1016/j.jneuroling.2018.02.003

Harris, L., Olson, A., & Humphreys, G. (2014). The link between STM and sentence comprehension: A neuropsychological rehabilitation study. *Neuropsychological Rehabilitation, 24*(5), 678-720. https://doi.org/10.1080/09602011.2014.892885

Hashimoto, N. (2012). The use of semantic- and phonological-based feature approaches to treat naming deficits in aphasia. *Clinical Linguistics & Phonetics, 26*(6), 518-553. https://doi.org/10.3109/02699206.2012.663051

Hashimoto, N., & Frome, A. (2011). The use of a modified semantic features analysis approach in aphasia. *Journal of Communication Disorders, 44*(4), 459-469. https://doi.org/10.1016/j.jcomdis.2011.02.004

Hebb, D. O. (1949). *The organization of behavior : a neuropsychological theory*. New York: New York : Wiley.

Heim, S., Opitz, B., Müller, K., & Friederici, A. D. (2003). Phonological processing during language production: fMRI evidence for a shared production-comprehension network. *Cognitive Brain Research, 16*(2), 285-296. https://doi.org/10.1016/S0926-6410(02)00284-7

Hernández-Sacristán, C., Rosell-Clari, V., Serra-Alegre, E., & Quiles-Climent, J. (2012). On natural metalinguistic abilities in aphasia: A preliminary study. *Aphasiology, 26*(2), 199-219. https://doi.org/10.1080/02687038.2011.630456

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron, 69*(3), 407-422. https://doi.org/10.1016/j.neuron.2011.01.019

Hickok, G., & Poeppel, D. (2004). Dorsal and Ventral Streams: A Framework for Understanding Aspects of the Functional Anatomy of Language. *Cognition, 92*(1-2), 67-62), p.67-99. https://doi.org/10.1016/j.cognition.2003.10.011

Hinckley, J. J., Patterson, J. P., & Carr, T. H. (2001). Differential effects of context- and skill-based treatment approaches: Preliminary findings. *Aphasiology, 15*(5), 463-476. doi.org/10.1080/02687040042000340

Hinton, G. E., & Plaut, D. C. (1987, July 16-18). *Using fast weights to deblur old memories* [Paper presentation]. Proceedings of the 9th Annual Conference of the Cognitive Society. https://cognitivesciencesociety.org/wp-content/uploads/2019/01/cogsci_9.pdf

Hoffman, P., Jefferies, E., Ehsan, S., Jones, R. W., & Lambon Ralph, M. A. (2012). How does linguistic knowledge contribute to short-term memory? Contrasting effects of impaired semantic knowledge and executive control. *Aphasiology, 26*(3-4), 383-403. https://doi.org/10.1080/02687038.2011.581798

Howard, D., Swinburn, K., & Porter, G. (2010). Putting the CAT out: What the Comprehensive Aphasia Test has to offer. *Aphasiology, 24*(1), 56-74. https://doi.org/10.1080/02687030802453202

Hula, W., & McNeil, M. R. (2008). Models of Attention and Dual- Task Performance As Explanatory Constructs in Aphasia. *Seminars in Speech and Language, 29*(3), 169-187. https://doi.org/10.1055/s-0028-1082882

Hunting Pompon, R., Bislick, L., Elliott, K., Madden, E., Minkina, I., Oelke, M., & Kendall, D. L. (2017). Influence of Linguistic and Nonlinguistic Variables on Generalization and Maintenance Following Phonomotor Treatment for Aphasia. *American Journal of Speech - Language Pathology (Online), 26*(4), 1092-1104. https://doi.org/10.1044/2017_AJSLP-16-0175

Hunting-Pompon, R., Kendall, D. L., & Bacon Moore, A. (2011). Examining attention and cognitive processing in participants with self-reported mild anomia. *Aphasiology, 25*(6-7), 800-812. https://doi.org/10.1080/02687038.2010.542562

Ingram, D. (1989). *First language acquisition : method, description, and explanation*. Cambridge [England] ; New York: Cambridge England ; New York : Cambridge University Press.

Jacquemot, C., Dupoux, E., Robotham, L., & Bachoud-Lévi, A.-C. (2012). Specificity in Rehabilitation of Word Production: A Meta-Analysis and a Case Study. *Behavioural Neurology, 25*(2), 73-101. https://doi.org/10.3233/BEN-2012-0358

Jokel, R., Rochon, E., & Leonard, C. (2004). Testing predictions of the interactive activation model in recovery from aphasia after treatment. *Brain and Cognition, 54*(3), 251-253. https://doi.org/10.1016/j.bandc.2004.02.033

Jones, G., & Macken, B. (2018). Long-term associative learning predicts verbal short-term memory performance. *Memory & Cognition, 46*(2), 216-229. https://doi.org/10.3758/s13421-017-0759-3

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, Mass.: Cambridge, Mass: MIT Press.

Just, M. A., & Carpenter, P. A. (1992). A Capacity Theory of Comprehension: Individual Differences in Working Memory. *Psychological Review, 99*(1), 122-149. https://doi.org/10.1037/0033-295X.99.1.122

Kalinyak-Fliszar, M., Kohen, F., & Martin, N. (2011). Remediation of language processing in aphasia: Improving activation and maintenance of linguistic representations in (verbal) short-term memory. *Aphasiology, 25*(10), 1095-1131. https://doi.org/10.1080/02687038.2011.577284

Kaplan, E., Goodglass, H., Segal, O., & Weintraub, S. (2001). Boston naming test. In (Second edition. ed.). Austin, Texas: Austin, Texas Pro-ed.

Kassambara, A. (2020). ggpubr: "ggplot2" based publication ready plots (Version 0.4.0).

Kay, J., Lesser, R., & Coltheart, M. (1992). *PALPA: Psycholinguistic Assessment of Language Processing in Aphasia*. Hove, England: Lawrence Erlbaum Associates.

Kendall, D. L., del Toro, C., Nadeau, S. E., Johnson, J., Rosenbek, J., & Velozo, C. (2010). *The development of a standardized assessment of phonology in aphasia*. Paper presented at the Clinical Aphasiology Conference, Isle of Palms, SC. http://aphasiology.pitt.edu/id/eprint/2112

Kendall, D. L., & Nadeau, S. E. (2016). The Phonomotor Approach to Treating Phonological-Based Language Deficits in People With Aphasia. *Topics in Language Disorders, 36*(2), 109-122. https://doi.org/10.1097/tld.0000000000000085

Kendall, D. L., Oelke, M., Allen, W., Torrence, J., & Nadeau, S. E. (2019). Phonomotor versus Semantic Feature Analysis Treatment for anomia in 58 persons with aphasia: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research*, *62*(12), 4464-4482. https://doi.org/10.1044/2019_JSLHR-L-18-0257

Kendall, D. L., Oelke, M., Brookshire, C. E., & Nadeau, S. E. (2015). The Influence of Phonomotor Treatment on Word Retrieval Abilities in 26 Individuals with Chronic Aphasia: An Open Trial. *Journal of Speech, Language, and Hearing Research, 58*(3), 798-812. https://doi.org/10.1044/2015_JSLHR-L-14-0131

Kendall, D. L., Pompon, R. H., Brookshire, C. E., Minkina, I., & Bislick, L. (2013). An Analysis of Aphasic Naming Errors as an Indicator of Improved Linguistic Processing Following Phonomotor Treatment. *American Journal of Speech-Language Pathology, 22*(2), S240-S249. https://doi.org/10.1044/1058-0360(2012/12-0078)

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 and Language, 105*(1), 1-17. https://doi.org/10.1016/j.bandl.2007.11.007

Kertesz, A. (1982). *The Western Aphasia Battery*. New York: Grune & Stratton.

Kertesz, A. (2007). *The Western Aphasia Battery - Revised*. San Antonio, TX: PsychCorp.

Kiran, S. (2016). How Does Severity of Aphasia Influence Individual Responsiveness to Rehabilitation? Using Big Data to Understand Theories of Aphasia Rehabilitation. *Seminars in speech and language, 37*(1), 48. https://doi.org/10.1055/s-0036-1571358

Kiran, S., & Thompson, C. K. (2003). The Role of Semantic Complexity in Treatment of Naming Deficits: Training Semantic Categories in Fluent Aphasia by Controlling Exemplar Typicality. *Journal of Speech, Language, and Hearing Research, 46*(3), 608-622. https://doi.org/10.1044/1092-4388(2003/048)

Kirk, R. E. (2013). *Experimental design: procedures for the behavioral sciences* (4th edition. ed.). Thousand Oaks: Thousand Oaks: Sage.

Kittredge, A. K., & Dell, G. S. (2016). Learning to Speak by Listening: Transfer of Phonotactics from Perception to Production. *Journal of Memory and Language, 89*, 8-22. https://doi.org/10.1016/j.jml.2015.08.001

Kleim, J. A., & Jones, T. A. (2008). Principles of Experience-Dependent Neural Plasticity: Implications for Rehabilitation after Brain Damage. *Journal of Speech, Language, and Hearing Research, 51*, S225-S239. https://doi.org/10.1044/1092-4388(2008/018)

Koenig-Bruhin, M., & Studer-Eichenberger, F. (2007). Therapy of short-term memory disorders in fluent aphasia: A single case study. *Aphasiology, 21*(5), 448-458. https://doi.org/10.1080/02687030600670593

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(13), 1-26. https://doi.org/https://doi.10.18637/jss.v082.i13

Kleim, J. A., & Jones, T. A. (2008). Principles of Experience-Dependent Neural Plasticity: Implications for Rehabilitation after Brain Damage. *Journal of Speech, Language, and Hearing Research, 51*, S225-S239. https://doi.org/10.1044/1092-4388(2008/018)

Kwak, S. K., & Kim, J. H. (2017). Statistical data preparation: management of missing values and outliers. *Korean journal of anesthesiology, 70*(4), 407-411. https://doi.org/10.4097/kjae.2017.70.4.407

Lambon Ralph, M. A., Snell, C., Fillingham, J. K., 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. *Neuropsychological Rehabilitation, 20*(2), 289-305. https://doi.org/10.1080/09602010903237875

Lichtheim, L. (1885). On aphasia. *Brain*, 433-484.

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software, 3*(26), 772. https://doi.org/https://doi.10.21105/joss.00772

Lüdecke, D. (2020). sjPlot: Data visualization for statistics in social science (Version R package version 2.8.5).

Macdonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24*(1), 56-98. https://doi.org/10.1016/0010-0285(92)90003-K

MacWhinney, B. (2000). *The CHILDES project : tools for analyzing talk* (3rd ed. ed.). Mahwah, NJ: Mahwah, NJ : Lawrence Erlbaum.

Madden, E. B., Conway, T., Henry, M. L., Spencer, K. A., Yorkston, K. M., & Kendall, D. L. (2018). The Relationship Between Non-Orthographic Language Abilities and Reading Performance in Chronic Aphasia: An Exploration of the Primary Systems Hypothesis. *Journal of speech, language, and hearing research, 61*(12), 3038-3054. https://doi.org/10.1044/2018_jslhr-l-18-0058

Majerus, S. (2018). Working memory treatment in aphasia: A theoretical and quantitative review. *Journal of Neurolinguistics, 48*, 157-175. https://doi.org/10.1016/j.jneuroling.2017.12.001

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*(1-2), 71-102. https://doi.org/10.1016/0010-0277(87)90005-9

Martin, N., & Ayala, J. (2004). Measurements of auditory-verbal STM span in aphasia: Effects of item, task, and lexical impairment. *Brain and Language, 89*(3), 464-483. https://doi.org/10.1016/j.bandl.2003.12.004

Martin, N., & Gupta, P. (2004). Exploring the relationship between word processing and verbal short-term memory: Evidence from associations and dissociations. *Cognitive Neuropsychology, 21*(2-4), 213-228. https://doi.org/10.1080/02643290342000447

Martin, N., Minkina, I., Kohen, F. P., & Kalinyak-Fliszar, M. (2018). Assessment of linguistic and verbal short-term memory components of language abilities in aphasia. *Journal of Neurolinguistics*. https://doi.org/10.1016/j.jneuroling.2018.02.006

Martin, N., & Saffran, E. M. (1992). A computational account of deep dyslexia – Evidence from a single case study. *Brain and Language, 43*(2), 240-274. https://doi.org/10.1016/0093-934x(92)90130-7

Martin, N., & Saffran, E. M. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology, 14*(5), 641-682.

Martin, N., & Saffran, E. M. (2002). The relationship of input and output phonological processing: An evaluation of models and evidence to support them. *Aphasiology, 16*(1-2), 107-150. https://doi.org/10.1080/02687040143000447

Martin, N., Schlesinger, J., Obermeyer, J., Minkina, I., & Rosenberg, S. (2020). Treatment of verbal short-term memory abilities to improve language function in aphasia: A case series treatment study. *Neuropsychological Rehabilitation, ahead-of-print*(ahead-of-print), 1-42. https://doi.org/10.1080/09602011.2020.1731554

Martin, N., Schwartz, M. F., & Kohen, F. P. (2006). Assessment of the ability to process semantic and phonological aspects of words in aphasia: A multi-measurement approach. *Aphasiology, 20*(2-4), 154-166. doi.org/10.1080/02687030500472520

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of Input and Output Phonology in Word Processing and Short-Term Memory. *Journal of Memory and Language, 41*(1), 3-29. https://doi.org/10.1006/jmla.1999.2637

McBride-Chang, C. (1995). What Is Phonological Awareness? *Journal of Educational Psychology, 87*(2), 179-192. https://doi.org/10.1037/0022-0663.87.2.179

McDonald, J. H. (2014). *Handbook of biological statistics* (3rd ed.). Baltimore, MD: Sparky House Publishing.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica, 22*(3), 276. https://doi.org/10.11613/BM.2012.031

McNeil, M. R. (1983). Aphasia: Neurological Considerations. *Topics in Language Disorders, 3*(4), 1-19.

McNeil, M. R., Doyle, P. J., Fossett, T. R. D., Park, G. H., & Goda, A. J. (2001). Reliability and concurrent validity of the information unit scoring metric for the story retelling procedure. *Aphasiology, 15*(10-11), 991-1006. https://doi.org/10.1080/02687040143000348

McNeil, M. R., Odell, K., & Tseng, C. H. (1991). Toward the integration of resource allocation into a general theory of aphasia. In T. Prescott (Ed.), *Clinical Aphasiology: Vol. 20* (pp. 21-39). Austin, TX: Pro-Ed Publishers.

Menyuk, P., Menu, L., & Silber, R. (1986). Early Strategies for the Perception and Production of Words and Sounds. In P. Fletcher & M. Garman (Eds.), *Language Acquisition* (2nd ed.), pp. 198-222. Cambridge Unuiversity Press. https://doi.org/10.1017/CBO9780511620683.012

Minkina, I., Oelke, M., Bislick, L. P., Brookshire, C. E., Hunting Pompon, R., Silkes, J. P., & Kendall, D. L. (2016). An investigation of aphasic naming error evolution following phonomotor treatment. *Aphasiology*, 1-19. https://doi.org/10.1080/02687038.2015.1081139

Minkina, I., Rosenberg, S., Kalinyak-Fliszar, M., & Martin, N. (2017). Short-Term Memory and Aphasia: From Theory to Treatment. *Seminars in Speech and Language, 38*(1), 17-28. doi.org/10.1055/s-0036-1597261

Minkina, I., Salis, C., & Martin, N. (2018). Short-term and working memory deficits in aphasia: Current issues in theory, evidence, and treatment. *Journal of Neurolinguistics, 48*, 1-3. https://doi.org/10.1016/j.jneuroling.2018.07.001

Minkina, I., Silkes, J. P., Bislick, L., Madden, E. B., Lai, V., Pompon, R. H., . . . Kendall, D. L. (2019). The Influence of Phonomotor Treatment on Word Retrieval: Insights From Naming Errors. *Journal of speech, language, and hearing research : JSLHR, 62*(11), 4080. https://doi.org/10.1044/2019_JSLHR-L-19-0014

Miyake, A., & Shah, P. (1999). *Models of working memory: mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Morris, R., Eccles, A., Ryan, B., & Kneebone, I. I. (2017). Prevalence of anxiety in people with aphasia after stroke. *Aphasiology, 31*(12), 1410-1415. https://doi.org/10.1080/02687038.2017.1304633

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

Murray, L., Salis, C., Martin, N., & Dralle, J. (2018). The use of standardised short-term and working memory tests in aphasia research: a systematic review. *Neuropsychological Rehabilitation, 28*(3), 309-351. https://doi.org/10.1080/09602011.2016.1174718

Nadeau, S. E. (2001). Phonology: A review and proposals from a connectionist perspective. *Brain and Language, 79*(3), 511-579.

Nakano, H., Saron, C., & Swaab, T. Y. (2010). Speech and Span: Working Memory Capacity Impacts the Use of Animacy but Not of World Knowledge during Spoken Sentence Comprehension. *Journal of Cognitive Neuroscience, 22*(12), 2886-2898. https://doi.org/10.1162/jocn.2009.21400

National Institute of Neurological Diseases and Stroke. (2011). *NIH stroke scale*. Bethesda, MD: Bethesda, MD: National Institute of Neurological Disorders and Stroke, Dept. of Health and Human Services, USA.

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology, 16*(10-11), 935-979. https://doi.org/10.1080/02687030244000563

Nicholas, M., Sinotte, M., & Helm-Estabrooks, N. (2005). Using a computer to communicate: Effect of executive function impairments in people with severe aphasia. *Aphasiology, 19*(10-11), 1052-1065. doi.org/10.1080/02687030544000245

Nozari, N., & Dell, G. S. (2013). How damaged brains repeat words: A computational approach. *Brain and Language, 126*(3), 327-337. https://doi.org/10.1016/j.bandl.2013.07.005

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology, 63*(1), 1-33. https://doi.org/10.1016/j.cogpsych.2011.05.001

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. https://doi.org/10.1016/j.jml.2010.08.001

Patterson, J., & Coppens, P. (2018). How is generalization managed in anomia research? An analysis of current practice and future directions. *Frontiers in human neuroscience, 12*. https://doi.org/10.3389/conf.fnhum.2018.228.00058

Peach, R., Nathan, M., & Beck, K. (2017). Language-Specific Attention Treatment for Aphasia: Description and Preliminary Findings. *Seminars in Speech and Language, 38*(1), 5.

Petitclerc, C. (2004). Normality: the unreachable star? *Clinical chemistry and laboratory medicine, 42*(7), 698-701. https://doi.org/10.1515/cclm.2004.119

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and brain sciences, 36*(4), 329-347. https://doi.org/10.1017/S0140525X12001495

Plaut, D. C. (1996). Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language, 52*(1), 25-82. https://doi.org/10.1006/brln.1996.0004

Prabhakaran, S. (2016). InformationValue: Performance analysis and companion functions for binary classification models (Version 1.2.3).

Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences - PNAS, 103*(20), 7865-7870. https://doi.org/10.1073/pnas.0509989103

Purdy, M. (2002). Executive function ability in persons with aphasia. *Aphasiology, 16*(4-6), 549-557. https://doi.org/10.1080/02687030244000176

Quene, H., & van den Bergh, H. (2008). Examples of Mixed-Effects Modeling with Crossed Random Effects and with Binomial Data. *Journal of Memory and Language, 59*(4), 413-425. https://doi.org/10.1016/j.jml.2008.02.002

Quique, Y. M., Evans, W. S., & Dickey, M. W. (2018). Acquisition and Generalization Responses in Aphasia Naming Treatment: A Meta-Analysis of Semantic Feature Analysis Outcomes. *American journal of speech-language pathology*, 1. https://doi.org/10.1044/2018_AJSLP-17-0155

Raven, J. C., Court, J. H., & Raven, J. E. (1998). Raven's coloured progressive matrices. In *Coloured progressive matrices*. San Antonio, Tex.: San Antonio, Tex. : Harcourt Assessment.

Raymer, A. M., Beeson, P., Holland, A., Kendall, D. L., Maher, L. M., Martin, N., . . . Gonzalez Rothi, L. J. (2008). Translational Research in Aphasia: From Neuroscience to Neurorehabilitation. *Journal of Speech, Language, and Hearing Research, 51*(1), S259-S275. doi.org/10.1044/1092-4388(2008/020)

Raymer, A. M., Thompson, C. K., Jacobs, B., & Legrand, H. R. (1993). Phonological treatment of naming deficits in aphasia: Model-based generalization analysis. *Aphasiology, 7*(1), 27-53. https://doi.org/10.1080/02687039308249498

Reilly, J., & Kean, J. (2007). Formal Distinctiveness of High- and Low-Imageability Nouns: Analyses and Theoretical Implications. *Cognitive Science, 31*(1), 157-168. https://doi.org/10.1080/03640210709336988

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.

Robey, R. R. (1994). The Efficacy of Treatment for Aphasic Persons: A Meta-Analysis. *Brain and Language, 47*(4), 582-608. https://doi.org/10.1006/brln.1994.1060

Robey, R. R. (1998). A Meta-Analysis of Clinical Outcomes in the Treatment of Aphasia. *Journal of Speech, Language, and Hearing Research, 41*(1), 172-187. https://doi.org/10.1044/jslhr.4101.172

Salis, C. (2012). Short-term memory treatment: patterns of learning and generalisation to sentence comprehension in a person with aphasia. *Neuropsychol Rehabil, 22*(3), 428-448. https://doi.org/10.1080/09602011.2012.656460

Salis, C., Kelly, H., & Code, C. (2015). Assessment and treatment of short-term and working memory impairments in stroke aphasia: a practical tutorial. *International Journal of*

Language & Communication Disorders, 50(6), 721-736. https://doi.org/10.1111/1460-6984.12172

Schatschneider, C., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Mehta, P. (1999). The Dimensionality of Phonological Awareness: An Application of Item Response Theory. *Journal of educational psychology, 91*(3), 439-449. https://doi.org/10.1037/0022-0663.91.3.439

Scherr, M., Kunz, A., Doll, A., Mutzenbach, J. S., Broussalis, E., Bergmann, H. J., . . . Killer-Oberpfalzer, M. (2016). Ignoring floor and ceiling effects may underestimate the effect of carotid artery stenting on cognitive performance. *Journal of NeuroInterventional Surgery, 8*(7), 747-751. https://doi.org/10.1136/neurintsurg-2014-011612

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of memory and language, 29*(1), 86-102. https://doi.org/10.1016/0749-596x(90)90011-n

Schuell, H. (1960). Clinical findings in aphasia. *The Journal-lancet, 80*, 482.

Segaert, K. R., Menenti, L. M. E., Weber, K. M., Petersson, K. M., & Hagoort, P. (2012). Shared Syntax in Language Production and Language Comprehension-An fMRI Study. *Cerebral cortex (New York, N.Y. 1991), 22*, 1662-1670.

Shehata, G. A., El Mistikawi, T., Risha, A. S. K., & Hassan, H. S. (2015). The effect of aphasia upon personality traits, depression and anxiety among stroke patients. *Journal of Affective Disorders, 172*, 312-314. https://doi.org/10.1016/j.jad.2014.10.027

Silkes, J. P., Fergadiotis, G., Graue, K., & Kendall, D. L. (2020). Effects of Phonomotor Therapy and Semantic Feature Analysis on Discourse Production. *American journal of speech-language pathology*, 1-14. https://doi.org/10.1044/2020_AJSLP-19-00111

Silkes, J. P., Fergadiotis, G., Hunting Pompon, R., Torrence, J., & Kendall, D. L. (2019). Effects of phonomotor treatment on discourse production. *Aphasiology, 33*(2), 125-139. https://doi.org/10.1080/02687038.2018.1512080

Šimkovic, M., & Träuble, B. (2019). Robustness of statistical methods when measure is affected by ceiling and/or floor effect. *PloS one, 14*(8), e0220889.

Sinyor, D., Amato, P., Kaloupek, D. G., Becker, R., Goldenberg, M., & Coopersmith, H. (1986). Post-stroke depression: relationships to functional impairment, coping strategies, and rehabilitation outcome. *Stroke, 17*(6), 1102. https://doi.org/10.1161/01.STR.17.6.1102

Stahl, S. A., & Murray, B. A. (1994). Defining Phonological Awareness and Its Relationship to Early Reading. *Journal of educational psychology, 86*(2), 221-234. https://doi.org/10.1037/0022-0663.86.2.221

Staines, G. L. (2008). The Causal Generalization Paradox: The Case of Treatment Outcome Research. *Review of General Psychology, 12*(3), 236-252. https://doi.org/10.1037/1089-2680.12.3.236

Storkel, H. L., Armbrüster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of speech, language, and hearing research: JSLHR, 49*(6), 1175. https://doi.org/10.1044/1092-4388(2006/085)

Sung, J. E., McNeil, M. R., Pratt, S. R., Dickey, M. W., Hula, W. D., Szuminsky, N. J., & Doyle, P. J. (2009). Verbal working memory and its relationship to sentence-level reading and listening comprehension in persons with aphasia. *Aphasiology, 23*(7-8), 1040-1052. https://doi.org/10.1080/02687030802592884

Swinburn, K., Porter, G., & Howard, D. (2004). *CAT: Comprehensive Aphasia Test*. New York: Psychology Press.

Sze, W. P., Hameau, S., Warren, J., & Best, W. (2020). Identifying the components of a successful spoken naming therapy: A meta-analysis of word-finding interventions for adults with aphasia. *Aphasiology*. Advance online pubication. https://doi.org/10.1080/02687038.2020.1781419

Threats, T. (2010). The ICF Framework and Third Party Disability: Application to the Spouses of Persons With Aphasia. *Topics in Stroke Rehabilitation, 17*(6), 451-457. https://doi.org/10.1310/tsr1706-451

Thompson, C. K., Shapiro, L. P., Kiran, S., & Sobecks, J. (2003). The Role of Syntactic Complexity in Treatment of Sentence Deficits in Agrammatic Aphasia: The Complexity Account of Treatment Efficacy (CATE). *Journal of Speech, Language, and Hearing Research, 46*(3), 591-607. https://doi.org/10.1044/1092-4388(2003/047)

Tseng, C.-H., McNeil, M. R., & Milenkovic, P. (1993). An Investigation of Attention Alloction Deficits in Aphasia. *Brain and Language, 45*(2), 276-296.

Vaden, K. I., Halpin, H. R., & Hickok, G. S. (2009). The Irvine Phonotactic Online Dictionary. In: http://www.iphod.com.

Villard, S., & Kiran, S. (2017). To what extent does attention underlie language in aphasia? In (Vol. 31, pp. 1226-1245). Abingdon: Taylor & Francis Ltd.

Vitevitch, M. S., & Luce, P. A. (1998). When Words Compete: Levels of Processing in Perception of Spoken Words. *Psychological science, 9*(4), 325-329. https://doi.org/10.1111/1467-9280.00064

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic Phonotactics and Neighborhood Activation in Spoken Word Recognition. *Journal of memory and language, 40*(3), 374-408. https://doi.org/10.1006/jmla.1998.2618

Vitevitch, M. S., & Luce, P. A. (2016). Phonological Neighborhood Effects in Spoken Word Perception and Production. *Annual review of linguistics, 2*(1), 75-94. https://doi.org/10.1146/annurev-linguistics-030514-124832

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & cognition, 31*(4), 491-504. https://doi.org/10.3758/bf03196091

Vitevitch, M. S., & Storkel, H. L. (2012). Examining the Acquisition of Phonological Word Forms with Computational Experiments. *Language and speech, 56*(4), 493-527. https://doi.org/10.1177/0023830912460513

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of Reading-Related Phonological Processing Abilities: New Evidence of Bidirectional Causality From a Latent Variable Longitudinal Study. *Developmental psychology, 30*(1), 73-87. https://doi.org/10.1037/0012-1649.30.1.73

Walker, G. M., & Hickok, G. (2015). Bridging computational approaches to speech production: The semantic–lexical–auditory–motor model (SLAM). *Psychonomic bulletin & review, 23*(2), 339-352. https://doi.org/10.3758/s13423-015-0903-7

Walker, G. M., Hickok, G., & Fridriksson, J. (2018). A cognitive psychometric model for assessment of picture naming abilities in aphasia. *Psychological assessment, 30*(6), 809-826. https://doi.org/10.1037/pas0000529

Wallace, S. J., Worrall, L., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., . . . Gauvreau, C. A. (2017). Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF. *Disability and Rehabilitation, 39*(14), 1364-1379. https://doi.org/10.1080/09638288.2016.1194899

Wambaugh, J. L. (2003). A comparison of the relative effects of phonologic and semantic cueing treatments. *Aphasiology, 17*(5), 433-441. doi.org/10.1080/02687030344000085

Wambaugh, J. L., Linebaugh, C. W., Doyle, P. J., & Martinez, A. L. (2001). Effects of two cueing treatments on lexical retrieval in aphasic speakers with different levels of deficit. *Aphasiology, 15*(10-11), 933-950. doi.org/pdf+

Wambaugh, J. L., Mauszycki, S., Cameron, R., Wright, S., & Nessler, C. (2013). Semantic Feature Analysis: Incorporating Typicality Treatment and Mediating Strategy Training to Promote Generalization. *American Journal of Speech-Language Pathology, 22*(2), Language Pathology, 2013, Vol.2022(2012). https://doi.org/10.1044/1058-0360(2013/12-0070)

Warker, J. A., Xu, Y., Dell, G. S., & Fisher, C. (2009). Speech errors reflect the phonotactic constraints in recently spoken syllables, but not in recently heard syllables. *Cognition, 112*(1), 81-96. https://doi.org/10.1016/j.cognition.2009.03.009

Wechsler, D. (2009). *Wechsler Memory Scale - 4th Edition*. San Antonio, TX: Pearson.

Wei, T., & Simko, V. (2017). R package "corrplot": Visualization of a correlation matrix (Version 0.84).

Weisberg, S., & Fox, J. (2019). *An R companion to applied regression* (Third edition. ed.). Thousand Oaks, California: Thousand Oaks, California : SAGE Publications, Inc.

Wiley, R. W., & Rapp, B. (2018). Statistical analysis in Small-N Designs: using linear mixed-effects modeling for evaluating intervention effectiveness. *Aphasiology, 33*(1), 1-30. https://doi.org/10.1080/02687038.2018.1454884

Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology, 23*(11), 1338-1352. https://doi.org/10.1080/02687030902732745

Yopp, H. K. (1988). The Validity and Reliability of Phonemic Awareness Tests. *Reading research quarterly, 23*(2), 159-177. https://doi.org/10.2307/747800

Zimmerman, R. M., Silkes, J. P., Kendall, D. L., & Minkina, I. (2019). The link between verbal short-term memory and anomia treatment gains. *American Journal of Speech-Language Pathology*, *28*(3), 1039-1052. https://doi.org/https://doi.org/10.1044/2019_AJSLP-18-0176

Appendix A: PMT Stimuli

Table 6

PMT Stimuli

Sounds in Isolation		**Real Words**				**Nonwords**			
Trained		**Trained**		**Untrained**		**Trained**		**Untrained**	
IPA	Graphemes	1 syllable	2 syllable	1 syllable	2 syllable	1 syllable	2 syllable	1 syllable	2 syllable
p	p	ape	clover	beef	baby	doi (dɔɪ)	chootee (ʧuti)	ain (eɪn)	wurkee (wɝki)
b	b	bird	diver	boot	iron	af (æf)	zhuree (ʒɝi)	poom (pum)	koetoe (koʊtoʊ)
f	f	bride	father	bow	jury	toos (tus)	foekoe (foʊkoʊ)	gee (gi)	wayzer (weɪzɚ)
v	v	bruise	genie	eel	ladder	sheev (ʃiv)	leber (lɛbɚ)	haje (heɪdʒ)	rooit (ruɪt)
t	t	cave	gravy	fig	lasso	ek (ɛk)	doem (doʊʌm)	loy (lɔɪ)	sayvay (seɪveɪ)
d	d	ditch	halo	hay	leather	dach (dætʃ)	mefoe (mɛfoʊ)	heeg (hig)	fooer (fuɚ)
k	k	fire	heater	thigh	razor	peenz (pinz)	shever (ʃɛvɚ)	jong (dʒɑŋ)	laybee (leɪbi)
g	g	fur	ivy	tire	shadow	poa (poʊə)	feether (fiðɚ)	poy (pɔɪ)	grayzee (greɪzi)
θ	th	jail	jockey	toy	turkey	meeth (miθ)	toiler (tɔɪlɚ)	awb (ɑb)	ekee (ɛki)
ð	th	jeans	level	whip	valet	ri (rɪ)	izel (ɑɪzl)	jeef (dʒif)	badow (bædoʊ)
s	s	knee	meadow	wire		ish (ɪʃ)	shaybee (ʃeɪbi)	tay (teɪ)	nider (nɑɪdɚ)
z	z	knob	movie			whup (wʌp)	veeder (vidɚ)	mirth (mɝθ)	eepoe (ipoʊ)
ʃ	sh	knot	polo			breek (brik)	zower (zaʊɚ)	vank (vænk)	vaylow (veɪloʊ)
ʒ	zh	maze	ranger			voo (vu)	tawthee (tɑθi)	bap (bæp)	sheefur (ʃifɚ)
ʧ	ch	mop	shoulder			eep (ip)	jiver (dʒɪvɚ)	ka (kæ)	hoower (huwɚ)
ʤ	j	owl	shower			reesh (riʃ)	wooter (wutɚ)	ool (ul)	eeshur (iʃɚ)
l	l	pie	speaker			nie (nɑɪ)	dungee (dʌŋi)	wog (wɑg)	rayger (reɪgɚ)
r	r	plane	teacher			iej (aɪdʒ)	turmee (tɝmi)	glane (gleɪn)	zopper (zɑpɚ)
h	h	wheel	tiger			zine (zɑɪn)	lekzher (lɛkʒɚ)	ieg (ɑɪg)	joah (dʒoʊə)
w	w	witch				broiz (brɔɪz)	lekee (lɛki)	dite (dɑɪt)	tawkee (tɑki)
wh	wh					thag (θæg)	juroe (dʒɝo)	grabe (greɪb)	zire (zɑɪər)
m	m					oit (ɔɪt)	shashoe (ʃæsoʊ)	jie (dʒɑɪ)	thiver (θɪvɚ)
n	n					kur (kɝ)	hoyter (hoɪtɚ)	wawj (wɑdʒ)	wiver (wɑɪvɚ)
ŋ	ng					froos (frus)	neenee (nini)	fie (fɑɪ)	uzher (ʌʒɚ)
i	ee					grake (greɪk)	rayzel (reɪzl)	oozh (uʒ)	chafter (ʧæftɚ)
ɪ	i					choy (ʧɔɪ)	highger (hɑɪgɚ)	whike (wɑɪk)	osay (oseɪ)
ɛ	e					oos (us)	woewuh (woʊwə)	gride (grɑɪd)	doojee (dudʒi)
eɪ	ae					wap (wæp)	unger (ʌngɚ)	loich (lɔɪʧ)	fayshur (feɪʃɚ)
æ	a					faps (fæps)	miver (mɑɪvɚ)	moy (mɔɪ)	shiloe (ʃɪlo)
ʌ, ə	u					woy (wɔɪ)	jawvee (dʒɑvi)	jurl (dʒɝl)	voker (voʊkɚ)
ɑ, ɔ	o, aw					awch (ɑʧ)	prezhur (prɛʒɚ)	thed (θɛd)	haybee (heɪbi)
o, oʊ	oe					plown (plɑʊn)	foover (fuvɚ)	eem (im)	rieger (rɑɪgɚ)
ʊ	oo					zae (zeɪ)	pire (pɑɪɚ)	riz (rɪz)	layfee (leɪfi)
u	oo					hob (hɑb)	dryper (drɑɪpɚ)		meevee (mivi)

Sounds in Isolation		**Real Words**				**Nonwords**			
Trained		**Trained**		**Untrained**		**Trained**		**Untrained**	
IPA	Graphemes	1 syllable	2 syllable	1 syllable	2 syllable	1 syllable	2 syllable	1 syllable	2 syllable
ɑɪ ju ɔɪ aʊ ɝ, ɚ ɔr ɑr	ie ue oi, oy ow, ou er, ir, ur or ar					veed (vid)	gower (gaʊɚ) teever (tivɚ) ibee (ɑɪbi)		tycher (tɑɪʧɚ) kloper (kloʊpɚ) nyer (naɪɚ) langee (leɪŋi) gainjer (geɪmdʒɚ) skonner (skɑnɚ)

Appendix B: PMT Protocol

Table 7

PMT Protocol

Treatment materials	• Small mirror • Line drawings of mouth postures, icons for voiced/voiceless consonants • Letter tiles • Wipe-off board with markers • Small colored blocks	
	Sounds in Isolation	**Sounds in Syllables**
Overview	The purpose of Stage 1 is to train sounds in *isolation* through multi-modal instruction using tasks designed to engage distributed articulatory-motor, acoustic, tactile-kinesthetic, and orthographic representations. **Consonant sounds** are introduced using mouth pictures and SLP model as cognate pairs by place/manner of articulation and grouped according to tactile-kinesthetic description (lip, tongue, air, nasal, wind). They are introduced in the following order: lip (*p/b, f/v*), tongue (*t/d, k/g, th/th*), air (*s/z, sh/zh, ch/j*), tongue (*l/r*), nasal (*m/n/ng*) and wind (*h/w/wh*). When mastery of a consonant pair is achieved (e.g. *p/b*) in perception and production (approximately 85% accuracy), the next sound pair is introduced (e.g. *t/d*). Once a sound pair is introduced, training continues on this pair in all subsequent sessions. Once a participant can perceive and produce all consonants in isolation, corresponding graphemes are introduced using the corresponding mouth picture. **Vowel sounds** are trained according to lip and jaw placement via mouth pictures and letter tiles. Vowel sounds (*ee, o, oo*) are introduced with consonants to allow for minimal pair discrimination (e.g., *eep, op, oop*). The remaining vowels are trained after consonants.	The purpose of Stage 2 is to extend skills acquired in Stage 1 to *phoneme sequences*. Treatment tasks remain similar to Stage 1 tasks, with the exception that sounds will be produced in combinations rather than isolation. Training progresses from shorter, monosyllabic sequences to longer, multisyllabic (more complex) sequences (e.g., VC, CV, CVC, CCV, VCC, CCVC, CVCC, CCVCC, CVCV). Both real and nonwords are trained using phonologic tasks (in other words, only phonological features, *not* semantic features, are trained for real words). Nonword training is introduced before real word training to allow for emphasis on phonology; however, as treatment progresses nonwords and real words are trained simultaneously.

	Sounds in Isolation	Sounds in Syllables
Introduction of sounds and sound sequences	Participant observes SLP producing a single sound (e.g. /*p*/). SLP asks participant what they observed (heard, saw) and if needed, describes what articulators are moving and how they move. For the sound /*p*/, for example, "the lips come together and blow apart, the sound is 'quiet', so the voice is turned off, the tongue is not moving." The participant is then shown the line drawing of the mouth posture corresponding to the sound. After looking at the mouth picture and hearing the SLP's production, the participant is then asked to repeat the sound while looking in the mirror. The participant is also asked to place their hand on their throat in order to feel for vocal fold vibration ("quiet" versus "noisy"). Following production, the SLP asks the participant what s/he saw and felt when the sound was made. Socratic questioning is used to enable the participant to "discover" the auditory, visual, articulatory, and tactile/kinesthetic attributes of the sound (e.g., "What do you feel when you make that sound? What moved? What did you see when you made that sound?" etc.). Within therapy progression for all levels is based on 85% accurate performance on task.	The process of "discovering" sounds primarily occurs in Stage 1; however, knowledge of the auditory, visual, articulatory and tactile/kinesthetic attributes of sounds can also be used later in the program as a cueing technique to identify individual phonemes within a phoneme sequence. For example, if a participant had trouble parsing the initial sound in *peef*, the SLP would use Socratic questioning (e.g., "What do you feel when you make that first sound? What moved? Did your lips or tongue move when you made that sound?" etc.) to help identify the initial sound /*p*/. Put differently, rather than give the participant a model and tell them what the initial sound is, the SLP assists the participant in self-awareness of errors and how to repair them.
Perception tasks	Perception of sounds in isolation can be trained through various multi-modal tasks. Examples: • **Mouth pictures:** SLP produces a sound (e.g., *p*) and asks the participant to choose that sound from an array of mouth pictures (e.g., *p, b, t, d*) • **Colored blocks:** SLP produces a string of individual sounds (e.g., *p, t, t, b*) and asks the participant to lay out blocks to demonstrate ability to discriminate sounds (e.g., blocks: red, blue, blue, green). • **Verbal:** SLP produces two sounds (e.g., *p, p* or *p, b*) and asks the participant "same or different." • **Letters:** SLP produces a sound and asks participant to point to the corresponding letter from an array of letters.	The SLP produces a real or nonword sound combination and asks the participant to depict the target through various tasks: • **Mouth pictures:** If the participant heard the CVC *peef*, they would select the pictures corresponding to *p, ee,* and *f*. • **Colored blocks:** If the participant heard the CVCV *peefee*, they would select three differently colored blocks arranged in the following order: white, black, red, black. • **Verbal:** If the participant heard the CCVCs *grook* and *glook*, the SLP would ask "same or different." • **Letters:** If the participant heard *chootee*, s/he would select the corresponding letter tiles.

	Sounds in Isolation	Sounds in Syllables
Production tasks	Production of sounds in isolation can be trained through various tasks. Here are some examples: • **Mouth pictures:** The SLP shows participant a mouth picture and asks the participant to produce that sound (e.g., *d*). • **Motor description:** The SLP describes a sound (e.g., "make the sound where your voice is noisy, and your tongue quickly taps the roof of your mouth") and asks the participant to say the sound. • **Verbal:** The SLP asks the participant to repeat a sound *p* or a string of individual sounds *p, p, s, d.* • **Letters:** The SLP shows the participant a letter to elicit production of the sound.	The SLP elicits a real or nonword sound combination by asking the participant to produce the target through various tasks: • **Mouth pictures:** The SLP lays out a series of mouth pictures and asks the participant to "touch and say" each sound (*f-ee-p*) and then blend the sounds to produce the target (*feep*). • **Verbal:** The SLP asks the participant to repeat a nonword *grook* and parse the word apart (*g-r-oo-k*). • **Letters:** The SLP lays out letter tiles (or writes letters on dry erase board). The participant parses out the sounds by underlining and verbalizing each grapheme and then blends the sounds to produce the target.

Appendix C: SFA Stimuli

Table 8

SFA Stimuli

High Frequency Stimuli								
Category	**Body Parts**	**Clothing and Accessories**	**Food and Beverages**	**Household**	**Hobbies, Recreation/Sports**	**Nature**	**Occupations**	**Transportation**
Targets	lungs	buttons	pudding	refrigerator	wrestling	rainbow	photographer	carriage
	chin	necklace	cereal	oven	bowling	hurricane	farmer	limo
	thumb	skirt	potatoes	bench	swimming	clouds	dentist	motorcycle
	toes	pockets	corn	rug	hockey	lawn	policeman	jeep
	palm	leather	rice	pillow	soccer	waves	chef	subway
	bone	tie	apple	plate	fishing	snow	pilot	elevator
	tongue	belt	soup	closet	painting	desert	priest	tank
	lips	shoe	juice	gate	baseball	coast	actor	bike
	ear	glasses	bread	roof	golf	trees	artist	taxi
	fingers	jacket	pizza	bedroom	football	mountain	nurse	traffic
Low Frequency Stimuli								
Category	**Body Parts**	**Clothing and Accessories**	**Food and Beverages**	**Household**	**Hobbies, Recreation/Sports**	**Nature**	**Occupations**	**Transportation**
Targets	trachea	bowtie	avocado	armoire	rafting	spiderweb	umpire	tugboat
	bellybutton	cardigan	pineapple	bookshelf	origami	geyser	veterinarian	rickshaw
	toenail	kilt	cinnamon	quilt	archery	sunflower	miner	boxcar
	pinkie	beret	macaroni	mixer	croquet	beehive	pianist	blimp
	calves	mittens	lime	silverware	fencing	petal	ballerina	rowboat
	heel	scarf	oatmeal	mattress	skiing	volcano	mechanic	canoe
	elbow	slippers	gravy	stove	photography	avalanche	musician	skates
	cheek	vest	lemonade	candle	yoga	autumn	rabbi	tractor
	forehead	sleeve	tomato	cabinet	camping	seeds	nun	ferry
	ankle	jeans	garlic	ceiling	chess	pond	magician	submarine

Appendix D: SFA Protocol

Table 9

SFA Protocol

Treatment materials	• Picture Cards • SFA Chart (see Figure 2) • Wipe-off board with markers and eraser
Overview	Semantic feature analysis is a treatment comprised of training whole words (e.g. nouns). The therapist teaches nouns by showing a picture (e.g. juice) and asking the participant a series of questions about that noun (e.g. what do you do with it?, where do you store it in your home?, etc.). The goal is to strengthen semantic networks within several categories (e.g. food and beverages, household items) and ultimately improve naming ability. A select set of words within each category is trained every session.
Feature Generation	The participant is first asked to name a given picture. Regardless of accuracy, the participant then verbally explores the semantic features of the picture in the context of five categories (see below). The therapist writes the generated features on a whiteboard. Upon completion of the feature generation, the participant is asked to name the picture three times in a row. If named accurately, the therapist will then show a new picture and repeat the same procedures. However, if the participant fails to accurately name the target, the therapist verbalizes the target and asks the participant to repeat the target three times. The therapist then reviews the chart, repeating the target word and one previously generated feature (e.g. 'juice belongs to the group *food and beverages*,' 'juice is a liquid,' etc.). Then the participant is again asked to repeat the target three times. Regardless of accuracy at this stage, the SLP proceeds to the subsequent item.
The Five Categories	
• Group	The *group* is the semantic category being trained (e.g. *food and beverages, household items*, etc.). The SLP asks, "What group does this belong to?" The participant generates only one feature for this category.
• Description	The *description* category explores the inherent properties of the pictured item. The SLP probes, "Let's describe it." The participant is encouraged to explore the color, texture, size, shape, and other associated perceptual characteristics. The participant generates a minimum of two features for this category.
• Function	The *function* category presents an opportunity to identify the uses and actions associated to the item. The SLP asks, "What is this used for?" or "What does this do?". The participant generates a minimum of two features for this category.
• Context	The *context* category is used to elicit responses related to the location or scene relating to the item. The SLP asks, "Where do you find it?" and "What places or other items are often associated with it?". The participant generates a minimum of two features for this category.
• Other/Personal	The *other/personal* category encourages the participant to share their own thoughts and personal stories related to the item. The SLP asks, "What does this remind you of?" or "What does this make you think of?". Only one personal association is required for this category.

Appendix E: Error Coding Guide

Complete Attempt (C): On each trial, score the subject's first "Complete Attempt", defined as: The first minimally CV (consonant-vowel) or VC response (schwa is not counted as a vowel).

Correct Responses

Correct ("1"): The target as given on the scoresheet (or acceptable probe responses).

Additional information

1. Allow for addition or deletion of plural morpheme.
2. Allow for addition of prepositional phrase *e.g.,* (can)→"can of peas".
3. Allow for addition of modifier, *e.g.,* (bone)→"dog's bone".
4. If the subject says "type of X" where X is the target, the answer is correct.
5. Do not penalize non-standard English speakers for omission of the final consonant in consonant clusters *e.g.,* (desk)→[dɛs], if it is consistent with their dialectal pattern.

PHONOLOGICAL ERRORS

Phonologically Related Real Word, No Obvious Semantic Relationship ("1"): A real word response that meets the criteria for Phonological Similarity to the target, but that is not obviously semantically related.

(mat) → "math"
(crab) → "cab"

Nonword - Phonologically Related ("1"): A nonword response that meets the criteria for Phonological Similarity to the target. (see "Phonological Similarity")

(dinosaur) → "/dɑɪ-nə-sɔɪ-əs/"

Phonological Similarity: Target and response share:

- the stressed vowel, initial or final phonemes;

(can) → "comb"
(well) → "ball"
(rope) → "hose"
(tractor) → "wagon"
(octopus) → "otter"
(vest) → "jacket"

- two or more phonemes (including stressed vowels but excluding unstressed vowels) at any position;

  (octopus) → "spider" (share two phonemes - /s/ and /p/), or

- one or more phonemes at a corresponding syllable and word positions, aligning words from left to right;

  (seal) → "dol-phin"(share the /l/)

(ca-me-ra) → "pho-to-graph" (/r/ at corresponding syllable position [onset of syllable] and corresponding word position [third syllable of target and response])
(ther-mo-me-ter) → "ca-mel"

NOTE:

- Plural morphemes and schwa are not counted in determining Phonological Similarity.
- Consonant clusters are treated as a unit when determining syllable position. The following do have Phonological Similarity:
  (cane) → "hand" (syllable position of /n/ corresponds as part of consonant cluster)
  (tractor) → "rug" (share the initial /r/)
  (plant) → "fern" (share the /n/)

Fails test for Phonological Similarity:

e.g., (type-wri-ter) → "/ɹɪŋ-kə/"
(/r/ is at a corresponding syllable position [onset of syllable] but not corresponding word position [second syllable of target versus first syllable of response])

- *e.g.*, (di-no-saur) → "pos-sum"
  (/s/ is at corresponding syllable position [onset of syllable] but not corresponding word position when aligned left-to-right [third syllable of target versus second syllable of response])
- *e.g.*, (map) → "pin"
  (One sound is in common, but it is not in a corresponding position.)

SEMANTIC ERRORS

Semantic Error ("+"): A real word noun response related to the target by one of the following:

synonym	(toilet)	→ "commode"
category coordinate	(banana)	→ "apple"
superordinate	(apple)	→ "fruit"
subordinate	(flower)	→ "rose"
associated	(bench)	→ "park"
diminutives	(dog)	→ " doggie"

- A related proper name that notes an association, *e.g.,* (star)→"Cowboys"
- The addition or substitution of a morpheme in a compound word:

(microscope) → "telescope"

NOT coded as semantic errors:

- A single verb, adjective, or adverb that has a semantic relationship to the target is not coded

(skis) → "skiing",
(dinosaur) → "scary "
(scarf) → "outside"

- A single verb, adjective or adverb that has no relationship to the target or does not characterize it in any way, *e.g.,* (beard)→"hit"

- A response in the form "type of X" where "X" is a superordinate of the target. *e.g.,* (apple)→"type of fruit"
- A response that negates the target, *e.g.,* (dinosaur)→"not a dinosaur"
- A response that includes a carrier phrase with the name of the target, *e.g.,* (door)→ "open the door

Appendix F: Linear Mixed-Effect Models

```
## READ IN DATA

# All
a <- read.csv("Linear_zAll_capped.csv")

all <- a %>%
  mutate_at(vars(Tx), list(factor)) %>%
  na.omit(a)
all2 <- subset(all, Time == "A2")
all3 <- subset(all, Time == "A3")


# PMT
b <- read.csv("Linear_zPMT.csv")

pmt <- na.omit(b)
pmt2 <- subset(pmt, Time == "A2")
pmt3 <- subset(pmt, Time == "A3")


# SFA
c <- read.csv("Linear_zSFA.csv")

sfa <- na.omit(c)
sfa2 <- subset(sfa, Time == "A2")
sfa3 <- subset(sfa, Time == "A3")
################################################################################
#### WHOLE GROUP ANALYSES
################################################################################

### Immediately Post-Treatment

## Treated Items

lmat2 <- lmer(Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx), data = all2)
summary(lmat2)
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx)
##    Data: all2
##
## REML criterion at convergence: 405.9
##
## Scaled residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -1.97718 -0.52363  0.03255  0.57829  2.14489
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Tx       (Intercept) 179.2    13.39
##  Residual             136.1    11.67
## Number of obs: 55, groups:  Tx, 2
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  31.3578     9.5983  0.9984   3.267 0.189416
## ZCAT_T        1.7924     2.2157 48.0018   0.809 0.422532
## ZSAPA2        2.2525     1.9039 48.0009   1.183 0.242605
## ZSAPA3        8.8336     2.1553 48.0426   4.099 0.000159 ***
## ZSem          4.4114     1.9664 48.0002   2.243 0.029524 *
## ZPhon         2.5312     1.8233 48.0582   1.388 0.171478
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr) ZCAT_T ZSAPA2 ZSAPA3 ZSem
## ZCAT_T -0.016
## ZSAPA2 -0.007 -0.224
## ZSAPA3 -0.003 -0.337 -0.234
## ZSem    0.009 -0.167  0.188 -0.423
## ZPhon   0.010 -0.283 -0.138 -0.083  0.290
## Untreated-Related Items

lmau2 <- lmer(Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx), data =
all2)
summary(lmau2)
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx)
##    Data: all2
##
## REML criterion at convergence: 326
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.95901 -0.60776 -0.01122  0.56121  2.16133
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Tx       (Intercept) 31.87    5.646
```

```
##  Residual             26.72    5.169
## Number of obs: 55, groups:  Tx, 2
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  13.0371     4.0538  0.9982   3.216   0.1923
## ZCAT_T        1.0130     0.9815 48.0020   1.032   0.3072
## ZSAPA2        1.0615     0.8434 48.0010   1.259   0.2143
## ZSAPA3        4.2867     0.9547 48.0468   4.490 4.46e-05 ***
## ZSem          1.1359     0.8711 48.0002   1.304   0.1985
## ZPhon         1.5198     0.8077 48.0640   1.882   0.0659 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr) ZCAT_T ZSAPA2 ZSAPA3 ZSem
## ZCAT_T -0.017
## ZSAPA2 -0.007 -0.224
## ZSAPA3 -0.003 -0.337 -0.234
## ZSem    0.009 -0.167  0.188 -0.423
## ZPhon   0.011 -0.283 -0.138 -0.083  0.290
## Untreated-Unrelated Items

lmap2 <- lmer(POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx), data =
all2)
## boundary (singular) fit: see ?isSingular
summary(lmap2)
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx)
##    Data: all2
##
## REML criterion at convergence: 308.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.44968 -0.70750 -0.04574  0.73913  1.93447
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Tx       (Intercept)  0.00    0.000
##  Residual             19.86    4.457
## Number of obs: 55, groups:  Tx, 2
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
```

```
## (Intercept)  11.6895     0.6070 49.0000  19.258  < 2e-16 ***
## ZCAT_T        0.7972     0.8460 49.0000   0.942 0.350670
## ZSAPA2        0.6420     0.7271 49.0000   0.883 0.381564
## ZSAPA3        3.0114     0.8168 49.0000   3.687 0.000568 ***
## ZSem          2.1812     0.7511 49.0000   2.904 0.005509 **
## ZPhon         1.4069     0.6890 49.0000   2.042 0.046566 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr) ZCAT_T ZSAPA2 ZSAPA3 ZSem
## ZCAT_T -0.095
## ZSAPA2 -0.042 -0.224
## ZSAPA3 -0.018 -0.337 -0.239
## ZSem    0.051 -0.167  0.188 -0.427
## ZPhon   0.064 -0.290 -0.137 -0.066  0.294
## convergence code: 0
## boundary (singular) fit: see ?isSingular
#####################################################
### Follow-Up
#####################################################

## Treated Items

lmat3 <- lmer(Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx), data = all3)
summary(lmat3)
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx)
##    Data: all3
##
## REML criterion at convergence: 399.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.4585 -0.4564 -0.1250  0.5587  2.5784
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Tx       (Intercept) 143.4    11.97
##  Residual             120.1    10.96
## Number of obs: 55, groups:  Tx, 2
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  26.9912     8.5968  0.9982   3.140 0.196650
```

```
## ZCAT_T       1.5763     2.0810 48.0020   0.757 0.452471
## ZSAPA2       2.6204     1.7881 48.0010   1.465 0.149326
## ZSAPA3       8.3798     2.0242 48.0468   4.140 0.000139 ***
## ZSem         3.5738     1.8468 48.0002   1.935 0.058882 .
## ZPhon        1.7426     1.7124 48.0639   1.018 0.313957
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##       (Intr) ZCAT_T ZSAPA2 ZSAPA3 ZSem
## ZCAT_T -0.017
## ZSAPA2 -0.007 -0.224
## ZSAPA3 -0.003 -0.337 -0.234
## ZSem    0.009 -0.167  0.188 -0.423
## ZPhon   0.011 -0.283 -0.138 -0.083  0.290
## Untreated-Related Items

lmau3 <- lmer(Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx), data =
all3)
summary(lmau3)
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx)
##    Data: all3
##
## REML criterion at convergence: 334.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.99604 -0.60115 -0.02208  0.49927  2.43871
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Tx       (Intercept) 23.41    4.839
##  Residual             32.24    5.678
## Number of obs: 55, groups:  Tx, 2
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  12.6096     3.5078  0.9972   3.595    0.173
## ZCAT_T       0.8908     1.0782 48.0032   0.826    0.413
## ZSAPA2       0.9795     0.9265 48.0017   1.057    0.296
## ZSAPA3       4.4589     1.0486 48.0750   4.252  9.7e-05 ***
## ZSem         0.6898     0.9569 48.0003   0.721    0.474
## ZPhon        0.7743     0.8870 48.1021   0.873    0.387
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr) ZCAT_T ZSAPA2 ZSAPA3 ZSem
## ZCAT_T -0.021
## ZSAPA2 -0.009 -0.224
## ZSAPA3 -0.004 -0.337 -0.234
## ZSem    0.011 -0.167  0.188 -0.423
## ZPhon   0.014 -0.283 -0.138 -0.083  0.290
## Untreated-Unrelated Items

lmap3 <- lmer(POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx), data =
all3)
## boundary (singular) fit: see ?isSingular
summary(lmap3)
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | Tx)
##    Data: all3
##
## REML criterion at convergence: 330.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.2449 -0.5040 -0.0655  0.6616  2.0274
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Tx       (Intercept)  0.00    0.000
##  Residual             31.19    5.585
## Number of obs: 55, groups:  Tx, 2
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  11.3802     0.7607 49.0000  14.960   <2e-16 ***
## ZCAT_T        1.4079     1.0603 49.0000   1.328   0.1904
## ZSAPA2        1.4621     0.9112 49.0000   1.605   0.1150
## ZSAPA3        2.3881     1.0237 49.0000   2.333   0.0238 *
## ZSem          1.5300     0.9412 49.0000   1.626   0.1105
## ZPhon         0.5477     0.8635 49.0000   0.634   0.5288
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr) ZCAT_T ZSAPA2 ZSAPA3 ZSem
## ZCAT_T -0.095
```

```
## ZSAPA2 -0.042 -0.224
## ZSAPA3 -0.018 -0.337 -0.239
## ZSem    0.051 -0.167  0.188 -0.427
## ZPhon   0.064 -0.290 -0.137 -0.066  0.294
## convergence code: 0
## boundary (singular) fit: see ?isSingular
################################################################################
#### PMT GROUP ANALYSES
################################################################################

### Immediately Post-Treatment

## Treated Items

lmpt2 <- lm(Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = pmt2)
summary(lmpt2)
##
## Call:
## lm(formula = Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = pmt2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.4635 -4.6673 -0.6277  3.6739 11.7319
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  20.9161     1.1643  17.965 3.17e-14 ***
## ZCAT_T        1.9491     1.6154   1.207  0.24102
## ZSAPA2       -0.8622     1.4231  -0.606  0.55108
## ZSAPA3        5.2558     1.4985   3.507  0.00210 **
## ZSem          4.3356     1.3342   3.250  0.00383 **
## ZPhon         4.5481     1.5829   2.873  0.00910 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.994 on 21 degrees of freedom
## Multiple R-squared:  0.7632, Adjusted R-squared:  0.7068
## F-statistic: 13.54 on 5 and 21 DF,  p-value: 5.671e-06
## Untreated-Related Items

lmpu2 <- lm(Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = pmt2)
summary(lmpu2)
##
## Call:
## lm(formula = Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
```

```
##     data = pmt2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6160 -1.6359  0.2473  1.8201  5.2962
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   8.6056     0.5551  15.503 5.68e-13 ***
## ZCAT_T        1.6690     0.7702   2.167 0.041885 *
## ZSAPA2        0.5428     0.6785   0.800 0.432666
## ZSAPA3        2.7340     0.7144   3.827 0.000982 ***
## ZSem          1.0809     0.6361   1.699 0.104014
## ZPhon         1.8869     0.7547   2.500 0.020764 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.858 on 21 degrees of freedom
## Multiple R-squared:  0.7978, Adjusted R-squared:  0.7496
## F-statistic: 16.57 on 5 and 21 DF,  p-value: 1.146e-06
## Untreated-Unrelated Items

lmpp2 <- lm(POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = pmt2)
summary(lmpp2)
##
## Call:
## lm(formula = POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = pmt2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.3608 -2.6031 -0.7038  2.8273  8.5003
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.8588     0.8901  12.199 5.37e-11 ***
## ZCAT_T        1.1326     1.2350   0.917   0.3695
## ZSAPA2        1.0783     1.0880   0.991   0.3329
## ZSAPA3        2.2127     1.1456   1.931   0.0670 .
## ZSem          2.0057     1.0200   1.966   0.0626 .
## ZPhon         0.4508     1.2102   0.373   0.7132
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.582 on 21 degrees of freedom
```

```
## Multiple R-squared:  0.4983, Adjusted R-squared:  0.3788
## F-statistic: 4.171 on 5 and 21 DF,  p-value: 0.008663
#####################################################
### Follow-Up
#####################################################

## Treated Items

lmpt3 <- lm(Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = pmt3)
summary(lmpt3)
##
## Call:
## lm(formula = Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = pmt3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.4507 -4.8960 -0.6597  4.1555 10.8884
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  17.4651     1.1385  15.340 6.98e-13 ***
## ZCAT_T        2.2651     1.5797   1.434  0.16634
## ZSAPA2        0.6303     1.3916   0.453  0.65525
## ZSAPA3        4.8379     1.4654   3.302  0.00340 **
## ZSem          3.7808     1.3047   2.898  0.00861 **
## ZPhon         3.5764     1.5479   2.310  0.03111 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.861 on 21 degrees of freedom
## Multiple R-squared:  0.7516, Adjusted R-squared:  0.6925
## F-statistic: 12.71 on 5 and 21 DF,  p-value: 9.17e-06
## Untreated-Related Items

lmpu3 <- lm(Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = pmt3)
summary(lmpu3)
##
## Call:
## lm(formula = Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = pmt3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0889 -2.0428 -0.0593  1.8972  6.0186
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   8.6210     0.5537  15.570 5.23e-13 ***
## ZCAT_T        1.8734     0.7682   2.439  0.02371 *
## ZSAPA2        0.6900     0.6768   1.020  0.31951
## ZSAPA3        2.5437     0.7126   3.569  0.00181 **
## ZSem          1.1184     0.6345   1.763  0.09250 .
## ZPhon         1.3462     0.7528   1.788  0.08817 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.85 on 21 degrees of freedom
## Multiple R-squared:  0.7789, Adjusted R-squared:  0.7263
## F-statistic:  14.8 on 5 and 21 DF,  p-value: 2.831e-06
## Untreated-Unrelated Items

lmpp3 <- lm(POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = pmt3)
summary(lmpp3)
##
## Call:
## lm(formula = POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = pmt3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.8307 -3.4283 -0.9288  2.5975  9.2737
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  11.1073     0.9451  11.752 1.07e-10 ***
## ZCAT_T        0.8295     1.3113   0.633   0.5338
## ZSAPA2        1.5746     1.1552   1.363   0.1873
## ZSAPA3        1.9834     1.2164   1.631   0.1179
## ZSem          2.7299     1.0830   2.521   0.0199 *
## ZPhon         0.8135     1.2849   0.633   0.5335
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.865 on 21 degrees of freedom
## Multiple R-squared:  0.5167, Adjusted R-squared:  0.4016
## F-statistic:  4.49 on 5 and 21 DF,  p-value: 0.006126
################################################################################
#### SFA GROUP ANALYSES
################################################################################

### Immediately Post-Treatment
```

```
## Treated Items

lmst2 <- lm(Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = sfa2)
summary(lmst2)
##
## Call:
## lm(formula = Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = sfa2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25.2823 -12.3298   0.9466   8.0516  28.4564
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  41.7756     2.9236  14.289  1.3e-12 ***
## ZCAT_T        0.7152     4.3489   0.164   0.8709
## ZSAPA2        4.8429     3.8061   1.272   0.2165
## ZSAPA3       11.5032     4.2952   2.678   0.0137 *
## ZSem          3.8382     3.9911   0.962   0.3467
## ZPhon         3.0583     3.0748   0.995   0.3307
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.29 on 22 degrees of freedom
## Multiple R-squared:  0.609,  Adjusted R-squared:  0.5201
## F-statistic: 6.853 on 5 and 22 DF,  p-value: 0.0005414
## Untreated-Related Items

lmsu2 <- lm(Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = sfa2)
summary(lmsu2)
##
## Call:
## lm(formula = Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = sfa2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.5714 -6.1507 -0.3902  4.8976 12.2213
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  17.4530     1.3243  13.179 6.44e-12 ***
## ZCAT_T        0.2412     1.9699   0.122  0.90367
## ZSAPA2        1.4415     1.7241   0.836  0.41208
```

```
## ZSAPA3       5.6652     1.9456   2.912  0.00808 **
## ZSem         0.9742     1.8078   0.539  0.59538
## ZPhon        1.6464     1.3928   1.182  0.24980
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.927 on 22 degrees of freedom
## Multiple R-squared:   0.58,  Adjusted R-squared:  0.4846
## F-statistic: 6.077 on 5 and 22 DF,  p-value: 0.001116
## Untreated-Unrelated Items

lmsp2 <- lm(POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = sfa2)
summary(lmsp2)
##
## Call:
## lm(formula = POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = sfa2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.6417 -3.2624 -0.3861  2.5254  8.1317
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  12.3940     0.8612  14.391 1.13e-12 ***
## ZCAT_T        0.7785     1.2811   0.608  0.54962
## ZSAPA2        0.3583     1.1212   0.320  0.75231
## ZSAPA3        4.1396     1.2653   3.272  0.00349 **
## ZSem          1.6507     1.1757   1.404  0.17426
## ZPhon         2.4387     0.9058   2.692  0.01330 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.504 on 22 degrees of freedom
## Multiple R-squared:  0.7142, Adjusted R-squared:  0.6493
## F-statistic:    11 on 5 and 22 DF,  p-value: 2.124e-05
#####################################################
### Follow-Up
#####################################################

## Treated Items

lmst3 <- lm(Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = sfa3)
summary(lmst3)
##
## Call:
```

```
## lm(formula = Trt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = sfa3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -25.7484  -9.7308  -0.6797   9.9455  30.4191
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  36.5065     2.7858  13.104  7.2e-12 ***
## ZCAT_T        0.3322     4.1440   0.080   0.9368
## ZSAPA2        4.0126     3.6268   1.106   0.2805
## ZSAPA3       11.1874     4.0928   2.733   0.0121 *
## ZSem          3.0394     3.8030   0.799   0.4327
## ZPhon         1.7859     2.9299   0.610   0.5484
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.57 on 22 degrees of freedom
## Multiple R-squared:  0.5803, Adjusted R-squared:  0.485
## F-statistic: 6.085 on 5 and 22 DF,  p-value: 0.001107
## Untreated-Related Items

lmsu3 <- lm(Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = sfa3)
summary(lmsu3)
##
## Call:
## lm(formula = Untrt ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = sfa3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -10.6208  -5.8876  -0.6948   5.0458  14.4748
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 16.586933   1.469392  11.288 1.27e-10 ***
## ZCAT_T      -0.004299   2.185752  -0.002  0.99845
## ZSAPA2       0.958904   1.912948   0.501  0.62116
## ZSAPA3       6.231679   2.158743   2.887  0.00856 **
## ZSem         0.174323   2.005914   0.087  0.93153
## ZPhon        0.669710   1.545379   0.433  0.66897
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.685 on 22 degrees of freedom
```

```
## Multiple R-squared:  0.4927, Adjusted R-squared:  0.3774
## F-statistic: 4.274 on 5 and 22 DF,  p-value: 0.007234
## Untreated-Unrelated Items

lmsp3 <- lm(POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon, data = sfa3)
summary(lmsp3)
##
## Call:
## lm(formula = POM ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon,
##     data = sfa3)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.9707  -3.5951  -0.3171   3.9282  11.2230
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  11.5357     1.2381   9.317  4.3e-09 ***
## ZCAT_T        2.5177     1.8417   1.367   0.1854
## ZSAPA2        1.0883     1.6118   0.675   0.5066
## ZSAPA3        3.2021     1.8189   1.760   0.0922 .
## ZSem         -0.1368     1.6902  -0.081   0.9362
## ZPhon         0.7128     1.3021   0.547   0.5896
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.476 on 22 degrees of freedom
## Multiple R-squared:  0.4774, Adjusted R-squared:  0.3586
## F-statistic: 4.019 on 5 and 22 DF,  p-value: 0.009647
```

Appendix G: R Code for Main and Post Hoc Analyses

```
z <- read.csv("Logistic_zAll.csv")

data <- z %>%
  mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
  na.omit(z)
data <- subset(data, Time != "A1")

t <- subset(data, Type_Key == "T")
u <- subset(data, Type_Key == "U")
p <- subset(data, Type_Key == "POM")

a <- read.csv("zAllA1.csv")
a <- a[, colnames(a) %in% c("ZCAT_T", "ZSAPA2", "ZSAPA3", "ZSem", "ZPhon")] %
>%
  na.omit(a)

################################################################################
## Whole Group Correlations
################################################################################
all_corr <- rcorr(as.matrix(a))

################################################################################
## Relations between Individual Predictors and Outcomes
################################################################################
## Treated

# A1-Outcome Relationship
t_A1 <- glmer(Response ~ 1 + ZpctA1 + (1|SubID) + (1|Item_No), data = t, fami
ly = binomial)

# CAT Comprehension Score Relationship
tCAT <- glmer(Response ~ 1 + ZCAT_T + (1|SubID) + (1|Item_No), data = t, fami
ly = binomial)

# SAPA2 Score Relationship
tSAPA2 <- glmer(Response ~ 1 + ZSAPA2 + (1|SubID) + (1|Item_No), data = t, fa
mily = binomial)

# SAPA3 Score Relationship
tSAPA3 <- glmer(Response ~ 1+ ZSAPA3 + (1|SubID) + (1|Item_No), data = t, fam
ily = binomial)

# Semantic Score Relationship
tSem <- glmer(Response ~ 1 + ZSem + (1|SubID) + (1|Item_No), data = t, family
= binomial)

# Phoological Score Relationship
tPhon <- glmer(Response ~ 1 + ZPhon + (1|SubID) + (1|Item_No), data = t, fami
```

```
ly = binomial)

tS_P <- glmer(Response ~ 1 + ZSem*ZPhon + (1|SubID) + (1|Item_No), data = t,
family = binomial)

##################################################

## Untreated

# A1-Outcome Relationship
u_A1 <- glmer(Response ~ 1 + ZpctA1 + (1|SubID) + (1|Item_No), data = u, fami
ly = binomial)

# CAT Comprehension Score Relationship
uCAT <- glmer(Response ~ 1 +  ZCAT_T + (1|SubID) + (1|Item_No), data = u, fam
ily = binomial)

# SAPA2 Score Relationship
uSAPA2 <- glmer(Response ~ 1 + ZSAPA2 + (1|SubID) + (1|Item_No), data = u, fa
mily = binomial)

# SAPA3 Score Relationship
uSAPA3 <- glmer(Response ~ 1 + ZSAPA3 + (1|SubID) + (1|Item_No), data = u, fa
mily = binomial)

# Semantic Score Relationship
uSem <- glmer(Response ~ 1 + ZSem + (1|SubID) + (1|Item_No), data = u, family
= binomial)

# Phoological Score Relationship
uPhon <- glmer(Response ~ 1 + ZPhon + (1|SubID) + (1|Item_No), data = u, fami
ly = binomial)

uS_P <- glmer(Response ~ 1 + ZSem*ZPhon + (1|SubID) + (1|Item_No), data = u,
family = binomial)

##################################################

## POM

# A1-Outcome Relationship
p_A1 <- glmer(Response ~ 1 + ZpctA1 + (1|SubID) + (1|Item_No), data = p, fami
ly = binomial)

# CAT Comprehension Score Relationship
pCAT <- glmer(Response ~ 1 + ZCAT_T + (1|SubID) + (1|Item_No), data = p, fami
ly = binomial)

# SAPA2 Score Relationship
pSAPA2 <- glmer(Response ~ 1 + ZSAPA2 + (1|SubID) + (1|Item_No), data = p, fa
```

```
mily = binomial)

# SAPA3 Score Relationship
pSAPA3 <- glmer(Response ~ 1 + ZSAPA3 + (1|SubID) + (1|Item_No), data = p, fa
mily = binomial)

# Semantic Score Relationship
pSem <- glmer(Response ~ 1 + ZSem + (1|SubID) + (1|Item_No), data = p, family
= binomial)

# Phoological Score Relationship
pPhon <- glmer(Response ~ 1 + ZPhon + (1|SubID) + (1|Item_No), data = p, fami
ly = binomial)

pS_P <- glmer(Response ~ 1 + ZSem*ZPhon + (1|SubID) + (1|Item_No), data = p,
family = binomial)

################################################################################
## Whole Group Correlations
################################################################################
pmt_corr <- rcorr(as.matrix(b))

################################################################################
## Relations between Individual Predictors and Outcomes
################################################################################


## Treated

# A1-Outcome Relationship
pt_A1 <- glmer(Response ~ 1 + ZpctA1 + (1|SubID) + (1|Item_No), data = pt, fa
mily = binomial)

# CAT Comprehension Score Relationship
ptCAT <- glmer(Response ~ 1 + ZCAT_T + (1|SubID) + (1|Item_No), data = pt, fa
mily = binomial)

# SAPA2 Score Relationship
ptSAPA2 <- glmer(Response ~ 1 + ZSAPA2 + (1|SubID) + (1|Item_No), data = pt,
family = binomial)

# SAPA3 Score Relationship
ptSAPA3 <- glmer(Response ~ 1+ ZSAPA3 + (1|SubID) + (1|Item_No), data = pt, f
amily = binomial)

# Semantic Score Relationship
ptSem <- glmer(Response ~ 1 + ZSem + (1|SubID) + (1|Item_No), data = pt, fami
ly = binomial)

# Phoological Score Relationship
```

```
ptPhon <- glmer(Response ~ 1 + ZPhon + (1|SubID) + (1|Item_No), data = pt, fa
mily = binomial)

ptS_P <- glmer(Response ~ 1 + ZSem*ZPhon + (1|SubID) + (1|Item_No), data = pt
, family = binomial)

##################################################

## Untreated

# A1-Outcome Relationship
pu_A1 <- glmer(Response ~ 1 + ZpctA1 + (1|SubID) + (1|Item_No), data = pu, fa
mily = binomial)

# CAT Comprehension Score Relationship
puCAT <- glmer(Response ~ 1 +  ZCAT_T + (1|SubID) + (1|Item_No), data = pu, f
amily = binomial)

# SAPA2 Score Relationship
puSAPA2 <- glmer(Response ~ 1 + ZSAPA2 + (1|SubID) + (1|Item_No), data = pu,
family = binomial)

# SAPA3 Score Relationship
puSAPA3 <- glmer(Response ~ 1 + ZSAPA3 + (1|SubID) + (1|Item_No), data = pu,
family = binomial)

# Semantic Score Relationship
puSem <- glmer(Response ~ 1 + ZSem + (1|SubID) + (1|Item_No), data = pu, fami
ly = binomial)

# Phoological Score Relationship
puPhon <- glmer(Response ~ 1 + ZPhon + (1|SubID) + (1|Item_No), data = pu, fa
mily = binomial)

puS_P <- glmer(Response ~ 1 + ZSem*ZPhon + (1|SubID) + (1|Item_No), data = pu
, family = binomial)

##################################################

## POM

# A1-Outcome Relationship
pp_A1 <- glmer(Response ~ 1 + ZpctA1 + (1|SubID) + (1|Item_No), data = pp, fa
mily = binomial)

# CAT Comprehension Score Relationship
ppCAT <- glmer(Response ~ 1 + ZCAT_T + (1|SubID) + (1|Item_No), data = pp, fa
mily = binomial)

# SAPA2 Score Relationship
```

```
ppSAPA2 <- glmer(Response ~ 1 + ZSAPA2 + (1|SubID) + (1|Item_No), data = pp,
family = binomial)

# SAPA3 Score Relationship
ppSAPA3 <- glmer(Response ~ 1 + ZSAPA3 + (1|SubID) + (1|Item_No), data = pp,
family = binomial)

# Semantic Score Relationship
ppSem <- glmer(Response ~ 1 + ZSem + (1|SubID) + (1|Item_No), data = pp, fami
ly = binomial)

# Phoological Score Relationship
ppPhon <- glmer(Response ~ 1 + ZPhon + (1|SubID) + (1|Item_No), data = pp, fa
mily = binomial)

ppS_P <- glmer(Response ~ 1 + ZSem*ZPhon + (1|SubID) + (1|Item_No), data = pp
, family = binomial)

################################################################################
## Whole Group Correlations
################################################################################
sfa_corr <- rcorr(as.matrix(c))

################################################################################
## Relations between Individual Predictors and Outcomes
################################################################################


## Treated

# A1-Outcome Relationship
st_A1 <- glmer(Response ~ 1 + ZpctA1 + (1|SubID) + (1|Item_No), data = st, fa
mily = binomial)

# CAT Comprehension Score Relationship
stCAT <- glmer(Response ~ 1 + ZCAT_T + (1|SubID) + (1|Item_No), data = st, fa
mily = binomial)

# SAPA2 Score Relationship
stSAPA2 <- glmer(Response ~ 1 + ZSAPA2 + (1|SubID) + (1|Item_No), data = st,
family = binomial)

# SAPA3 Score Relationship
stSAPA3 <- glmer(Response ~ 1+ ZSAPA3 + (1|SubID) + (1|Item_No), data = st, f
amily = binomial)

# Semantic Score Relationship
stSem <- glmer(Response ~ 1 + ZSem + (1|SubID) + (1|Item_No), data = st, fami
ly = binomial)
```

```
# Phoological Score Relationship
stPhon <- glmer(Response ~ 1 + ZPhon + (1|SubID) + (1|Item_No), data = st, fa
mily = binomial)

stS_P <- glmer(Response ~ 1 + ZSem*ZPhon + (1|SubID) + (1|Item_No), data = st
, family = binomial)

################################################

## Untreated

# A1-Outcome Relationship
su_A1 <- glmer(Response ~ 1 + ZpctA1 + (1|SubID) + (1|Item_No), data = su, fa
mily = binomial)

# CAT Comprehension Score Relationship
suCAT <- glmer(Response ~ 1 +  ZCAT_T + (1|SubID) + (1|Item_No), data = su, f
amily = binomial)

# SAPA2 Score Relationship
suSAPA2 <- glmer(Response ~ 1 + ZSAPA2 + (1|SubID) + (1|Item_No), data = su,
family = binomial)

# SAPA3 Score Relationship
suSAPA3 <- glmer(Response ~ 1 + ZSAPA3 + (1|SubID) + (1|Item_No), data = su,
family = binomial)

# Semantic Score Relationship
suSem <- glmer(Response ~ 1 + ZSem + (1|SubID) + (1|Item_No), data = su, fami
ly = binomial)

# Phoological Score Relationship
suPhon <- glmer(Response ~ 1 + ZPhon + (1|SubID) + (1|Item_No), data = su, fa
mily = binomial)

suS_P <- glmer(Response ~ 1 + ZSem*ZPhon + (1|SubID) + (1|Item_No), data = su
, family = binomial)

################################################

## POM

# A1-Outcome Relationship
sp_A1 <- glmer(Response ~ 1 + ZpctA1 + (1|SubID) + (1|Item_No), data = sp, fa
mily = binomial)

# CAT Comprehension Score Relationship
spCAT <- glmer(Response ~ 1 + ZCAT_T + (1|SubID) + (1|Item_No), data = sp, fa
mily = binomial)
```

```
# SAPA2 Score Relationship
spSAPA2 <- glmer(Response ~ 1 + ZSAPA2 + (1|SubID) + (1|Item_No), data = sp,
family = binomial)

# SAPA3 Score Relationship
spSAPA3 <- glmer(Response ~ 1 + ZSAPA3 + (1|SubID) + (1|Item_No), data = sp,
family = binomial)

# Semantic Score Relationship
spSem <- glmer(Response ~ 1 + ZSem + (1|SubID) + (1|Item_No), data = sp, fami
ly = binomial)

# Phoological Score Relationship
spPhon <- glmer(Response ~ 1 + ZPhon + (1|SubID) + (1|Item_No), data = sp, fa
mily = binomial)

spS_P <- glmer(Response ~ 1 + ZSem*ZPhon + (1|SubID) + (1|Item_No), data = sp
, family = binomial)
```

```
################################################################################
## Whole Group -- A2 Data
################################################################################
# z <- read.csv("Logistic_zAll_capped1.csv")

# data <- z %>%
#  mutate_at(vars(c(Tx_Code, Time_Code, Item_No)), list(factor)) %>%
#  na.omit(z)
# data <- subset(data, Time == "A2")
# head(data)
# summary(data)

# t2 <- subset(data, Type_Key == "T")
# u2 <- subset(data, Type_Key == "U")
# p2 <- subset(data, Type_Key == "POM")

# head(t2)
# summary(t2)
# head(u2)
# head(p2)
################################################################################
## Whole Group -- A2 Treated
################################################################################

## Treated Regressions

# Failed to converge
# allt2.0 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No) + (1|Tx_Code), data = t2, family = binomial)
```

```
# allt2.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = t2, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.allt2 <- influence(allt2.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_allt2.jpg")
# infIndexPlot(estex.allt2, var = "Cook",id=list(method="y", n=4), main = NULL)
# dev.off()

# Exclude Influential Data

# t2.1 <- subset(t2, SubID != c("SF09SEA", "SF10PDX", "SF12SEA", "SF18SEA"))
# summary(t2.1)

# Re-analysis without influential data

# allt2.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = t2.1, family = binomial)

# allt2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = t2.1, family = binomial)

# Summary and Model Comparisons

# summary(allt2.2)
# anova(allt2.2, allt2.null2)
# at2.models <- list(allt2.1, allt2.2, allt2.null, allt2.null2)
# at2.model.names <- c("allt2.1", "allt2.2", "allt2.null", "allt2.null2")
# aictab(cand.set = at2.models, modnames = at2.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(allt2.2)

# at2.predicted <- plogis(predict(allt2.2, t2.1))

# optCutOff <- optimalCutoff(t2.1$Response, at2.predicted)[1]

# sensitivity(t2.1$Response, at2.predicted, threshold = optCutOff)

# specificity(t2.1$Response, at2.predicted, threshold = optCutOff)
```

```
# Predicted Probabilities

# pp.at21 <- ggpredict(allt2.2, "ZCAT_T [-2:2]")
# pp.at22 <- ggpredict(allt2.2, "ZSAPA2[-2:2]")
# pp.at23 <- ggpredict(allt2.2, "ZSAPA3[-2:2]")
# pp.at24 <- ggpredict(allt2.2, "ZSem[-2:2]")
# pp.at25 <- ggpredict(allt2.2, "ZPhon[-2:2]")

# pp.at21
# pp.at22
# pp.at23
# pp.at24
# pp.at25

# Plot Results

# at.plot21 <-plot(pp.at21) + labs(title = "Treated Items", x = "CAT Comprehension z-Score", y
= "Predicted Probability")
# at.plot22 <-plot(pp.at22) + labs(title = "Treated Items", x = "SAPA 2 z-Score", y = "Predicted
Probability")
# at.plot23 <-plot(pp.at23) + labs(title = "Treated Items", x = "SAPA 3 z-Score", y = "Predicted
Probability")
# at.plot24 <-plot(pp.at24) + labs(title = "Treated Items", x = "Semantic Impairment z-Score", y
= "Predicted Probability")
# at.plot25 <-plot(pp.at25) + labs(title = "Treated Items", x = "Phonological Impairment z-
Score", y = "Predicted Probability")

# Table

# tab_model(allt2.2, allt2.null2,
#   dv.labels = c("All Treated Immed. Post-Tx Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
##############################################################################
## Whole Group -- A2 Untreated-Related
##############################################################################

## Untreated Regressions

# allu2.0 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No) + (1|Tx_Code), data = u2, family = binomial)
```

```
# allu2.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = u2, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.allu2 <- influence(allu2.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_allu2.jpg")
# infIndexPlot(estex.allu2, id=list(method="y", n=5), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data

# u2.1 <- subset(u2, SubID != c("PM15SEA", "SF02PDX", "SF10PDX", "SF12SEA",
"SF18SEA"))

# Re-analysis without influential data

# allu2.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = u2.1, family = binomial)

# allu2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = u2.1, family = binomial)

# Summary and Model Comparisons

# summary(allu2.2)
# anova(allu2.2, allu2.null)
# au2.models <- list(allu2.0, allu2.1, allu2.2, allu2.null)
# au2.model.names <- c("allu2.0", "allu2.1", "allu2.2", "allu2.null")
# aictab(cand.set = au2.models, modnames = au2.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(allu2.2)

# au2.predicted <- plogis(predict(allu2.2, u2.1))

# optCutOff <- optimalCutoff(u2.1$Response, au2.predicted)[1]

# sensitivity(u2.1$Response, au2.predicted, threshold = optCutOff)

# specificity(u2.1$Response, au2.predicted, threshold = optCutOff)
```

```
# Predicted Probabilities

# pp.au21 <- ggpredict(allu2.2, "ZCAT_T [-2:2]")
# pp.au22 <- ggpredict(allu2.2, "ZSAPA2[-2:2]")
# pp.au23 <- ggpredict(allu2.2, "ZSAPA3[-2:2]")
# pp.au24 <- ggpredict(allu2.2, "ZSem[-2:2]")
# pp.au25 <- ggpredict(allu2.2, "ZPhon[-2:2]")

# pp.au21
# pp.au22
# pp.au23
# pp.au24
# pp.au25

# Plot Results

# au.plot21 <-plot(pp.au21) + labs(title = "Untreated-Related Items", x = "CAT Comprehension
z-Score", y = "Predicted Probability")
# au.plot22 <-plot(pp.au22) + labs(title = "Untreated-Related Items", x = "SAPA 2 z-Score", y =
"Predicted Probability")
# au.plot23 <-plot(pp.au23) + labs(title = "Untreated-Related Items", x = "SAPA 3 z-Score", y =
"Predicted Probability")
# au.plot24 <-plot(pp.au24) + labs(title = "Untreated-Related Items", x = "Semantic Impairment
z-Score", y = "Predicted Probability")
# au.plot25 <-plot(pp.au25) + labs(title = "Untreated-Related Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(allu2.2, allu2.null,
#   dv.labels = c("All Untreated-Related Immed. Post-Tx Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## Whole Group -- A2 Untreated-Unrelated
################################################################################

## Untreated Regressions

# allp2.0 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No) + (1|Tx_Code), data = p2, family = binomial)
```

```
# allp2.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = p2, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.allp2 <- influence(allp2.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_allp2.jpg")
# infIndexPlot(estex.allp2, id=list(method="y", n=4), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data

# p2.1 <- subset(p2, SubID != c("PM12SEA", "SF04PDX"))

# Re-analysis without influential data

# allp2.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = p2.1, family = binomial)

# allp2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = p2.1, family = binomial)

# Summary and Model Comparisons

# summary(allp2.2)
# anova(allp2.2, allp2.null)
# ap2.models <- list(allp2.0, allp2.1, allp2.2, allp2.null)
# ap2.model.names <- c("allp2.0", "allp2.1", "allp2.2", "allp2.null")
# aictab(cand.set = ap2.models, modnames = ap2.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(allp2.2)

# ap2.predicted <- plogis(predict(allp2.2, p2.1))

# optCutOff <- optimalCutoff(p2.1$Response, ap2.predicted)[1]

# sensitivity(p2.1$Response, ap2.predicted, threshold = optCutOff)

# specificity(p2.1$Response, ap2.predicted, threshold = optCutOff)

# Predicted Probabilities
```

```
# pp.ap21 <- ggpredict(allp2.2, "ZCAT_T [-2:2]")
# pp.ap22 <- ggpredict(allp2.2, "ZSAPA2[-2:2]")
# pp.ap23 <- ggpredict(allp2.2, "ZSAPA3[-2:2]")
# pp.ap24 <- ggpredict(allp2.2, "ZSem[-2:2]")
# pp.ap25 <- ggpredict(allp2.2, "ZPhon[-2:2]")

# pp.ap21
# pp.ap22
# pp.ap23
# pp.ap24
# pp.ap25

# Plot Results

# ap.plot21 <-plot(pp.ap21) + labs(title = "Untreated-Unrelated Items", x = "CAT
Comprehension z-Score", y = "Predicted Probability")
# ap.plot22 <-plot(pp.ap22) + labs(title = "Untreated-Unrelated Items", x = "SAPA 2 z-Score", y
= "Predicted Probability")
# ap.plot23 <-plot(pp.ap23) + labs(title = "Untreated-Unrelated Items", x = "SAPA 3 z-Score", y
= "Predicted Probability")
# ap.plot24 <-plot(pp.ap24) + labs(title = "Untreated-Unrelated Items", x = "Semantic
Impairment z-Score", y = "Predicted Probability")
# ap.plot25 <-plot(pp.ap25) + labs(title = "Untreated-Unrelated Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(allp2.2, allp2.null,
#   dv.labels = c("All Untreated-Unrelated Immed. Post-Tx Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
##############################################################################
## Whole Group -- A3 Data
##############################################################################
# y <- read.csv("Logistic_zAll_capped1.csv")
# head(y)

# data1 <- y %>%
#   mutate_at(vars(c(Tx_Code, Time_Code, Item_No)), list(factor)) %>%
#   na.omit(y)
# data1 <- subset(data1, Time == "A3")
```

```
# head(data1)
# summary(data1)

# t3 <- subset(data1, Type_Key == "T")
# u3 <- subset(data1, Type_Key == "U")
# p3 <- subset(data1, Type_Key == "POM")

# head(t3)
# head(u3)
# head(p3)
################################################################################
## Whole Group -- A3 Treated
################################################################################

## Treated Regressions

# allt3.0 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No) + (1|Tx_Code), data = t3, family = binomial)

# allt3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = t3, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.allt3 <- influence(allt3.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_allt3.jpg")
# infIndexPlot(estex.allt3, var = "Cook",id=list(method="y", n=5), main = NULL)
# dev.off()

# Exclude Influential Data

# t3.1 <- subset(t3, SubID != c("PM15SEA", "SF10PDX", "SF12SEA", "SF18SEA"))
# summary(t3.1)

# Re-analysis without influential data

# allt3.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = t3.1, family = binomial)

# allt3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = t3.1, family = binomial)

# Summary and Model Comparisons
```

```
# summary(allt3.2)
# anova(allt3.2, allt3.null)
# at3.models <- list(allt3.1, allt3.2, allt3.null)
# at3.model.names <- c("allt3.1", "allt3.2", "allt3.null")
# aictab(cand.set = at3.models, modnames = at3.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(allt3.2)

# at3.predicted <- plogis(predict(allt3.2, t3.1))

# optCutOff <- optimalCutoff(t3.1$Response, at3.predicted)[1]

# sensitivity(t3.1$Response, at3.predicted, threshold = optCutOff)

# specificity(t3.1$Response, at3.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.at31 <- ggpredict(allt3.2, "ZCAT_T [-2:2]")
# pp.at32 <- ggpredict(allt3.2, "ZSAPA2[-2:2]")
# pp.at33 <- ggpredict(allt3.2, "ZSAPA3[-2:2]")
# pp.at34 <- ggpredict(allt3.2, "ZSem[-2:2]")
# pp.at35 <- ggpredict(allt3.2, "ZPhon[-2:2]")

# pp.at31
# pp.at32
# pp.at33
# pp.at34
# pp.at35

# Plot Results

# at.plot31 <-plot(pp.at31) + labs(title = "Treated Items", x = "CAT Comprehension z-Score", y
= "Predicted Probability")
# at.plot32 <-plot(pp.at32) + labs(title = "Treated Items", x = "SAPA 2 z-Score", y = "Predicted
Probability")
# at.plot33 <-plot(pp.at33) + labs(title = "Treated Items", x = "SAPA 3 z-Score", y = "Predicted
Probability")
# at.plot34 <-plot(pp.at34) + labs(title = "Treated Items", x = "Semantic Impairment z-Score", y
= "Predicted Probability")
# at.plot35 <-plot(pp.at35) + labs(title = "Treated Items", x = "Phonological Impairment z-
Score", y = "Predicted Probability")
```

```
# Table

# tab_model(allt3.2, allt3.null,
#   dv.labels = c("All Treated Follow-Up Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## Whole Group -- A3 Untreated-Related
################################################################################

## Untreated Regressions

# allu3.0 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No) + (1|Tx_Code), data = u3, family = binomial)

# allu3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = u3, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.allu3 <- influence(allu3.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_allu3.jpg")
# infIndexPlot(estex.allu3, id=list(method="y", n=5), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data

# u3.1 <- subset(u3, SubID != c("SF10PDX", "SF18SEA"))

# Re-analysis without influential data

# allu3.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = u3.1, family = binomial)

# allu3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = u3.1, family = binomial)

# Summary and Model Comparisons

# summary(allu3.2)
```

```
# anova(allu3.2, allu3.null)
# au3.models <- list(allu3.0, allu3.1, allu3.2, allu3.null)
# au3.model.names <- c("allu3.0", "allu3.1", "allu3.2", "allu3.null")
# aictab(cand.set = au3.models, modnames = au3.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(allu3.2)

# au3.predicted <- plogis(predict(allu3.2, u3.1))

# optCutOff <- optimalCutoff(u3.1$Response, au3.predicted)[1]

# sensitivity(u3.1$Response, au3.predicted, threshold = optCutOff)

# specificity(u3.1$Response, au3.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.au31 <- ggpredict(allu3.2, "ZCAT_T [-2:2]")
# pp.au32 <- ggpredict(allu3.2, "ZSAPA2[-2:2]")
# pp.au33 <- ggpredict(allu3.2, "ZSAPA3[-2:2]")
# pp.au34 <- ggpredict(allu3.2, "ZSem[-2:2]")
# pp.au35 <- ggpredict(allu3.2, "ZPhon[-2:2]")

# pp.au31
# pp.au32
# pp.au33
# pp.au34
# pp.au35

# Plot Results

# au.plot31 <-plot(pp.au31) + labs(title = "Untreated-Related Items", x = "CAT Comprehension
z-Score", y = "Predicted Probability")
# au.plot32 <-plot(pp.au32) + labs(title = "Untreated-Related Items", x = "SAPA 2 z-Score", y =
"Predicted Probability")
# au.plot33 <-plot(pp.au33) + labs(title = "Untreated-Related Items", x = "SAPA 3 z-Score", y =
"Predicted Probability")
# au.plot34 <-plot(pp.au34) + labs(title = "Untreated-Related Items", x = "Semantic Impairment
z-Score", y = "Predicted Probability")
# au.plot35 <-plot(pp.au35) + labs(title = "Untreated-Related Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table
```

```
# tab_model(allu3.2, allu3.null,
#   dv.labels = c("All Untreated-Related Follow-Up Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## Whole Group -- A3 Untreated-Unrelated
################################################################################

## Untreated Regressions

# allp3.0 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No) + (1|Tx_Code), data = p3, family = binomial)

# allp3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = p3, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.allp3 <- influence(allp3.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_allp3.jpg")
# infIndexPlot(estex.allp3, id=list(method="y", n=5), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data

# p3.1 <- subset(p3, SubID != c("PM12SEA", "SF04SEA"))

# Re-analysis without influential data

# allp3.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = p3.1, family = binomial)

# allp3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = p3.1, family = binomial)

# Summary and Model Comparisons

# summary(allp3.2)
# anova(allp3.2, allp3.null)
```

```
# ap3.models <- list(allp3.0, allp3.1, allp3.2, allp3.null)
# ap3.model.names <- c("allp3.0", "allp3.1", "allp3.2", "allp3.null")
# aictab(cand.set = ap3.models, modnames = ap3.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(allp3.2)

# ap3.predicted <- plogis(predict(allp3.2, p3.1))

# optCutOff <- optimalCutoff(p3.1$Response, ap3.predicted)[1]

# sensitivity(p3.1$Response, ap3.predicted, threshold = optCutOff)

# specificity(p3.1$Response, ap3.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.ap31 <- ggpredict(allp3.2, "ZCAT_T [-2:2]")
# pp.ap32 <- ggpredict(allp3.2, "ZSAPA2[-2:2]")
# pp.ap33 <- ggpredict(allp3.2, "ZSAPA3[-2:2]")
# pp.ap34 <- ggpredict(allp3.2, "ZSem[-2:2]")
# pp.ap35 <- ggpredict(allp3.2, "ZPhon[-2:2]")

# pp.ap31
# pp.ap32
# pp.ap33
# pp.ap34
# pp.ap35

# Plot Results

# ap.plot31 <-plot(pp.ap21) + labs(title = "Untreated-Unrelated Items", x = "CAT
Comprehension z-Score", y = "Predicted Probability")
# ap.plot32 <-plot(pp.ap22) + labs(title = "Untreated-Unrelated Items", x = "SAPA 2 z-Score", y
= "Predicted Probability")
# ap.plot33 <-plot(pp.ap23) + labs(title = "Untreated-Unrelated Items", x = "SAPA 3 z-Score", y
= "Predicted Probability")
# ap.plot34 <-plot(pp.ap24) + labs(title = "Untreated-Unrelated Items", x = "Semantic
Impairment z-Score", y = "Predicted Probability")
# ap.plot35 <-plot(pp.ap25) + labs(title = "Untreated-Unrelated Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table
```

```
# tab_model(allp3.2, allp3.null,
#   dv.labels = c("All Untreated-Unrelated Follow-Up Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
##############################################################################
## PMT Group -- A2 Data
##############################################################################
# a <- read.csv("Logistic_zPMT_capped1.csv")

# a <- a %>%
#   mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
#   na.omit(a)
# a <- subset(a, Time == "A2")
# head(a)
# summary(a)

# pt2 <- subset(a, Type_Key == "T")
# pu2 <- subset(a, Type_Key == "U")
# pp2 <- subset(a, Type_Key == "POM")

# head(pt2)
# summary(pt2)
# head(pu2)
# head(pp2)
##############################################################################
## PMT Group -- A2 Treated
##############################################################################

## Treated Regressions

# pmtt2.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pt2, family = binomial)

# summary(pmtt2.1)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.pmtt2 <- influence(pmtt2.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_pmtt2.jpg")
```

```
# infIndexPlot(estex.pmtt2, var = "Cook",id=list(method="y", n=5), main = NULL)
# dev.off()

# Exclude Influential Data

# pt2.1 <- subset(pt2, SubID != c("PM08SEA", "PM12PDX", "PMTSEA"))
# head(pt2.1)

# Re-analysis without influential data

# pmtt2.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pt2.1, family = binomial)

# pmtt2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = pt2.1, family = binomial)

# Summary and Model Comparisons

# summary(pmtt2.2)
# anova(pmtt2.2, pmtt2.null)
# pt2.models <- list(pmtt2.1, pmtt2.2, pmtt2.null)
# pt2.model.names <- c("pmtt2.1", "pmtt2.2", "pmtt2.null")
# aictab(cand.set = pt2.models, modnames = pt2.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(pmtt2.2)

# pt2.predicted <- plogis(predict(pmtt2.2, pt2.1))

# optCutOff <- optimalCutoff(pt2.1$Response, pt2.predicted)[1]

# sensitivity(pt2.1$Response, pt2.predicted, threshold = optCutOff)

# specificity(pt2.1$Response, pt2.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.pt21 <- ggpredict(pmtt2.2, "ZCAT_T [-2:2]")
# pp.pt22 <- ggpredict(pmtt2.2, "ZSAPA2[-2:2]")
# pp.pt23 <- ggpredict(pmtt2.2, "ZSAPA3[-2:2]")
# pp.pt24 <- ggpredict(pmtt2.2, "ZSem[-2:2]")
# pp.pt25 <- ggpredict(pmtt2.2, "ZPhon[-2:2]")

# pp.pt21
# pp.pt22
# pp.pt23
```

```
# pp.pt24
# pp.pt25

# Plot Results

# pt.plot21 <-plot(pp.pt21) + labs(title = "Treated Items", x = "CAT Comprehension z-Score", y
= "Predicted Probability")
# pt.plot22 <-plot(pp.pt22) + labs(title = "Treated Items", x = "SAPA 2 z-Score", y = "Predicted
Probability")
# pt.plot23 <-plot(pp.pt23) + labs(title = "Treated Items", x = "SAPA 3 z-Score", y = "Predicted
Probability")
# pt.plot24 <-plot(pp.pt24) + labs(title = "Treated Items", x = "Semantic Impairment z-Score", y
= "Predicted Probability")
# pt.plot25 <-plot(pp.pt25) + labs(title = "Treated Items", x = "Phonological Impairment z-
Score", y = "Predicted Probability")

# Table

# tab_model(pmtt2.2, pmtt2.null,
#   dv.labels = c("PMT Treated Immed. Post-Tx Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
##############################################################################
## PMT Group -- A2 Untreated-Related
##############################################################################

## Untreated Regressions

# pmtu2.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pu2, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.pmtu2 <- influence(pmtu2.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_pmtu2.jpg")
# infIndexPlot(estex.pmtu2, id=list(method="y", n=5), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data
```

```
# pu2.1 <- subset(pu2, SubID != c("PM14SEA", "PM15SEA", "PM17SEA"))

# Re-analysis without influential data

# pmtu2.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pu2.1, family = binomial)

# pmtu2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = pu2.1, family =
binomial)

# Summary and Model Comparisons

# summary(pmtu2.2)
# anova(pmtu2.2, pmtu2.null)
# pu2.models <- list(pmtu2.1, pmtu2.2, pmtu2.null)
# pu2.model.names <- c("pmtu2.1", "pmtu2.2", "pmtu2.null")
# aictab(cand.set = pu2.models, modnames = pu2.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(pmtu2.2)

# pu2.predicted <- plogis(predict(pmtu2.2, pu2.1))

# optCutOff <- optimalCutoff(pu2.1$Response, pu2.predicted)[1]

# sensitivity(pu2.1$Response, pu2.predicted, threshold = optCutOff)

# specificity(pu2.1$Response, pu2.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.pu21 <- ggpredict(pmtu2.2, "ZCAT_T [-2:2]")
# pp.pu22 <- ggpredict(pmtu2.2, "ZSAPA2[-2:2]")
# pp.pu23 <- ggpredict(pmtu2.2, "ZSAPA3[-2:2]")
# pp.pu24 <- ggpredict(pmtu2.2, "ZSem[-2:2]")
# pp.pu25 <- ggpredict(pmtu2.2, "ZPhon[-2:2]")

# pp.pu21
# pp.pu22
# pp.pu23
# pp.pu24
# pp.pu25

# Plot Results
```

```
# pu.plot21 <-plot(pp.pu21) + labs(title = "Untreated-Related Items", x = "CAT Comprehension z-Score", y = "Predicted Probability")
# pu.plot22 <-plot(pp.pu22) + labs(title = "Untreated-Related Items", x = "SAPA 2 z-Score", y = "Predicted Probability")
# pu.plot23 <-plot(pp.pu23) + labs(title = "Untreated-Related Items", x = "SAPA 3 z-Score", y = "Predicted Probability")
# pu.plot24 <-plot(pp.pu24) + labs(title = "Untreated-Related Items", x = "Semantic Impairment z-Score", y = "Predicted Probability")
# pu.plot25 <-plot(pp.pu25) + labs(title = "Untreated-Related Items", x = "Phonological Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(pmtu2.2, pmtu2.null,
#   dv.labels = c("PMT Untreated-Related Immed. Post-Tx Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment", "Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## PMT Group -- A2 Untreated-Unrelated
################################################################################

## Untreated Regressions

# pmtp2.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) + (1|Item_No), data = pp2, family = binomial)

# summary(pmtp2.1)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.pmtp2 <- influence(pmtp2.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_pmtp2.jpg")
# infIndexPlot(estex.pmtp2, id=list(method="y", n=4), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data
```

```
# pp2.1 <- subset(pp2, SubID != c("PM01PDX", "PM12SEA", "PM15SEA"))

# Re-analysis without influential data

# pmtp2.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pp2.1, family = binomial)

# pmtp2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = pp2.1, family =
binomial)

# Summary and Model Comparisons

# summary(pmtp2.2)
# anova(pmtp2.2, pmtp2.null)
# pp2.models <- list(pmtp2.1, pmtp2.2, pmtp2.null)
# pp2.model.names <- c("pmtp2.1", "pmtp2.2", "pmtp2.null")
# aictab(cand.set = pp2.models, modnames = pp2.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(pmtp2.2)

# pp2.predicted <- plogis(predict(pmtp2.2, pp2.1))

# optCutOff <- optimalCutoff(pp2.1$Response, pp2.predicted)[1]

# sensitivity(pp2.1$Response, pp2.predicted, threshold = optCutOff)

# specificity(pp2.1$Response, pp2.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.pp21 <- ggpredict(pmtp2.2, "ZCAT_T [-2:2]")
# pp.pp22 <- ggpredict(pmtp2.2, "ZSAPA2[-2:2]")
# pp.pp23 <- ggpredict(pmtp2.2, "ZSAPA3[-2:2]")
# pp.pp24 <- ggpredict(pmtp2.2, "ZSem[-2:2]")
# pp.pp25 <- ggpredict(pmtp2.2, "ZPhon[-2:2]")

# pp.pp21
# pp.pp22
# pp.pp23
# pp.pp24
# pp.pp25

# Plot Results
```

```
# pp.plot21 <-plot(pp.pp21) + labs(title = "Untreated-Unrelated Items", x = "CAT
Comprehension z-Score", y = "Predicted Probability")
# pp.plot22 <-plot(pp.pp22) + labs(title = "Untreated-Unrelated Items", x = "SAPA 2 z-Score", y
= "Predicted Probability")
# pp.plot23 <-plot(pp.pp23) + labs(title = "Untreated-Unrelated Items", x = "SAPA 3 z-Score", y
= "Predicted Probability")
# pp.plot24 <-plot(pp.pp24) + labs(title = "Untreated-Unrelated Items", x = "Semantic
Impairment z-Score", y = "Predicted Probability")
# pp.plot25 <-plot(pp.pp25) + labs(title = "Untreated-Unrelated Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(pmtp2.2, pmtp2.null,
#   dv.labels = c("PMT Untreated-Unrelated Immed. Post-Tx Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#  string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## PMT Group -- A3 Data
################################################################################
# b <- read.csv("Logistic_zPMT_capped1.csv")
# head(b)

# b <- b %>%
#   mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
#   na.omit(b)
# b <- subset(b, Time == "A3")
# head(b)
# summary(b)

# pt3 <- subset(b, Type_Key == "T")
# pu3 <- subset(b, Type_Key == "U")
# pp3 <- subset(b, Type_Key == "POM")

# head(pt3)
# head(pu3)
# head(pp3)
################################################################################
## PMT Group -- A3 Treated
################################################################################

## Treated Regressions
```

```
# pmtt3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pt3, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.pmtt3 <- influence(pmtt3.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_pmtt3.jpg")
# infIndexPlot(estex.pmtt3, var = "Cook",id=list(method="y", n=5), main = NULL)
# dev.off()

# Exclude Influential Data

# pt3.1 <- subset(pt3, SubID != c("PM12PDX", "PM15SEA", "PM16SEA", "PM17SEA"))

# Re-analysis without influential data

# pmtt3.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pt3.1, family = binomial)

# pmtt3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = pt3.1, family = binomial)

# Summary and Model Comparisons

# summary(pmtt3.2)
# anova(pmtt3.2, pmtt3.null)
# pt3.models <- list(pmtt3.1, pmtt3.2, pmtt3.null)
# pt3.model.names <- c("pmtt3.1", "pmtt3.2", "pmtt3.null")
# aictab(cand.set = pt3.models, modnames = pt3.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(pmtt3.2)

# pt3.predicted <- plogis(predict(pmtt3.2, pt3.1))

# optCutOff <- optimalCutoff(pt3.1$Response, pt3.predicted)[1]

# sensitivity(pt3.1$Response, pt3.predicted, threshold = optCutOff)

# specificity(pt3.1$Response, pt3.predicted, threshold = optCutOff)
```

```
# Predicted Probabilities

# pp.pt31 <- ggpredict(pmtt3.2, "ZCAT_T [-2:2]")
# pp.pt32 <- ggpredict(pmtt3.2, "ZSAPA2[-2:2]")
# pp.pt33 <- ggpredict(pmtt3.2, "ZSAPA3[-2:2]")
# pp.pt34 <- ggpredict(pmtt3.2, "ZSem[-2:2]")
# pp.pt35 <- ggpredict(pmtt3.2, "ZPhon[-2:2]")

# pp.pt31
# pp.pt32
# pp.pt33
# pp.pt34
# pp.pt35

# Plot Results

# pt.plot31 <-plot(pp.pt31) + labs(title = "Treated Items", x = "CAT Comprehension z-Score", y
= "Predicted Probability")
# pt.plot32 <-plot(pp.pt32) + labs(title = "Treated Items", x = "SAPA 2 z-Score", y = "Predicted
Probability")
# pt.plot33 <-plot(pp.pt33) + labs(title = "Treated Items", x = "SAPA 3 z-Score", y = "Predicted
Probability")
# pt.plot34 <-plot(pp.pt34) + labs(title = "Treated Items", x = "Semantic Impairment z-Score", y
= "Predicted Probability")
# pt.plot35 <-plot(pp.pt35) + labs(title = "Treated Items", x = "Phonological Impairment z-
Score", y = "Predicted Probability")

# Table

# tab_model(pmtt3.2, pmtt3.null,
#   dv.labels = c("PMT Treated Follow-Up Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
 #  string.p = "p-Value")
################################################################################
## PMT Group -- A3 Untreated-Related
################################################################################

## Untreated Regressions

# pmtu3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pu3, family = binomial)

## Identify and Remove Influential Data
```

```
# Identify Infuential Data
# estex.pmtu3 <- influence(pmtu3.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_pmtu3.jpg")
# infIndexPlot(estex.pmtu3, id=list(method="y", n=5), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data

# pu3.1 <- subset(pu3, SubID != c("PM12PDX", "PM15SEA", "PM17SEA"))

# Re-analysis without influential data

# pmtu3.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pu3.1, family = binomial)

# pmtu3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = pu3.1, family =
binomial)

# Summary and Model Comparisons

# summary(pmtu3.2)
# anova(pmtu3.2, pmtu3.null)
# pu3.models <- list(pmtu3.1, pmtu3.2, pmtu3.null)
# pu3.model.names <- c("pmtu3.1", "pmtu3.2", "pmtu3.null")
# aictab(cand.set = pu3.models, modnames = pu3.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(pmtu3.2)

# pu3.predicted <- plogis(predict(pmtu3.2, pu3.1))

# optCutOff <- optimalCutoff(pu3.1$Response, pu3.predicted)[1]

# sensitivity(pu3.1$Response, pu3.predicted, threshold = optCutOff)

# specificity(pu3.1$Response, pu3.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.pu31 <- ggpredict(pmtu3.2, "ZCAT_T [-2:2]")
# pp.pu32 <- ggpredict(pmtu3.2, "ZSAPA2[-2:2]")
```

```
# pp.pu33 <- ggpredict(pmtu3.2, "ZSAPA3[-2:2]")
# pp.pu34 <- ggpredict(pmtu3.2, "ZSem[-2:2]")
# pp.pu35 <- ggpredict(pmtu3.2, "ZPhon[-2:2]")

# pp.pu31
# pp.pu32
# pp.pu33
# pp.pu34
# pp.pu35

# Plot Results

# pu.plot31 <-plot(pp.pu31) + labs(title = "Untreated-Related Items", x = "CAT Comprehension
z-Score", y = "Predicted Probability")
# pu.plot32 <-plot(pp.pu32) + labs(title = "Untreated-Related Items", x = "SAPA 2 z-Score", y =
"Predicted Probability")
# pu.plot33 <-plot(pp.pu33) + labs(title = "Untreated-Related Items", x = "SAPA 3 z-Score", y =
"Predicted Probability")
# pu.plot34 <-plot(pp.pu34) + labs(title = "Untreated-Related Items", x = "Semantic Impairment
z-Score", y = "Predicted Probability")
# pu.plot35 <-plot(pp.pu35) + labs(title = "Untreated-Related Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(pmtu3.2, pmtu3.null,
#   dv.labels = c("PMT Untreated-Related Follow-Up Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## PMT Group -- A3 Untreated-Unrelated
################################################################################

## POM Regressions

# pmtp3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pp3, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.pmtp3 <- influence(pmtp3.1, "SubID")
```

```
# Cook's Distance

# jpeg("CooksDistance_pmtp3.jpg")
# infIndexPlot(estex.pmtp3, id=list(method="y", n=5), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data

# pp3.1 <- subset(pp3, SubID != c("PM01PDX", "PM17SEA"))

# Re-analysis without influential data

# pmtp3.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = pp3.1, family = binomial)

# pmtp3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = pp3.1, family =
binomial)

# Summary and Model Comparisons

# summary(pmtp3.2)
# anova(pmtp3.2, pmtp3.null)
# pp3.models <- list(pmtp3.1, pmtp3.2, pmtp3.null)
# pp3.model.names <- c("pmtp3.1", "pmtp3.2", "pmtp3.null")
# aictab(cand.set = pp3.models, modnames = pp3.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(pmtp3.2)

# pp3.predicted <- plogis(predict(pmtp3.2, pp3.1))

# optCutOff <- optimalCutoff(pp3.1$Response, pp3.predicted)[1]

# sensitivity(pp3.1$Response, pp3.predicted, threshold = optCutOff)

# specificity(pp3.1$Response, pp3.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.pp31 <- ggpredict(pmtp3.2, "ZCAT_T [-2:2]")
# pp.pp32 <- ggpredict(pmtp3.2, "ZSAPA2[-2:2]")
# pp.pp33 <- ggpredict(pmtp3.2, "ZSAPA3[-2:2]")
# pp.pp34 <- ggpredict(pmtp3.2, "ZSem[-2:2]")
# pp.pp35 <- ggpredict(pmtp3.2, "ZPhon[-2:2]")
```

```
# pp.pp31
# pp.pp32
# pp.pp33
# pp.pp34
# pp.pp35

# Plot Results

# pp.plot31 <-plot(pp.pp21) + labs(title = "Untreated-Unrelated Items", x = "CAT
Comprehension z-Score", y = "Predicted Probability")
# pp.plot32 <-plot(pp.pp22) + labs(title = "Untreated-Unrelated Items", x = "SAPA 2 z-Score", y
= "Predicted Probability")
# pp.plot33 <-plot(pp.pp23) + labs(title = "Untreated-Unrelated Items", x = "SAPA 3 z-Score", y
= "Predicted Probability")
# pp.plot34 <-plot(pp.pp24) + labs(title = "Untreated-Unrelated Items", x = "Semantic
Impairment z-Score", y = "Predicted Probability")
# pp.plot35 <-plot(pp.pp25) + labs(title = "Untreated-Unrelated Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(pmtp3.2, pmtp3.null,
#   dv.labels = c("PMT Untreated-Unrelated Follow-Up Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
##############################################################################
## SFA Group -- A2 Data
##############################################################################
# c <- read.csv("Logistic_zSFA_capped1.csv")

# c <- c %>%
#   mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
#   na.omit(c)
# c <- subset(c, Time == "A2")
# head(c)
# summary(c)

# st2 <- subset(c, Type_Key == "T")
# su2 <- subset(c, Type_Key == "U")
# sp2 <- subset(c, Type_Key == "POM")
```

```
# head(st2)
# head(su2)
# head(sp2)
##############################################################################
## SFA Group -- A2 Treated
##############################################################################

## Treated Regressions

# sfat2.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = st2, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.sfat2 <- influence(sfat2.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_sfat2.jpg")
# infIndexPlot(estex.sfat2, var = "Cook",id=list(method="y", n=5), main = NULL)
# dev.off()

# Exclude Influential Data

# st2.1 <- subset(st2, SubID != c("SF07PDX", "SF10PDX"))

# Re-analysis without influential data

# sfat2.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = st2.1, family = binomial)

# sfat2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = st2.1, family = binomial)

# Summary and Model Comparisons

# summary(sfat2.2)
# anova(sfat2.2, sfat2.null)
# st2.models <- list(sfat2.1, sfat2.2, sfat2.null)
# st2.model.names <- c("sfat2.1", "sfat2.2", "sfat2.null")
# aictab(cand.set = st2.models, modnames = st2.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(sfat2.2)
```

```
# st2.predicted <- plogis(predict(sfat2.2, st2.1))

# optCutOff <- optimalCutoff(st2.1$Response, st2.predicted)[1]

# sensitivity(st2.1$Response, st2.predicted, threshold = optCutOff)

# specificity(st2.1$Response, st2.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.st21 <- ggpredict(sfat2.2, "ZCAT_T [-2:2]")
# pp.st22 <- ggpredict(sfat2.2, "ZSAPA2[-2:2]")
# pp.st23 <- ggpredict(sfat2.2, "ZSAPA3[-2:2]")
# pp.st24 <- ggpredict(sfat2.2, "ZSem[-2:2]")
# pp.st25 <- ggpredict(sfat2.2, "ZPhon[-2:2]")

# pp.st21
# pp.st22
# pp.st23
# pp.st24
# pp.st25

# Plot Results

# st.plot21 <-plot(pp.pt21) + labs(title = "Treated Items", x = "CAT Comprehension z-Score", y
= "Predicted Probability")
# st.plot22 <-plot(pp.pt22) + labs(title = "Treated Items", x = "SAPA 2 z-Score", y = "Predicted
Probability")
# st.plot23 <-plot(pp.pt23) + labs(title = "Treated Items", x = "SAPA 3 z-Score", y = "Predicted
Probability")
# st.plot24 <-plot(pp.pt24) + labs(title = "Treated Items", x = "Semantic Impairment z-Score", y
= "Predicted Probability")
# st.plot25 <-plot(pp.pt25) + labs(title = "Treated Items", x = "Phonological Impairment z-
Score", y = "Predicted Probability")

# Table

# tab_model(sfat2.2, sfat2.null,
#  dv.labels = c("SFA Treated Immed. Post-Tx Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## SFA Group -- A2 Untreated-Related
```

```
################################################################################

## Untreated Regressions

# sfau2.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = su2, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.sfau2 <- influence(sfau2.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_sfau2.jpg")
# infIndexPlot(estex.sfau2, id=list(method="y", n=5), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data

# su2.1 <- subset(su2, SubID != c("SF10PDX"))

# Re-analysis without influential data

# sfau2.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = su2.1, family = binomial)

# sfau2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = su2.1, family = binomial)

# Summary and Model Comparisons

# summary(sfau2.2)
# anova(sfau2.2, sfau2.null)
# su2.models <- list(sfau2.1, sfau2.2, sfau2.null)
# su2.model.names <- c("sfau2.1", "sfau2.2", "sfau2.null")
# aictab(cand.set = su2.models, modnames = su2.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(sfau2.2)

# su2.predicted <- plogis(predict(sfau2.2, su2.1))

# optCutOff <- optimalCutoff(su2.1$Response, su2.predicted)[1]

# sensitivity(su2.1$Response, su2.predicted, threshold = optCutOff)
```

```
# specificity(su2.1$Response, su2.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.su21 <- ggpredict(sfau2.2, "ZCAT_T [-2:2]")
# pp.su22 <- ggpredict(sfau2.2, "ZSAPA2[-2:2]")
# pp.su23 <- ggpredict(sfau2.2, "ZSAPA3[-2:2]")
# pp.su24 <- ggpredict(sfau2.2, "ZSem[-2:2]")
# pp.su25 <- ggpredict(sfau2.2, "ZPhon[-2:2]")

# pp.su21
# pp.su22
# pp.su23
# pp.su24
# pp.su25

# Plot Results

# su.plot21 <-plot(pp.su21) + labs(title = "Untreated-Related Items", x = "CAT Comprehension
z-Score", y = "Predicted Probability")
# su.plot22 <-plot(pp.su22) + labs(title = "Untreated-Related Items", x = "SAPA 2 z-Score", y =
"Predicted Probability")
# su.plot23 <-plot(pp.su23) + labs(title = "Untreated-Related Items", x = "SAPA 3 z-Score", y =
"Predicted Probability")
# su.plot24 <-plot(pp.su24) + labs(title = "Untreated-Related Items", x = "Semantic Impairment
z-Score", y = "Predicted Probability")
# su.plot25 <-plot(pp.su25) + labs(title = "Untreated-Related Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(sfau2.2, sfau2.null,
#   dv.labels = c("SFA Untreated-Related Immed. Post-Tx Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
##############################################################################
## SFA Group -- A2 Untreated-Unrelated
##############################################################################

## POM Regressions
```

```
# sfap2.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = sp2, family = binomial)

# summary(sfap2.1)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.sfap2 <- influence(sfap2.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_sfap2.jpg")
# infIndexPlot(estex.sfap2, id=list(method="y", n=4), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data

# sp2.1 <- subset(sp2, SubID != c("SF04PDX", "SF07PDX", "SF10PDX", "SF18SEA"))

# Re-analysis without influential data

# sfap2.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = sp2.1, family = binomial)

# sfap2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = sp2.1, family = binomial)

# Summary and Model Comparisons

# summary(sfap2.2)
# anova(sfap2.2, sfap2.null)
# sp2.models <- list(sfap2.1, sfap2.2, sfap2.null)
# sp2.model.names <- c("sfap2.1", "sfap2.2", "sfap2.null")
# aictab(cand.set = sp2.models, modnames = sp2.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(sfap2.2)

# sp2.predicted <- plogis(predict(sfap2.2, sp2.1))

# optCutOff <- optimalCutoff(sp2.1$Response, sp2.predicted)[1]

# sensitivity(sp2.1$Response, sp2.predicted, threshold = optCutOff)

# specificity(sp2.1$Response, sp2.predicted, threshold = optCutOff)
```

```
# Predicted Probabilities

# pp.sp21 <- ggpredict(sfap2.2, "ZCAT_T [-2:2]")
# pp.sp22 <- ggpredict(sfap2.2, "ZSAPA2[-2:2]")
# pp.sp23 <- ggpredict(sfap2.2, "ZSAPA3[-2:2]")
# pp.sp24 <- ggpredict(sfap2.2, "ZSem[-2:2]")
# pp.sp25 <- ggpredict(sfap2.2, "ZPhon[-2:2]")

# pp.sp21
# pp.sp22
# pp.sp23
# pp.sp24
# pp.sp25

# Plot Results

# sp.plot21 <-plot(pp.sp21) + labs(title = "Untreated-Unrelated Items", x = "CAT
Comprehension z-Score", y = "Predicted Probability")
# sp.plot22 <-plot(pp.sp22) + labs(title = "Untreated-Unrelated Items", x = "SAPA 2 z-Score", y
= "Predicted Probability")
# sp.plot23 <-plot(pp.sp23) + labs(title = "Untreated-Unrelated Items", x = "SAPA 3 z-Score", y
= "Predicted Probability")
# sp.plot24 <-plot(pp.sp24) + labs(title = "Untreated-Unrelated Items", x = "Semantic
Impairment z-Score", y = "Predicted Probability")
# sp.plot25 <-plot(pp.sp25) + labs(title = "Untreated-Unrelated Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(sfap2.2, sfap2.null,
#   dv.labels = c("SFA Untreated-Unrelated Immed. Post-Tx Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## SFA Group -- A3 Data

# d <- read.csv("Logistic_zSFA_capped1.csv")
# head(d)

# d <- d %>%
#   mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
```

```
#   na.omit(d)
# d <- subset(d, Time == "A3")
# head(d)
# summary(d)

# st3 <- subset(d, Type_Key == "T")
# su3 <- subset(d, Type_Key == "U")
# sp3 <- subset(d, Type_Key == "POM")

# head(st3)
# head(su3)
# head(sp3)
################################################################################
## SFA Group -- A3 Treated
################################################################################

## Treated Regressions

# sfat3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = st3, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.sfat3 <- influence(sfat3.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_sfat3.jpg")
# infIndexPlot(estex.sfat3, var = "Cook",id=list(method="y", n=5), main = NULL)
# dev.off()

# Exclude Influential Data

# st3.1 <- subset(st3, SubID != c("SF07PDX", "SF10PDX", "SF12SEA"))

# Re-analysis without influential data

# sfat3.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = st3.1, family = binomial)

# sfat3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = st3.1, family = binomial)

# Summary and Model Comparisons

# summary(sfat3.2)
```

```
# anova(sfat3.2, sfat3.null)
# st3.models <- list(sfat3.1, sfat3.2, sfat3.null)
# st3.model.names <- c("sfat3.1", "sfat3.2", "sfat3.null")
# aictab(cand.set = st3.models, modnames = st3.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(sfat3.2)

# st3.predicted <- plogis(predict(sfat3.2, st3.1))

# optCutOff <- optimalCutoff(st3.1$Response, st3.predicted)[1]

# sensitivity(st3.1$Response, st3.predicted, threshold = optCutOff)

# specificity(st3.1$Response, st3.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.st31 <- ggpredict(sfat3.2, "ZCAT_T [-2:2]")
# pp.st32 <- ggpredict(sfat3.2, "ZSAPA2[-2:2]")
# pp.st33 <- ggpredict(sfat3.2, "ZSAPA3[-2:2]")
# pp.st34 <- ggpredict(sfat3.2, "ZSem[-2:2]")
# pp.st35 <- ggpredict(sfat3.2, "ZPhon[-2:2]")

# pp.st31
# pp.st32
# pp.st33
# pp.st34
# pp.st35

# Plot Results

# st.plot31 <-plot(pp.st31) + labs(title = "Treated Items", x = "CAT Comprehension z-Score", y
= "Predicted Probability")
# st.plot32 <-plot(pp.st32) + labs(title = "Treated Items", x = "SAPA 2 z-Score", y = "Predicted
Probability")
# st.plot33 <-plot(pp.st33) + labs(title = "Treated Items", x = "SAPA 3 z-Score", y = "Predicted
Probability")
# st.plot34 <-plot(pp.st34) + labs(title = "Treated Items", x = "Semantic Impairment z-Score", y
= "Predicted Probability")
# st.plot35 <-plot(pp.st35) + labs(title = "Treated Items", x = "Phonological Impairment z-
Score", y = "Predicted Probability")

# Table
```

```
# tab_model(sfat3.2, sfat3.null,
#   dv.labels = c("SFA Treated Follow-Up Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## SFA Group -- A3 Untreated-Related
################################################################################

## Untreated Regressions

# sfau3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = su3, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.sfau3 <- influence(sfau3.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_sfau3.jpg")
# infIndexPlot(estex.sfau3, id=list(method="y", n=5), var = "Cook", main = NULL)
# dev.off()

# Exclude Influential Data

# su3.1 <- subset(su3, SubID != c("SF04SEA", "SF07PDX", "SF10PDX", "SF18SEA"))

# Re-analysis without influential data

# sfau3.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon +
(1|SubID) + (1|Item_No), data = su3.1, family = binomial)

# sfau3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = su3.1, family = binomial)

# Summary and Model Comparisons

# summary(sfau3.2)
# anova(sfau3.2, sfau3.null)
# su3.models <- list(sfau3.1, sfau3.2, sfau3.null)
# su3.model.names <- c("sfau3.1", "sfau3.2", "sfau3.null")
# aictab(cand.set = su3.models, modnames = su3.model.names)
```

```
# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(sfau3.2)

# su3.predicted <- plogis(predict(sfau3.2, su3.1))

# optCutOff <- optimalCutoff(su3.1$Response, su3.predicted)[1]

# sensitivity(su3.1$Response, su3.predicted, threshold = optCutOff)

# specificity(su3.1$Response, su3.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.su31 <- ggpredict(sfau3.2, "ZCAT_T [-2:2]")
# pp.su32 <- ggpredict(sfau3.2, "ZSAPA2[-2:2]")
# pp.su33 <- ggpredict(sfau3.2, "ZSAPA3[-2:2]")
# pp.su34 <- ggpredict(sfau3.2, "ZSem[-2:2]")
# pp.su35 <- ggpredict(sfau3.2, "ZPhon[-2:2]")

# pp.su31
# pp.su32
# pp.su33
# pp.su34
# pp.su35

# Plot Results

# su.plot31 <-plot(pp.su31) + labs(title = "Untreated-Related Items", x = "CAT Comprehension
z-Score", y = "Predicted Probability")
# su.plot32 <-plot(pp.su32) + labs(title = "Untreated-Related Items", x = "SAPA 2 z-Score", y =
"Predicted Probability")
# su.plot33 <-plot(pp.su33) + labs(title = "Untreated-Related Items", x = "SAPA 3 z-Score", y =
"Predicted Probability")
# su.plot34 <-plot(pp.su34) + labs(title = "Untreated-Related Items", x = "Semantic Impairment
z-Score", y = "Predicted Probability")
# su.plot35 <-plot(pp.su35) + labs(title = "Untreated-Related Items", x = "Phonological
Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(sfau3.2, sfau3.null,
#   dv.labels = c("SFA Untreated-Related Follow-Up Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic Impairment",
```

```
"Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
################################################################################
## SFA Group -- A3 Untreated-Unrelated
################################################################################

## POM Regressions

# sfap3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (
1|SubID) + (1|Item_No), data = sp3, family = binomial)

## Identify and Remove Influential Data

# Identify Infuential Data
# estex.sfap3 <- influence(sfap3.1, "SubID")

# Cook's Distance

# jpeg("CooksDistance_sfap3.jpg")
# infIndexPlot(estex.sfap3, id=list(method="y", n=5), var = "Cook", main = NU
LL)
# dev.off()

# Exclude Influential Data

# sp3.1 <- subset(sp3, SubID != c("SF05SEA", "SF09SEA", "SF10SEA"))

# Re-analysis without influential data

# sfap3.2 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (
1|SubID) + (1|Item_No), data = sp3.1, family = binomial)

# sfap3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = sp3.1, f
amily = binomial)

# Summary and Model Comparisons

# summary(sfap3.2)
# anova(sfap3.2, sfap3.null)
# sp3.models <- list(sfap3.1, sfap3.2, sfap3.null)
# sp3.model.names <- c("sfap3.1", "sfap3.2", "sfap3.null")
# aictab(cand.set = sp3.models, modnames = sp3.model.names)

# Model Measures of Multicollinearity, Sensitivity & Specificity

# vif(sfap3.2)

# sp3.predicted <- plogis(predict(sfap3.2, sp3.1))
```

```
# optCutOff <- optimalCutoff(sp3.1$Response, sp3.predicted)[1]

# sensitivity(sp3.1$Response, sp3.predicted, threshold = optCutOff)

# specificity(sp3.1$Response, sp3.predicted, threshold = optCutOff)

# Predicted Probabilities

# pp.sp31 <- ggpredict(sfap3.2, "ZCAT_T [-2:2]")
# pp.sp32 <- ggpredict(sfap3.2, "ZSAPA2[-2:2]")
# pp.sp33 <- ggpredict(sfap3.2, "ZSAPA3[-2:2]")
# pp.sp34 <- ggpredict(sfap3.2, "ZSem[-2:2]")
# pp.sp35 <- ggpredict(sfap3.2, "ZPhon[-2:2]")

# pp.sp31
# pp.sp32
# pp.sp33
# pp.sp34
# pp.sp35

# Plot Results

# sp.plot31 <-plot(pp.sp21) + labs(title = "Untreated-Unrelated Items", x = "
CAT Comprehension z-Score", y = "Predicted Probability")
# sp.plot32 <-plot(pp.sp22) + labs(title = "Untreated-Unrelated Items", x = "
SAPA 2 z-Score", y = "Predicted Probability")
# sp.plot33 <-plot(pp.sp23) + labs(title = "Untreated-Unrelated Items", x = "
SAPA 3 z-Score", y = "Predicted Probability")
# sp.plot34 <-plot(pp.sp24) + labs(title = "Untreated-Unrelated Items", x = "
Semantic Impairment z-Score", y = "Predicted Probability")
# sp.plot35 <-plot(pp.sp25) + labs(title = "Untreated-Unrelated Items", x = "
Phonological Impairment z-Score", y = "Predicted Probability")


# Table

# tab_model(sfap3.2, sfap3.null,
#   dv.labels = c("SFA Untreated-Unrelated Follow-Up Model", "Null Model"),
#   string.pred = "Coeffcient",
#   pred.labels = c("Intercept", "CAT T-score", "SAPA 2", "SAPA 3", "Semantic
Impairment", "Phonological Impairment"),
#   string.ci = "CI (95%)",
#   string.p = "p-Value")
```

```
##############################################################################
## SAPA 3 POST HOC
##############################################################################
```

```
##############################################################################
# Whole Group -- A2 Data
##############################################################################z
<- read.csv("Logistic_zAll_capped1.csv")

data <- z %>%
  mutate_at(vars(c(Tx_Code, Time_Code, Item_No)), list(factor)) %>%
  na.omit(z)
data <- subset(data, Time == "A2")

at2 <- subset(data, Type_Key == "T")
au2 <- subset(data, Type_Key == "U")
ap2 <- subset(data, Type_Key == "POM")


##############################################################################
# Whole Group SAPA 3 -- A2 Treated
##############################################################################
## Treated Regressions

atsap2 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = at2, family = binomial)

# Exclude Influential Data

at2.1 <- subset(at2, SubID != c("PM08SEA", "PM17SEA", "SF02PDX"))
## Warning in `!=.default`(SubID, c("PM08SEA", "PM17SEA", "SF02PDX")): longer
## object length is not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

atsap2.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1 | SubID) + (1 | Item_No), data = at2.1, family = binomial)

atsap2.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = at2.1, family = binomial)
##############################################################################
# Whole Group SAPA 3-- A2 Untreated-Related
##############################################################################
## Untreated Regressions

ausap2 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = au2, family = binomial)

# Exclude Influential Data
```

```
au2.1 <- subset(au2, SubID != c("PM08SEA", "PM17SEA", "SF02PDX", "SF06PDX"))

# Re-analysis without influential data

ausap2.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1|SubID) + (1|Item_No), data = au2.1, family = binomial)

ausap2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = au2.1, family = binomial)
##############################################################################
# Whole Group SAPA 3 -- A2 Untreated-Unrelated
##############################################################################
## POM Regressions

apsap2 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1|SubID) + (1|Item_No), data = ap2, family = binomial)

# Exclude Influential Data

ap2.1 <- subset(ap2, SubID != c("PM01SEA", "PM12SEA", "PM17SEA"))

# Re-analysis without influential data

apsap2.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1|SubID) + (1|Item_No), data = ap2.1, family = binomial)

apsap2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = ap2.1, family = binomial)
##############################################################################
# Whole Group -- A3 Data
##############################################################################y
<- read.csv("Logistic_zAll_capped1.csv")

data1 <- y %>%
  mutate_at(vars(c(Tx_Code, Time_Code, Item_No)), list(factor)) %>%
  na.omit(y)
data1 <- subset(data1, Time == "A3")

at3 <- subset(data1, Type_Key == "T")
au3 <- subset(data1, Type_Key == "U")
ap3 <- subset(data1, Type_Key == "POM")
##############################################################################
# Whole Group SAPA 3 -- A3 Treated
##############################################################################
## Treated Regressions

atsap3 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1|SubID) + (1|Item_No), data = at3, family = binomial)
```

```
# Exclude Influential Data

at3.1 <- subset(at3, SubID != c("PM08SEA", "PM17SEA", "SF02PDX", "SF05SEA"))
## Warning in `!=.default`(SubID, c("PM08SEA", "PM17SEA", "SF02PDX", "SF05SEA")):
## longer object length is not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

atsap3.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1 | SubID) + (1 | Item_No), data = at3.1, family = binomial)

atsap3.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = at3.1, family = binomial)
################################################################################
# Whole Group SAPA 3 -- A3 Untreated-Related
################################################################################
## Untreated Regressions

ausap3 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = au3, family = binomial)

# Exclude Influential Data

au3.1 <- subset(au3, SubID != c("PM08SEA", "PM17SEA", "SF02PDX", "SF05SEA"))
## Warning in `!=.default`(SubID, c("PM08SEA", "PM17SEA", "SF02PDX", "SF05SEA")):
## longer object length is not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

ausap3.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1 | SubID) + (1 | Item_No), data = au3.1, family = binomial)

ausap3.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = au3.1, family = binomial)
################################################################################
# Whole Group SAPA 3 -- A3 Untreated-Unrelated
################################################################################
## POM Regressions

apsap3 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = ap3, family = binomial)

# Exclude Influential Data

ap3.1 <- subset(ap3, SubID != c("PM01SEA", "PM12SEA", "PM17SEA", "SF18SEA"))
```

```
## Warning in `!=.default`(SubID, c("PM01SEA", "PM12SEA", "PM17SEA", "SF18SEA")):
## longer object length is not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

apsap3.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1|SubID) + (1|Item_No), data = ap3.1, family = binomial)

apsap3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = ap3.1, family = binomial)
################################################################################
# PMT Group -- A2 Data
################################################################################a
<- read.csv("Logistic_zPMT_capped1.csv")

a <- a %>%
  mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
  na.omit(a)
a <- subset(a, Time == "A2")

pt2 <- subset(a, Type_Key == "T")
pu2 <- subset(a, Type_Key == "U")
pp2 <- subset(a, Type_Key == "POM")
################################################################################
# PMT Group SAPA 3 -- A2 Treated
################################################################################
## Treated Regressions

ptsap2 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1|SubID) + (1|Item_No), data = pt2, family = binomial)

# Exclude Influential Data

pt2.1 <- subset(pt2, SubID != c("PM06SEA", "PM08SEA", "PM17SEA"))

# Re-analysis without influential data

ptsap2.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1|SubID) + (1|Item_No), data = pt2.1, family = binomial)

ptsap2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = pt2.1, family = binomial)
################################################################################
# PMT Group SAPA 3 -- A2 Untreated-Related
################################################################################
## Untreated Regressions
```

```
pusap2 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = pu2, family = binomial)

# Exclude Influential Data

pu2.1 <- subset(pu2, SubID != c("PM06SEA","PM08SEA", "PM14SEA", "PM17SEA"))

# Re-analysis without influential data

pusap2.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1 | SubID) + (1 | Item_No), data = pu2.1, family = binomial)

pusap2.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = pu2.1, family =
binomial)
################################################################################
# PMT Group SAPA 3 -- A2 Untreated-Unrelated
################################################################################
## POM Regressions

ppsap2 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = pp2, family = binomial)

# Exclude Influential Data

pp2.1 <- subset(pp2, SubID != c("PM01PDX", "PM12SEA", "PM17SEA"))
## Warning in `!=.default`(SubID, c("PM01PDX", "PM12SEA", "PM17SEA")): longer
## object length is not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

ppsap2.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1 | SubID) + (1 | Item_No), data = pp2.1, family = binomial)

ppsap2.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = pp2.1, family =
binomial)
################################################################################
# PMT Group -- A3 Data
################################################################################b
<- read.csv("Logistic_zPMT_capped1.csv")

b <- b %>%
  mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
  na.omit(b)
b <- subset(b, Time == "A3")
```

```
pt3 <- subset(b, Type_Key == "T")
pu3 <- subset(b, Type_Key == "U")
pp3 <- subset(b, Type_Key == "POM")
##############################################################################
# PMT Group SAPA 3 -- A3 Treated
##############################################################################
## Treated Regressions

ptsap3 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = pt3, family = binomial)

# Exclude Influential Data

pt3.1 <- subset(pt3, SubID != c("PM06SEA", "PM08SEA", "PM17SEA"))

# Re-analysis without influential data

ptsap3.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1 | SubID) + (1 | Item_No), data = pt3.1, family = binomial)

ptsap3.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = pt3.1, family = binomial)
##############################################################################
# PMT Group SAPA 3 -- A3 Untreated-Related
##############################################################################
## Untreated Regressions

pusap3 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = pu3, family = binomial)

# Exclude Influential Data

pu3.1 <- subset(pu3, SubID != c("PM08SEA", "PM06SEA", "PM17SEA"))

# Re-analysis without influential data

pusap3.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1 | SubID) + (1 | Item_No), data = pu3.1, family = binomial)

pusap3.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = pu3.1, family =
binomial)
##############################################################################
# PMT Group -- A3 Untreated-Unrelated
##############################################################################
## POM Regressions

ppsap3 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
```

```
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = pp3, family = binomial)

# Exclude Influential Data

pp3.1 <- subset(pp3, SubID != c("PM01PDX", "PM08SEA", "PM12SEA", "PM17SEA"))
## Warning in `!=.default`(SubID, c("PM01PDX", "PM08SEA", "PM12SEA", "PM17SEA")):
## longer object length is not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

ppsap3.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1 | SubID) + (1 | Item_No), data = pp3.1, family = binomial)

ppsap3.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = pp3.1, family =
binomial)
################################################################################
# SFA Group -- A2 Data
################################################################################c
<- read.csv("Logistic_zSFA_capped1.csv")

c <- c %>%
  mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
  na.omit(c)
c <- subset(c, Time == "A2")

st2 <- subset(c, Type_Key == "T")
su2 <- subset(c, Type_Key == "U")
sp2 <- subset(c, Type_Key == "POM")
################################################################################
# SFA Group SAPA 3 -- A2 Treated
################################################################################
## Treated Regressions

stsap2 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1 | SubID) + (1 | Item_No), data = st2, family = binomial)

# Exclude Influential Data

st2.1 <- subset(st2, SubID != c("SF02PDX", "SF06PDX"))

# Re-analysis without influential data

stsap2.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1 | SubID) + (1 | Item_No), data = st2.1, family = binomial)
```

```
stsap2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = st2.1, family = binomial)
##############################################################################
# SFA Group SAPA 3 -- A2 Untreated-Related
##############################################################################
## Untreated Regressions

susap2 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1|SubID) + (1|Item_No), data = su2, family = binomial)

# Exclude Influential Data

su2.1 <- subset(su2, SubID != c("SF02PDX","SF06PDX"))

# Re-analysis without influential data

susap2.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1|SubID) + (1|Item_No), data = su2.1, family = binomial)

susap2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = su2.1, family = binomial)
##############################################################################
# SFA Group SAPA 3 -- A2 Untreated-Unrelated
##############################################################################
## POM Regressions

spsap2 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1|SubID) + (1|Item_No), data = sp2, family = binomial)

# Exclude Influential Data

sp2.1 <- subset(sp2, SubID != c("SF02PDX", "SF06SEA", "SF07PDX", "SF18SEA"))

# Re-analysis without influential data

spsap2.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1|SubID) + (1|Item_No), data = sp2.1, family = binomial)

spsap2.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = sp2.1, family = binomial)
##############################################################################
# SFA Group -- A3 Data
##############################################################################d
<- read.csv("Logistic_zSFA_capped1.csv")

d <- d %>%
  mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
  na.omit(d)
```

```
d <- subset(d, Time == "A3")

st3 <- subset(d, Type_Key == "T")
su3 <- subset(d, Type_Key == "U")
sp3 <- subset(d, Type_Key == "POM")
##############################################################################
# SFA Group SAPA 3 -- A3 Treated
##############################################################################
## Treated Regressions

stsap3 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1|SubID) + (1|Item_No), data = st3, family = binomial)

# Exclude Influential Data

st3.1 <- subset(st3, SubID != c("SF02PDX", "SF05SEA", "SF07PDX"))
## Warning in `!=.default`(SubID, c("SF02PDX", "SF05SEA", "SF07PDX")): longer
## object length is not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

stsap3.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1|SubID) + (1|Item_No), data = st3.1, family = binomial)

stsap3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = st3.1, family = binomial)
##############################################################################
# SFA Group SAPA 3 -- A3 Untreated-Related
##############################################################################
## Untreated Regressions

susap3 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1|SubID) + (1|Item_No), data = su3, family = binomial)

# Exclude Influential Data

su3.1 <- subset(su3, SubID != c("SF02PDX", "SF05SEA"))

# Re-analysis without influential data

susap3.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1|SubID) + (1|Item_No), data = su3.1, family = binomial)

susap3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = su3.1, family = binomial)
##############################################################################
# SFA Group -- A3 Untreated-Unrelated
```

```
##############################################################################
## POM Regressions

spsap3 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend + ZRWParse
+ ZNWParse + (1|SubID) + (1|Item_No), data = sp3, family = binomial)


# Exclude Influential Data

sp3.1 <- subset(sp3, SubID != c("SF02PDX", "SF06PDX"))
## Warning in `!=.default`(SubID, c("SF02PDX", "SF06PDX")): longer object length is
## not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

spsap3.1 <- glmer(Response ~ 1 + ZRWRep + ZNWRep + ZRWBlend + ZNWBlend +
ZRWParse + ZNWParse + (1|SubID) + (1|Item_No), data = sp3.1, family = binomial)

spsap3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = sp3.1, family = binomial)
##############################################################################
## SEVERITY POST HOC
##############################################################################


#### A2 DATA

## Read in Whole Group Data

z <- read.csv("Logistic_zAll.csv")

z <- z %>%
  mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
  na.omit(z)
z <- subset(z, Time == "A2")

mildat <- subset(z, Type_Key == "T" & Severity == "Mild")
mildau <- subset(z, Type_Key == "U" & Severity == "Mild")
mildap <- subset(z, Type_Key == "POM" & Severity == "Mild")

sevat <- subset(z, Type_Key == "T" & Severity == "Severe")
sevau <- subset(z, Type_Key == "U" & Severity == "Severe")
sevap <- subset(z, Type_Key == "POM" & Severity == "Severe")

## Read in PMT Group Data

y <- read.csv("Logistic_zPMT.csv")
```

```
y <- y %>%
  mutate_at(vars(c(Item_No)), list(factor)) %>%
  na.omit(y)
y <- subset(y, Time == "A2")

mildpt <- subset(y, Type_Key == "T" & Severity == "Mild")
mildpu <- subset(y, Type_Key == "U" & Severity == "Mild")
mildpp <- subset(y, Type_Key == "POM" & Severity == "Mild")

sevpt <- subset(y, Type_Key == "T" & Severity == "Severe")
sevpu <- subset(y, Type_Key == "U" & Severity == "Severe")
sevpp <- subset(y, Type_Key == "POM" & Severity == "Severe")

## Read in SFA Group Data

x <- read.csv("Logistic_zSFA.csv")
x <- x %>%
  mutate_at(vars(Item_No), list(factor)) %>%
  na.omit(x)
x <- subset(x, Time == "A2")

mildst <- subset(x, Type_Key == "T" & Severity == "Mild")
mildsu <- subset(x, Type_Key == "U" & Severity == "Mild")
mildsp <- subset(x, Type_Key == "POM" & Severity == "Mild")

sevst <- subset(x, Type_Key == "T" & Severity == "Severe")
sevsu <- subset(x, Type_Key == "U" & Severity == "Severe")
sevsp <- subset(x, Type_Key == "POM" & Severity == "Severe")
################################################################################
## Whole Group Analyses -- A2 Severity Post Hoc
################################################################################

#####################################################
## Treated
#####################################################

## Mild

mat <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildat, family = binomial)

# Exclude Influential Data

wtm <- subset(mildat, SubID != c("PM08SEA", "SF05SEA"))
## Warning in `!=.default`(SubID, c("PM08SEA", "SF05SEA")): longer object length is
## not a multiple of shorter object length
```

```
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

mat.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|Sub
ID) + (1|Item_No), data = wtm, family = binomial)

mat.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = wtm, family
= binomial)


## Severe

sat <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubI
D) + (1|Item_No), data = sevat, family = binomial)

# Exclude Influential Data

wts <- subset(sevat, SubID != c("SF10PDX"))

# Re-analysis without influential data

sat.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|Sub
ID) + (1|Item_No), data = wts, family = binomial)

sat.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = wts, family
= binomial)


#####################################################
## Untreated
#####################################################

## Mild

mau <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildau, family = binomial)

# Exclude Influential Data

wum <- subset(mildau, SubID != c("SF05SEA"))

# Re-analysis without influential data

mau.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = wum, family = binomial)
```

```
mau.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = wum, family = binomial)


## Severe

sau <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevau, family = binomial)

# Exclude Influential Data

wus <- subset(sevau, SubID != c("PM14SEA", "SF10PDX"))

# Re-analysis without influential data

sau.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = wus, family = binomial)

sau.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = wus, family = binomial)


#####################################################
## POM
#####################################################

## Mild

map.0 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = mildap, family = binomial)

# Exclude Influential Data

wpm <- subset(mildap, SubID != c("PM12SEA", "SF05SEA"))
## Warning in `!=.default`(SubID, c("PM12SEA", "SF05SEA")): longer object length is
## not a multiple of shorter object length

## Warning in `!=.default`(SubID, c("PM12SEA", "SF05SEA")): longer object length is
## not a multiple of shorter object length
# Re-analysis without influential data

map.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = wpm, family = binomial)

map.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = wpm, family = binomial)


## Severe
```

```
sap <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevap, family = binomial)

# Exclude Influential Data

wps <- subset(sevap, SubID != c("PM04SEA", "PM14SEA", "SF10PDX"))
## Warning in `!=.default`(SubID, c("PM04SEA", "PM14SEA", "SF10PDX")): longer
## object length is not a multiple of shorter object length

## Warning in `!=.default`(SubID, c("PM04SEA", "PM14SEA", "SF10PDX")): longer
## object length is not a multiple of shorter object length
# Re-analysis without influential data

sap.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = wps, family = binomial)

sap.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = wps, family = binomial)
################################################################################
## PMT Group Analyses -- A2 Severity Post Hoc
################################################################################


#################################################
## Treated
#################################################

## Mild

mpt <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildpt, family = binomial)

# Exclude Influential Data

vtm <- subset(mildpt, SubID != c("PM07PDX", "PM08SEA", "SF15SEA"))

# Re-analysis without influential data

mpt.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = vtm, family = binomial)

mpt.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vtm, family = binomial)


## Severe
```

```
spt <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevpt, family = binomial)

# Exclude Influential Data

vts <- subset(sevpt, SubID != c("PM04SEA"))

# Re-analysis without influential data

spt.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = vts, family = binomial)

spt.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vts, family = binomial)


#####################################################
## Untreated
#####################################################

## Mild -- Model Failed to Converge

mpu <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildpu, family = binomial)

# Exclude Influential Data

vum <- subset(mildpu, SubID != c("PM07PDX", "PM16SEA"))
## Warning in `!=.default`(SubID, c("PM07PDX", "PM16SEA")): longer object length is
## not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

mpu.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = vum, family = binomial)
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0042717 (tol = 0.002, component 1)
mpu.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vum, family = binomial)


## Severe

spu <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevpu, family = binomial)

# Exclude Influential Data
```

```
vus <- subset(sevpu, SubID != c("PM14SEA"))

# Re-analysis without influential data

spu.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = vus, family = binomial)

spu.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vus, family = binomial)


#####################################################
## POM
#####################################################

## Mild

mpp <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildpp, family = binomial)

# Exclude Influential Data

vpm <- subset(mildpp, SubID != c("PM01PDX", "PM02PDX"))
## Warning in `!=.default`(SubID, c("PM01PDX", "PM02PDX")): longer object length is
## not a multiple of shorter object length

## Warning in `!=.default`(SubID, c("PM01PDX", "PM02PDX")): longer object length is
## not a multiple of shorter object length
# Re-analysis without influential data

mpp.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = vpm, family = binomial)

mpp.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vpm, family = binomial)


## Severe

spp <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevpp, family = binomial)

# Exclude Influential Data

vps <- subset(sevpp, SubID != c("PM14SEA"))

# Re-analysis without influential data
```

```
spp.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = vps, family = binomial)

spp.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vps, family = binomial)
################################################################################
## SFA Group Analyses -- A2 Severity Post Hoc
################################################################################

#################################################
## Treated
#################################################

## Mild -- Null Model Best Fit

mst <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildst, family = binomial)

# Exclude Influential Data

utm <- subset(mildst, SubID != c("SF05SEA", "SF07PDX"))

# Re-analysis without influential data

mst.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = utm, family = binomial)

mst.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = utm, family = binomial)


## Severe -- Null Model Best Fit

sst <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevst, family = binomial)
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00361039 (tol = 0.002, component 1)
# Exclude Influential Data

uts <- subset(sevst, SubID != c("SF10SEA", "SF17SEA"))

# Re-analysis without influential data

sst.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = uts, family = binomial)

sst.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = uts, family = binomial)
```

```
#####################################################
## Untreated
#####################################################

## Mild

msu <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildsu, family = binomial)

# Exclude Influential Data

uum <- subset(mildsu, SubID != c("SF05SEA", "PSF07PDX"))

# Re-analysis without influential data

msu.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = uum, family = binomial)

msu.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = uum, family = binomial)


## Severe -- Null Model Best Fit

ssu <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevsu, family = binomial)

# Exclude Influential Data

uus <- subset(sevsu, SubID != c("SF06PDX", "SF10PDX", "SF10SEA", "SF17SEA"))

# Re-analysis without influential data

ssu.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = uus, family = binomial)

ssu.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = uus, family = binomial)


#####################################################
## POM
#####################################################

## Mild
```

```
msp <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildsp, family = binomial)

# Exclude Influential Data

upm <- subset(mildsp, SubID != c("SF05SEA", "PSF07PDX"))

# Re-analysis without influential data

msp.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = upm, family = binomial)

msp.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = upm, family = binomial)


## Severe -- Failed to Converge

ssp <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevsp, family = binomial)
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00232128 (tol = 0.002, component 1)
# Exclude Influential Data

ups <- subset(sevsp, SubID != c("SF10PDX", "SF10SEA", "SF17SEA"))
## Warning in `!=.default`(SubID, c("SF10PDX", "SF10SEA", "SF17SEA")): longer
## object length is not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

ssp.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = ups, family = binomial)
## boundary (singular) fit: see ?isSingular
ssp.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = ups, family = binomial)
#### A3 DATA

## Read in Whole Group Data

z <- read.csv("Logistic_zAll_capped1.csv")

z <- z %>%
  mutate_at(vars(c(Time_Code, Item_No)), list(factor)) %>%
  na.omit(z)
z <- subset(z, Time == "A3")

mildat3 <- subset(z, Type_Key == "T" & Severity == "Mild")
```

```
mildau3 <- subset(z, Type_Key == "U" & Severity == "Mild")
mildap3 <- subset(z, Type_Key == "POM" & Severity == "Mild")

sevat3 <- subset(z, Type_Key == "T" & Severity == "Severe")
sevau3 <- subset(z, Type_Key == "U" & Severity == "Severe")
sevap3 <- subset(z, Type_Key == "POM" & Severity == "Severe")

## Read in PMT Group Data

y <- read.csv("Logistic_zPMT_capped1.csv")
y <- y %>%
  mutate_at(vars(c(Item_No)), list(factor)) %>%
  na.omit(y)
y <- subset(y, Time == "A3")

mildpt3 <- subset(y, Type_Key == "T" & Severity == "Mild")
mildpu3 <- subset(y, Type_Key == "U" & Severity == "Mild")
mildpp3 <- subset(y, Type_Key == "POM" & Severity == "Mild")

sevpt3 <- subset(y, Type_Key == "T" & Severity == "Severe")
sevpu3 <- subset(y, Type_Key == "U" & Severity == "Severe")
sevpp3 <- subset(y, Type_Key == "POM" & Severity == "Severe")

## Read in SFA Group Data

x <- read.csv("Logistic_zSFA_capped1.csv")
x <- x %>%
  mutate_at(vars(Item_No), list(factor)) %>%
  na.omit(x)
x <- subset(x, Time == "A3")

mildst3 <- subset(x, Type_Key == "T" & Severity == "Mild")
mildsu3 <- subset(x, Type_Key == "U" & Severity == "Mild")
mildsp3 <- subset(x, Type_Key == "POM" & Severity == "Mild")

sevst3 <- subset(x, Type_Key == "T" & Severity == "Severe")
sevsu3 <- subset(x, Type_Key == "U" & Severity == "Severe")
sevsp3 <- subset(x, Type_Key == "POM" & Severity == "Severe")
#############################################################################
## Whole Group Analyses -- A3 Severity Post Hoc
#############################################################################

#####################################################
## Treated A3
#####################################################
```

```
## Mild

mat3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | SubID) +
(1 | Item_No), data = mildat3, family = binomial)

# Exclude Influential Data

wtm3 <- subset(mildat3, SubID != c("SF05SEA", "SF12PDX"))

# Re-analysis without influential data

mat3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | SubID)
+ (1 | Item_No), data = wtm3, family = binomial)

mat3.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = wtm3, family = binomial)


## Severe

sat3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | SubID) +
(1 | Item_No), data = sevat3, family = binomial)

# Exclude Influential Data

wts3 <- subset(sevat3, SubID != c("PM03SEA", "PM04SEA", "PM14SEA", "SF04SEA",
"SF10PDX", "SF10SEA"))
## Warning in `!=.default`(SubID, c("PM03SEA", "PM04SEA", "PM14SEA", "SF04SEA", :
## longer object length is not a multiple of shorter object length
## Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
## shorter object length
# Re-analysis without influential data

sat3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | SubID) +
(1 | Item_No), data = wts3, family = binomial)

sat3.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = wts3, family = binomial)


#####################################################
## Untreated
#####################################################

## Mild

mau3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | SubID) +
(1 | Item_No), data = mildau3, family = binomial)
```

```
# Exclude Influential Data

wum3 <- subset(mildau3, SubID != c("SF05SEA", "SF07PDX", "SF12PDX", "SF18SEA"))
## Warning in `!=.default`(SubID, c("SF05SEA", "SF07PDX", "SF12PDX", "SF18SEA")):
## longer object length is not a multiple of shorter object length

## Warning in `!=.default`(SubID, c("SF05SEA", "SF07PDX", "SF12PDX", "SF18SEA")):
## longer object length is not a multiple of shorter object length
# Re-analysis without influential data

mau3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | SubID)
+ (1 | Item_No), data = wum3, family = binomial)

mau3.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = wum3, family = binomial)


## Severe

sau3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | SubID) +
(1 | Item_No), data = sevau3, family = binomial)

# Exclude Influential Data

wus3 <- subset(sevau3, SubID != c("SF04SEA"))

# Re-analysis without influential data

sau3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | SubID)
+ (1 | Item_No), data = wus3, family = binomial)

sau3.null <- glmer(Response ~ 1 + (1 | SubID) + (1 | Item_No), data = wus3, family = binomial)


#####################################################
## POM
#####################################################

## Mild

map3.0 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1 | SubID)
+ (1 | Item_No), data = mildap3, family = binomial)

# Exclude Influential Data

wpm3 <- subset(mildap3, SubID != c("SF05SEA", "SF06SEA"))
```

```
## Warning in `!=.default`(SubID, c("SF05SEA", "SF06SEA")): longer object length is
## not a multiple of shorter object length

## Warning in `!=.default`(SubID, c("SF05SEA", "SF06SEA")): longer object length is
## not a multiple of shorter object length
# Re-analysis without influential data

map3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = wpm3, family = binomial)

map3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = wpm3, family = binomial)


## Severe

sap3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevap3, family = binomial)

# Exclude Influential Data

wps3 <- subset(sevap, SubID != c("PM06SEA", "PM14SEA", "SF04SEA", "SF10PDX",
"SF10SEA"))

# Re-analysis without influential data

sap3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = wps3, family = binomial)

sap3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = wps3, family = binomial)
################################################################################
## PMT Group Analyses -- A3 Severity Post Hoc
################################################################################

#####################################################
## Treated
#####################################################

## Mild

mpt3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildpt3, family = binomial)
## boundary (singular) fit: see ?isSingular
# Exclude Influential Data

vtm3 <- subset(mildpt3, SubID != c("PM07PDX"))
```

Re-analysis without influential data

mpt3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) + (1|Item_No), data = vtm3, family = binomial)
boundary (singular) fit: see ?isSingular
mpt3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vtm3, family = binomial)

Severe

spt3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) + (1|Item_No), data = sevpt3, family = binomial)

Exclude Influential Data

vts3 <- subset(sevpt3, SubID != c("PM04SEA", "PM06PDX", "PM06SEA", "PM14SEA"))
Warning in `!=.default`(SubID, c("PM04SEA", "PM06PDX", "PM06SEA", "PM14SEA")):
longer object length is not a multiple of shorter object length
Warning in is.na(e1) | is.na(e2): longer object length is not a multiple of
shorter object length

```
# Re-analysis without influential data

spt3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) + (1|Item_No), data = vts3, family = binomial)

spt3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vts3, family = binomial)
```

##
Untreated
##

Mild -- Model Failed to Converge

mpu3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) + (1|Item_No), data = mildpu3, family = binomial)
boundary (singular) fit: see ?isSingular
Exclude Influential Data

vum3 <- subset(mildpu3, SubID != c("PM01PDX", "PM02SEA", "PM07PDX", "PM15SEA"))
Warning in `!=.default`(SubID, c("PM01PDX", "PM02SEA", "PM07PDX", "PM15SEA")):
longer object length is not a multiple of shorter object length

```
## Warning in `!=.default`(SubID, c("PM01PDX", "PM02SEA", "PM07PDX", "PM15SEA")):
## longer object length is not a multiple of shorter object length
# Re-analysis without influential data

mpu3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = vum3, family = binomial)
## boundary (singular) fit: see ?isSingular
mpu3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vum3, family = binomial)


## Severe

spu3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevpu3, family = binomial)

# Exclude Influential Data

vus3 <- subset(sevpu3, SubID != c("PM04SEA", "PM14SEA"))
## Warning in `!=.default`(SubID, c("PM04SEA", "PM14SEA")): longer object length is
## not a multiple of shorter object length

## Warning in `!=.default`(SubID, c("PM04SEA", "PM14SEA")): longer object length is
## not a multiple of shorter object length
# Re-analysis without influential data

spu3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = vus3, family = binomial)

spu3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vus3, family = binomial)


#####################################################
## POM
#####################################################

## Mild

mpp3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildpp3, family = binomial)

# Exclude Influential Data

vpm3 <- subset(mildpp3, SubID != c("PM01PDX", "PM09PDX"))

# Re-analysis without influential data
```

```
mpp3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = vpm3, family = binomial)

mpp3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vpm3, family = binomial)


## Severe

spp3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevpp3, family = binomial)

# Exclude Influential Data

vps3 <- subset(sevpp3, SubID != c("PM03SEA", "PM04SEA", "PM05SEA", "PM06SEA",
"PM14SEA"))

# Re-analysis without influential data

spp3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = vps3, family = binomial)

spp3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = vps3, family = binomial)
################################################################################
## SFA Group Analyses -- A3 Severity Post Hoc
################################################################################

#################################################
## Treated
#################################################

## Mild -- Null Model Best Fit

mst3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildst3, family = binomial)

# Exclude Influential Data

utm3 <- subset(mildst3, SubID != c("SF05SEA", "SF07PDX"))

# Re-analysis without influential data

mst3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = utm3, family = binomial)

mst3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = utm3, family = binomial)
```

```
## Severe -- Null Model Best Fit

sst3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevst3, family = binomial)
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00522848 (tol = 0.002, component 1)
# Exclude Influential Data

uts3 <- subset(sevst3, SubID != c("SF10SEA", "SF17SEA"))

# Re-analysis without influential data

sst3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = uts3, family = binomial)

sst3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = uts3, family = binomial)


#######################################################
## Untreated
#######################################################

## Mild

msu3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildsu3, family = binomial)

# Exclude Influential Data

uum3 <- subset(mildsu3, SubID != c("SF05SEA", "PSF07PDX", "SF18SEA"))

# Re-analysis without influential data

msu3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = uum3, family = binomial)

msu3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = uum3, family = binomial)


## Severe -- Null Model Best Fit

ssu3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevsu3, family = binomial)

# Exclude Influential Data
```

```
uus3 <- subset(sevsu3, SubID != c("SF06PDX", "SF10PDX", "SF10SEA", "SF17SEA"))

# Re-analysis without influential data

ssu3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = uus3, family = binomial)
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0195789 (tol = 0.002, component 1)
ssu3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = uus3, family = binomial)


#####################################################
## POM
#####################################################

## Mild

msp3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = mildsp3, family = binomial)

# Exclude Influential Data

upm3 <- subset(mildsp3, SubID != c("SF05PDX", "SF05SEA", "PSF07PDX"))

# Re-analysis without influential data

msp3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = upm3, family = binomial)

msp3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = upm3, family = binomial)


## Severe -- Failed to Converge

ssp3 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID) +
(1|Item_No), data = sevsp3, family = binomial)

# Exclude Influential Data

ups3 <- subset(sevsp3, SubID != c("SF10PDX", "SF10SEA", "SF17SEA"))

# Re-analysis without influential data

ssp3.1 <- glmer(Response ~ 1 + ZCAT_T + ZSAPA2 + ZSAPA3 + ZSem + ZPhon + (1|SubID)
+ (1|Item_No), data = ups3, family = binomial)
```

```
ssp3.null <- glmer(Response ~ 1 + (1|SubID) + (1|Item_No), data = ups3, family = binomial)
```

Appendix H: Results Tables for Post Hoc 1 – Severity

Table 29

Logistic Regression Results: All Mild Group - Treated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.80	1.33 – 2.44	**<0.001**	<0.001
CAT Comp. T-score	1.21	0.90 – 1.64	0.353	0.212
SAPA 2	1.01	0.75 – 1.36	0.978	0.964
SAPA 3	1.56	1.11 – 2.19	**0.059**	0.011
Semantic Impairment	0.91	0.69 – 1.20	0.680	0.524
Phonological Impairment	1.04	0.77 – 1.40	0.898	0.821
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.13			
$\tau_{00\ SubID}$	0.41			
ICC	0.14			
N_{SubID}	33			
N_{Item_No}	95			
Observations	5704			
Marginal R^2 / Conditional R^2	0.052 / 0.185			

Table 30

Logistic Regression Results: All Mild Group - Untreated-Related Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.91	1.47 – 2.48	**<0.001**	<0.001
CAT Comp. T-score	1.64	1.29 – 2.09	**<0.001**	<0.001
SAPA 2	1.00	0.79 – 1.25	0.994	0.987
SAPA 3	1.34	1.04 – 1.73	**0.092**	0.026
Semantic Impairment	0.78	0.62 – 0.97	**0.091**	0.025
Phonological Impairment	0.71	0.56 – 0.92	**0.050**	0.009
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.22			
$\tau_{00\ SubID}$	0.19			
ICC	0.11			
N_{SubID}	32			
N_{Item_No}	56			
Observations	2897			
Marginal R^2 / Conditional R^2	0.065 / 0.168			

Table 31

Logistic Regression Results: All Mild Group - Untreated-Unrelated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.12	0.75 – 1.69	0.730	0.581
CAT Comp. T-score	1.44	1.00 – 2.07	0.152	0.050
SAPA 2	1.05	0.73 – 1.49	0.895	0.808
SAPA 3	2.10	1.39 – 3.17	**<0.001**	<0.001
Semantic Impairment	1.03	0.74 – 1.44	0.912	0.853
Phonological Impairment	0.77	0.53 – 1.10	0.295	0.151
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.56			
$\tau_{00\ Item\_No}$	0.28			
ICC	0.2			
N_{SubID}	33			
N_{Item_No}	25			
Observations	2262			
Marginal R^2 / Conditional R^2	0.129 / 0.305			

Table 32

Logistic Regression Results: All Mild Group - Treated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.18	0.87 – 1.62	0.449	0.292
CAT T-score	1.15	0.84 – 1.56	0.541	0.389
SAPA 2	1.06	0.78 – 1.44	0.842	0.717
SAPA 3	1.53	1.09 – 2.17	**0.071**	0.015
Semantic Impairment	0.97	0.72 – 1.29	0.898	0.824
Phonological Impairment	1.01	0.74 – 1.37	0.969	0.940
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.12			
$\tau_{00\ SubID}$	0.43			
ICC	0.14			
N_{SubID}	33			
N_{Item_No}	95			
Observations	5406			
Marginal R^2 / Conditional R^2	0.050 / 0.186			

Table 33

Logistic Regression Results: All Mild Group - Untreated-Related Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.19	0.84 – 1.68	0.482	0.33
CAT T-score	1.13	0.81 – 1.57	0.626	0.464
SAPA 2	1.14	0.82 – 1.59	0.590	0.428
SAPA 3	1.64	1.13 – 2.37	**0.050**	0.009
Semantic Impairment	0.99	0.72 – 1.35	0.965	0.926
Phonological Impairment	1.07	0.77 – 1.48	0.823	0.698
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.21			
$\tau_{00\ SubID}$	0.47			
ICC	0.17			
N_{SubID}	33			
N_{Item_No}	56			
Observations	2764			
Marginal R^2 / Conditional R^2	0.067 / 0.226			

Table 34

Logistic Regression Results: All Mild Group - Untreated-Unrelated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.04	0.66 – 1.64	0.912	0.857
CAT T-score	1.30	0.87 – 1.96	0.349	0.207
SAPA 2	1.01	0.69 – 1.50	0.970	0.948
SAPA 3	2.09	1.37 – 3.19	**0.010**	0.001
Semantic Impairment	1.32	0.92 – 1.89	0.282	0.137
Phonological Impairment	0.79	0.54 – 1.14	0.349	0.205
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.60			
$\tau_{00\ Item\_No}$	0.37			
ICC	0.23			
N_{SubID}	32			
N_{Item_No}	25			
Observations	2149			
Marginal R^2 / Conditional R^2	0.149 / 0.343			

Table 35

Logistic Regression Results: All Severe Group - Treated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.80	0.44 – 1.43	0.604	0.445
CAT Comp. T-score	1.05	0.62 – 1.78	0.912	0.858
SAPA 2	1.26	0.93 – 1.71	0.282	0.137
SAPA 3	1.95	1.30 – 2.90	**0.010**	0.001
Semantic Impairment	1.65	1.10 – 2.50	**0.071**	0.017
Phonological Impairment	1.94	1.33 – 2.83	**0.010**	0.001
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.26			
$\tau_{00\ SubID}$	0.24			
ICC	0.13			
N_{SubID}	21			
N_{Item_No}	95			
Observations	2881			
Marginal R^2 / Conditional R^2	0.112 / 0.229			

Table 36

Logistic Regression Results: All Severe Group - Untreated-Related Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.67	0.34 – 1.36	0.424	0.27
CAT Comp. T-score	0.83	0.43 – 1.61	0.732	0.587
SAPA 2	1.30	0.88 – 1.93	0.336	0.188
SAPA 3	1.95	1.13 – 3.39	**0.071**	0.017
Semantic Impairment	1.61	0.92 – 2.80	0.218	0.093
Phonological Impairment	1.82	1.20 – 2.77	**0.034**	0.005
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.12			
$\tau_{00\ SubID}$	0.44			
ICC	0.14			
N_{SubID}	22			
N_{Item_No}	56			
Observations	1508			
Marginal R^2 / Conditional R^2	0.099 / 0.229			

Table 37

Logistic Regression Results: All Severe Group - Untreated-Unrelated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.73	0.38 – 1.41	0.496	0.351
CAT Comp. T-score	0.82	0.45 – 1.51	0.685	0.53
SAPA 2	1.27	0.87 – 1.83	0.353	0.212
SAPA 3	1.67	1.01 – 2.76	0.140	0.044
Semantic Impairment	2.02	1.23 – 3.33	**0.039**	0.006
Phonological Impairment	1.89	1.28 – 2.79	**0.010**	0.001
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.2			
$\tau_{00\ SubID}$	0.36			
ICC	0.15			
N_{SubID}	22			
N_{Item_No}	25			
Observations	1434			
Marginal R^2 / Conditional R^2	0.115 / 0.244			

Table 38

Logistic Regression Results: All Severe Group - Treated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.41	0.21 – 0.77	**0.039**	0.006
CAT T-score	0.84	0.46 – 1.53	0.713	0.564
SAPA 2	1.40	0.97 – 2.02	0.183	0.071
SAPA 3	1.60	0.96 – 2.65	0.179	0.069
Semantic Impairment	1.44	0.86 – 2.41	0.306	0.161
Phonological Impairment	1.36	0.93 – 1.99	0.252	0.115
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.13			
$\tau_{00\ SubID}$	0.40			
ICC	0.14			
N_{SubID}	22			
N_{Item_No}	95			
Observations	2817			
Marginal R^2 / Conditional R^2	0.062 / 0.194			

Table 39

Logistic Regression Results: All Severe Group - Untreated-Related Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.41	0.22 – 0.79	**0.043**	0.007
CAT T-score	0.73	0.39 – 1.39	0.484	0.339
SAPA 2	2.83	1.78 – 4.50	**<0.001**	<0.001
SAPA 3	1.67	1.03 – 2.69	0.119	0.036
Semantic Impairment	2.45	1.45 – 4.14	**0.010**	0.001
Phonological Impairment	1.21	0.83 – 1.76	0.476	0.318
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.20			
$\tau_{00\ SubID}$	0.30			
ICC	0.13			
N_{SubID}	21			
N_{Item_No}	56			
Observations	1457			
Marginal R^2 / Conditional R^2	0.179 / 0.288			

Table 40

Logistic Regression Results: All Severe Group - Untreated-Unrelated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.83	0.40 – 1.76	0.777	0.632
CAT T-score	0.97	0.49 – 1.91	0.965	0.929
SAPA 2	1.39	0.92 – 2.11	0.253	0.117
SAPA 3	1.46	0.83 – 2.58	0.341	0.192
Semantic Impairment	2.06	1.16 – 3.65	**0.070**	0.014
Phonological Impairment	1.69	1.10 – 2.60	**0.071**	0.017
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.25			
$\tau_{00\ SubID}$	0.49			
ICC	0.18			
N_{SubID}	22			
N_{Item_No}	25			
Observations	1390			
Marginal R^2 / Conditional R^2	0.097 / 0.263			

Table 41

Logistic Regression Results: PMT Mild Group - Treated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.43	0.92 – 2.23	0.252	0.115
CAT Comp. T-score	1.31	0.95 – 1.81	0.224	0.098
SAPA 2	0.73	0.59 – 0.90	**0.030**	0.004
SAPA 3	2.99	2.05 – 4.34	**<0.001**	<0.001
Semantic Impairment	1.16	0.90 – 1.48	0.406	0.25
Phonological Impairment	0.99	0.75 – 1.32	0.980	0.969
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.46			
$\tau_{00\ SubID}$	0.05			
ICC	0.13			
N_{SubID}	13			
N_{Item_No}	39			
Observations	1365			
Marginal R^2 / Conditional R^2	0.083 / 0.206			

Table 42

Logistic Regression Results: PMT Mild Group - Untreated-Unrelated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.74	0.39 – 1.38	0.484	0.339
CAT Comp. T-score	1.94	1.20 – 3.14	**0.043**	0.007
SAPA 2	0.97	0.70 – 1.34	0.912	0.856
SAPA 3	2.34	1.41 – 3.89	**0.010**	0.001
Semantic Impairment	0.96	0.67 – 1.37	0.895	0.809
Phonological Impairment	0.43	0.28 – 0.67	**<0.001**	<0.001
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.2			
$\tau_{00\ SubID}$	0.17			
ICC	0.1			
N_{SubID}	13			
N_{Item_No}	25			
Observations	850			
Marginal R^2 / Conditional R^2	0.133 / 0.219			

Table 43

Logistic Regression Results: PMT Mild Group - Untreated-Unrelated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.86	0.44 – 1.68	0.792	0.66
CAT T-score	1.53	0.94 – 2.49	0.213	0.09
SAPA 2	1.17	0.84 – 1.63	0.496	0.352
SAPA 3	2.82	1.66 – 4.78	**<0.001**	<0.001
Semantic Impairment	1.00	0.69 – 1.44	0.996	0.996
Phonological Impairment	0.37	0.23 – 0.59	**<0.001**	<0.001
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.50			
$\tau_{00\ SubID}$	0.16			
ICC	0.17			
N_{SubID}	13			
N_{Item_No}	25			
Observations	774			
Marginal R^2 / Conditional R^2	0.162 / 0.301			

Table 44

Logistic Regression Results: PMT Severe Group - Treated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.60	0.38 – 0.94	**0.092**	0.027
CAT Comp. T-score	0.70	0.46 – 1.05	0.204	0.084
SAPA 2	1.09	0.75 – 1.58	0.786	0.648
SAPA 3	2.05	1.51 – 2.79	**<0.001**	<0.001
Semantic Impairment	1.43	1.08 – 1.91	**0.066**	0.013
Phonological Impairment	2.59	1.70 – 3.93	**<0.001**	<0.001
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.43			
$\tau_{00\ SubID}$	0.08			
ICC	0.13			
N_{SubID}	13			
N_{Item_No}	39			
Observations	1404			
Marginal R^2 / Conditional R^2	0.148 / 0.262			

Table 45

Logistic Regression Results: PMT Severe Group - Untreated-Related Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.70	0.40 – 1.22	0.349	0.206
CAT Comp. T-score	1.20	0.71 – 2.01	0.659	0.498
SAPA 2	1.14	0.75 – 1.73	0.700	0.55
SAPA 3	1.41	0.99 – 2.00	0.162	0.057
Semantic Impairment	1.62	1.13 – 2.34	**0.050**	0.009
Phonological Impairment	2.10	1.30 – 3.39	**0.017**	0.002
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.18			
$\tau_{00\ SubID}$	0.12			
ICC	0.08			
N_{SubID}	13			
N_{Item_No}	21			
Observations	756			
Marginal R^2 / Conditional R^2	0.089 / 0.164			

Table 46

Logistic Regression Results: PMT Severe Group - Untreated-Unrelated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.76	0.50 – 1.14	0.336	0.186
CAT Comp. T-score	1.13	0.78 – 1.62	0.677	0.519
SAPA 2	1.33	0.98 – 1.80	0.171	0.064
SAPA 3	1.30	1.03 – 1.63	**0.092**	0.026
Semantic Impairment	1.92	1.51 – 2.43	**<0.001**	<0.001
Phonological Impairment	1.93	1.40 – 2.67	**<0.001**	<0.001
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.28			
$\tau_{00\ SubID}$	0.02			
ICC	0.08			
N_{SubID}	13			
N_{Item_No}	25			
Observations	925			
Marginal R^2 / Conditional R^2	0.106 / 0.179			

Table 47

Logistic Regression Results: PMT Severe Group - Treated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.48	0.31 – 0.76	**0.017**	0.002
CAT T-score	1.11	0.73 – 1.69	0.777	0.633
SAPA 2	1.66	1.16 – 2.38	**0.034**	0.005
SAPA 3	1.31	0.99 – 1.74	0.162	0.057
Semantic Impairment	1.73	1.28 – 2.34	**<0.001**	<0.001
Phonological Impairment	1.67	1.14 – 2.46	**0.050**	0.009
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.17			
$\tau_{00\ SubID}$	0.08			
ICC	0.07			
N_{SubID}	14			
N_{Item_No}	39			
Observations	1337			
Marginal R^2 / Conditional R^2	0.124 / 0.188			

Table 48

Logistic Regression Results: PMT Severe Group - Untreated-Related Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.38	0.21 – 0.70	**0.017**	0.002
CAT T-score	0.94	0.52 – 1.67	0.898	0.825
SAPA 2	2.67	1.56 – 4.58	**<0.001**	<0.001
SAPA 3	1.25	0.89 – 1.77	0.345	0.198
Semantic Impairment	2.36	1.58 – 3.50	**<0.001**	<0.001
Phonological Impairment	1.45	0.87 – 2.43	0.295	0.151
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.29			
$\tau_{00\ SubID}$	0.11			
ICC	0.11			
N_{SubID}	14			
N_{Item_No}	21			
Observations	724			
Marginal R^2 / Conditional R^2	0.218 / 0.302			

Table 49

Logistic Regression Results: PMT Severe Group - Untreated-Unrelated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.10	0.65 – 1.86	0.842	0.733
CAT T-score	1.25	0.75 – 2.08	0.551	0.398
SAPA 2	1.46	0.96 – 2.20	0.188	0.075
SAPA 3	1.26	0.91 – 1.75	0.317	0.169
Semantic Impairment	2.44	1.70 – 3.50	**<0.001**	<0.001
Phonological Impairment	2.83	1.81 – 4.44	**<0.001**	<0.001
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.12			
$\tau_{00\ SubID}$	0.11			
ICC	0.07			
N_{SubID}	14			
N_{Item_No}	25			
Observations	855			
Marginal R^2 / Conditional R^2	0.210 / 0.262			

Table 50

Logistic Regression Results: SFA Mild Group - Untreated-Related Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.40	1.04 – 1.90	**0.095**	0.028
CAT Comp. T-score	1.34	1.00 – 1.79	0.149	0.048
SAPA 2	1.50	1.07 – 2.10	**0.071**	0.017
SAPA 3	1.20	0.87 – 1.65	0.418	0.263
Semantic Impairment	0.98	0.74 – 1.32	0.959	0.913
Phonological Impairment	0.98	0.73 – 1.30	0.913	0.866
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.26			
$\tau_{00\ SubID}$	0.2			
ICC	0.12			
N_{SubID}	20			
N_{Item_No}	41			
Observations	2160			
Marginal R^2 / Conditional R^2	0.072 / 0.185			

Table 51

Logistic Regression Results: SFA Mild Group - Untreated-Unrelated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.34	0.89 – 2.03	0.308	0.163
CAT Comp. T-score	1.47	0.97 – 2.21	0.176	0.067
SAPA 2	1.09	0.68 – 1.74	0.842	0.722
SAPA 3	2.06	1.29 – 3.28	**0.017**	0.002
Semantic Impairment	1.05	0.70 – 1.59	0.895	0.802
Phonological Impairment	1.37	0.92 – 2.05	0.257	0.12
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.27			
$\tau_{00\ SubID}$	0.39			
ICC	0.17			
N_{SubID}	20			
N_{Item_No}	25			
Observations	1375			
Marginal R^2 / Conditional R^2	0.190 / 0.325			

Table 52

Logistic Regression Results: SFA Mild Group - Untreated-Unrelated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	1.22	0.74 – 2.00	0.591	0.431
CAT T-score	1.26	0.77 – 2.07	0.496	0.353
SAPA 2	0.88	0.49 – 1.60	0.809	0.68
SAPA 3	1.97	1.19 – 3.25	**0.048**	0.008
Semantic Impairment	1.43	0.90 – 2.28	0.278	0.133
Phonological Impairment	1.25	0.81 – 1.94	0.476	0.318
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.26			
$\tau_{00\ SubID}$	0.52			
ICC	0.19			
N_{SubID}	19			
N_{Item_No}	25			
Observations	1300			
Marginal R^2 / Conditional R^2	0.183 / 0.340			

Appendix I: Results Tables for Post Hoc 2 – SAPA 3

Table 53

Logistic Regression Results: All Group - SAPA 3 Treated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.94	0.73 – 1.21	0.786	0.649
Real Word Repetition	1.40	1.02 – 1.92	0.125	0.039
Nonword Repetition	1.25	0.91 – 1.70	0.318	0.171
Real Word Blending	1.30	0.92 – 1.85	0.284	0.139
Nonword Blending	0.71	0.49 – 1.05	0.207	0.086
Real Word Parsing	1.56	1.06 – 2.29	**0.086**	0.023
Nonword Parsing	1.34	0.90 – 1.99	0.295	0.152
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.15			
$\tau_{00\ SubID}$	0.76			
ICC	0.22			
N_{SubID}	55			
N_{Item_No}	95			
Observations	8726			
Marginal R^2 / Conditional R^2	0.213 / 0.383			

Table 54

Logistic Regression Results: All Group - SAPA 3 Untreated-Related Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.88	0.67 – 1.16	0.502	0.359
Real Word Repetition	1.37	0.98 – 1.92	0.171	0.064
Nonword Repetition	1.25	0.90 – 1.74	0.327	0.179
Real Word Blending	1.22	0.84 – 1.77	0.449	0.29
Nonword Blending	0.71	0.48 – 1.07	0.231	0.102
Real Word Parsing	1.61	1.07 – 2.40	**0.080**	0.021
Nonword Parsing	1.23	0.81 – 1.87	0.482	0.329
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.14			
$\tau_{00\ SubID}$	0.81			
ICC	0.22			
N_{SubID}	55			
N_{Item_No}	56			
Observations	4475			
Marginal R^2 / Conditional R^2	0.185 / 0.367			

Table 55

Logistic Regression Results: All Group - SAPA 3 Untreated-Unrelated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.83	0.63 – 1.09	0.327	0.179
Real Word Repetition	1.21	0.87 – 1.67	0.407	0.255
Nonword Repetition	1.37	1.00 – 1.88	0.154	0.052
Real Word Blending	1.15	0.81 – 1.64	0.597	0.438
Nonword Blending	0.66	0.45 – 0.97	0.117	0.035
Real Word Parsing	1.60	1.09 – 2.36	**0.071**	0.017
Nonword Parsing	1.58	1.05 – 2.36	**0.092**	0.027
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.73			
$\tau_{00\ Item\_No}$	0.12			
ICC	0.21			
N_{SubID}	55			
N_{Item_No}	25			
Observations	3750			
Marginal R^2 / Conditional R^2	0.215 / 0.376			

Table 56

Logistic Regression Results: All Group - SAPA 3 Treated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.68	0.54 – 0.86	**0.010**	0.001
Real Word Repetition	1.22	0.90 – 1.65	0.348	0.201
Nonword Repetition	1.34	1.00 – 1.80	0.154	0.053
Real Word Blending	1.39	1.00 – 1.93	0.154	0.053
Nonword Blending	0.74	0.51 – 1.05	0.219	0.094
Real Word Parsing	1.42	0.99 – 2.04	0.166	0.059
Nonword Parsing	1.24	0.85 – 1.80	0.418	0.265
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.10			
$\tau_{00\ SubID}$	0.67			
ICC	0.19			
N_{SubID}	55			
N_{Item_No}	95			
Observations	8450			
Marginal R^2 / Conditional R^2	0.185 / 0.339			

Table 57

Logistic Regression Results: All Group - SAPA 3 Untreated-Related Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.66	0.49 – 0.88	**0.034**	0.005
Real Word Repetition	1.19	0.84 – 1.70	0.476	0.321
Nonword Repetition	1.36	0.96 – 1.91	0.204	0.083
Real Word Blending	1.44	0.98 – 2.12	0.171	0.064
Nonword Blending	0.67	0.44 – 1.02	0.171	0.064
Real Word Parsing	1.67	1.09 – 2.54	**0.071**	0.017
Nonword Parsing	1.25	0.81 – 1.93	0.466	0.307
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.20			
$\tau_{00\ SubID}$	0.87			
ICC	0.25			
N_{SubID}	55			
N_{Item_No}	56			
Observations	4331			
Marginal R^2 / Conditional R^2	0.219 / 0.411			

Table 58

Logistic Regression Results: All Group - SAPA 3 Untreated-Unrelated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.85	0.63 – 1.15	0.466	0.307
Real Word Repetition	1.19	0.84 – 1.69	0.484	0.334
Nonword Repetition	1.23	0.86 – 1.74	0.406	0.251
Real Word Blending	1.37	0.92 – 2.02	0.253	0.117
Nonword Blending	0.76	0.50 – 1.15	0.345	0.197
Real Word Parsing	1.57	1.03 – 2.39	0.124	0.038
Nonword Parsing	1.27	0.82 – 1.97	0.443	0.284
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.87			
$\tau_{00\ Item\_No}$	0.13			
ICC	0.23			
N_{SubID}	54			
N_{Item_No}	25			
Observations	3613			
Marginal R^2 / Conditional R^2	0.186 / 0.376			

Table 59

Logistic Regression Results: PMT Group - SAPA 3 Treated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.98	0.67 – 1.45	0.969	0.937
Real Word Repetition	1.12	0.69 – 1.83	0.778	0.637
Nonword Repetition	1.10	0.72 – 1.70	0.787	0.653
Real Word Blending	1.85	1.22 – 2.81	**0.030**	0.004
Nonword Blending	0.97	0.58 – 1.64	0.962	0.919
Real Word Parsing	1.99	1.11 – 3.57	**0.080**	0.021
Nonword Parsing	1.08	0.63 – 1.86	0.877	0.776
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.45			
$\tau_{00\ SubID}$	0.68			
ICC	0.25			
N_{SubID}	27			
N_{Item_No}	39			
Observations	2847			
Marginal R^2 / Conditional R^2	0.264 / 0.451			

Table 60

Logistic Regression Results: PMT Group - SAPA 3 Untreated-Related Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.95	0.65 – 1.38	0.877	0.774
Real Word Repetition	1.09	0.66 – 1.80	0.847	0.74
Nonword Repetition	1.17	0.75 – 1.83	0.656	0.492
Real Word Blending	1.67	1.08 – 2.57	**0.079**	0.020
Nonword Blending	0.94	0.55 – 1.59	0.895	0.807
Real Word Parsing	2.09	1.14 – 3.84	**0.071**	0.017
Nonword Parsing	0.86	0.49 – 1.50	0.744	0.601
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.69			
$\tau_{00\ Item\_No}$	0.15			
ICC	0.2			
N_{SubID}	27			
N_{Item_No}	21			
Observations	1533			
Marginal R^2 / Conditional R^2	0.221 / 0.379			

Table 61

Logistic Regression Results: PMT Group - SAPA 3 Untreated-Unrelated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.72	0.51 – 1.00	0.154	0.052
Real Word Repetition	1.09	0.68 – 1.73	0.842	0.724
Nonword Repetition	1.13	0.75 – 1.70	0.713	0.565
Real Word Blending	1.15	0.77 – 1.70	0.656	0.493
Nonword Blending	0.71	0.44 – 1.14	0.306	0.16
Real Word Parsing	1.99	1.14 – 3.47	**0.071**	0.015
Nonword Parsing	0.93	0.56 – 1.55	0.891	0.792
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.59			
$\tau_{00\ Item\_No}$	0.1			
ICC	0.17			
N_{SubID}	27			
N_{Item_No}	25			
Observations	1850			
Marginal R^2 / Conditional R^2	0.112 / 0.265			

Table 62

Logistic Regression Results: PMT Group - SAPA 3 Treated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.69	0.50 – 0.94	**0.079**	0.020
Real Word Repetition	1.01	0.65 – 1.57	0.978	0.960
Nonword Repetition	1.13	0.77 – 1.67	0.688	0.535
Real Word Blending	1.84	1.26 – 2.68	**0.010**	0.001
Nonword Blending	0.85	0.54 – 1.34	0.651	0.485
Real Word Parsing	1.90	1.12 – 3.21	**0.071**	0.016
Nonword Parsing	0.95	0.59 – 1.54	0.898	0.828
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.15			
$\tau_{00\ SubID}$	0.53			
ICC	0.17			
N_{SubID}	27			
N_{Item_No}	39			
Observations	2678			
Marginal R^2 / Conditional R^2	0.221 / 0.355			

Table 63

Logistic Regression Results: PMT Group - SAPA 3 Untreated-Related Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.68	0.44 – 1.06	0.213	0.090
Real Word Repetition	0.91	0.55 – 1.52	0.842	0.727
Nonword Repetition	1.30	0.83 – 2.05	0.407	0.254
Real Word Blending	1.95	1.25 – 3.04	**0.024**	0.003
Nonword Blending	0.91	0.53 – 1.56	0.842	0.733
Real Word Parsing	2.40	1.30 – 4.43	**0.034**	0.005
Nonword Parsing	0.89	0.51 – 1.55	0.803	0.672
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.68			
$\tau_{00\ Item\_No}$	0.40			
ICC	0.25			
N_{SubID}	27			
N_{Item_No}	21			
Observations	1442			
Marginal R^2 / Conditional R^2	0.280 / 0.458			

Table 64

Logistic Regression Results: PMT Group - SAPA 3 Untreated-Unrelated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.73	0.50 – 1.07	0.236	0.105
Real Word Repetition	1.00	0.58 – 1.71	0.996	0.996
Nonword Repetition	0.98	0.61 – 1.59	0.969	0.944
Real Word Blending	1.34	0.84 – 2.12	0.361	0.218
Nonword Blending	0.66	0.38 – 1.15	0.290	0.145
Real Word Parsing	2.67	1.39 – 5.14	**0.024**	0.003
Nonword Parsing	0.75	0.41 – 1.35	0.484	0.335
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.81			
$\tau_{00\ Item\_No}$	0.11			
ICC	0.22			
N_{SubID}	27			
N_{Item_No}	25			
Observations	1712			
Marginal R^2 / Conditional R^2	0.146 / 0.333			

Table 65

Logistic Regression Results: SFA Group - SAPA 3 Treated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.92	0.69 – 1.24	0.732	0.588
Real Word Repetition	1.59	1.10 – 2.27	**0.064**	0.012
Nonword Repetition	1.57	1.06 – 2.33	**0.091**	0.025
Real Word Blending	1.11	0.65 – 1.90	0.823	0.698
Nonword Blending	0.58	0.33 – 0.99	0.148	0.047
Real Word Parsing	1.39	0.84 – 2.27	0.345	0.197
Nonword Parsing	1.48	0.83 – 2.63	0.336	0.188
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.11			
$\tau_{00\ SubID}$	0.56			
ICC	0.17			
N_{SubID}	28			
N_{Item_No}	81			
Observations	5720			
Marginal R^2 / Conditional R^2	0.227 / 0.357			

Table 66

Logistic Regression Results: SFA Group - SAPA 3 Untreated-Related Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.85	0.62 – 1.17	0.476	0.321
Real Word Repetition	1.59	1.09 – 2.30	**0.071**	0.015
Nonword Repetition	1.46	0.97 – 2.18	0.176	0.067
Real Word Blending	0.92	0.53 – 1.59	0.868	0.762
Nonword Blending	0.60	0.34 – 1.05	0.186	0.073
Real Word Parsing	1.50	0.91 – 2.49	0.252	0.114
Nonword Parsing	1.55	0.86 – 2.81	0.290	0.146
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.18			
$\tau_{00\ SubID}$	0.55			
ICC	0.18			
N_{SubID}	28			
N_{Item_No}	41			
Observations	2860			
Marginal R^2 / Conditional R^2	0.203 / 0.349			

Table 67

Logistic Regression Results: SFA Group - SAPA 3 Untreated-Unrelated Items Immed. Post-Tx

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.96	0.69 – 1.34	0.895	0.812
Real Word Repetition	1.33	0.92 – 1.91	0.272	0.129
Nonword Repetition	1.46	0.98 – 2.17	0.171	0.062
Real Word Blending	1.19	0.70 – 2.04	0.677	0.517
Nonword Blending	0.43	0.24 – 0.75	0.024	0.003
Real Word Parsing	1.48	0.90 – 2.42	0.257	0.121
Nonword Parsing	3.05	1.70 – 5.49	**<0.001**	<0.001
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.49			
$\tau_{00\ Item\_No}$	0.16			
ICC	0.17			
N_{SubID}	28			
N_{Item_No}	25			
Observations	1824			
Marginal R^2 / Conditional R^2	0.336 / 0.446			

Table 68

Logistic Regression Results: SFA Group - SAPA 3 Treated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.70	0.51 – 0.95	**0.080**	0.021
Real Word Repetition	1.37	0.94 – 1.99	0.224	0.097
Nonword Repetition	1.63	1.08 – 2.46	**0.078**	0.019
Real Word Blending	1.06	0.61 – 1.85	0.909	0.842
Nonword Blending	0.71	0.40 – 1.25	0.381	0.233
Real Word Parsing	1.37	0.82 – 2.30	0.381	0.233
Nonword Parsing	1.35	0.74 – 2.47	0.482	0.328
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.14			
$\tau_{00\ SubID}$	0.60			
ICC	0.18			
N_{SubID}	28			
N_{Item_No}	81			
Observations	5600			
Marginal R^2 / Conditional R^2	0.188 / 0.337			

Table 69

Logistic Regression Results: SFA Group - SAPA 3 Untreated-Related Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.64	0.45 – 0.91	**0.066**	0.013
Real Word Repetition	1.36	0.90 – 2.05	0.288	0.142
Nonword Repetition	1.65	1.06 – 2.57	**0.092**	0.027
Real Word Blending	1.23	0.67 – 2.25	0.665	0.505
Nonword Blending	0.57	0.31 – 1.05	0.187	0.074
Real Word Parsing	1.47	0.84 – 2.57	0.327	0.179
Nonword Parsing	1.42	0.74 – 2.71	0.449	0.293
Random Effects				
σ^2	3.29			
$\tau_{00\ Item\_No}$	0.17			
$\tau_{00\ SubID}$	0.68			
ICC	0.21			
N_{SubID}	28			
N_{Item_No}	41			
Observations	2800			
Marginal R^2 / Conditional R^2	0.210 / 0.373			

Table 70

Logistic Regression Results: SFA Group - SAPA 3 Untreated-Unrelated Items at Follow-Up

Coefficient	*Odds Ratios*	*CI (95%)*	*B-H Corrected p-Value*	*p-Value*
Intercept	0.97	0.66 – 1.41	0.913	0.863
Real Word Repetition	1.36	0.89 – 2.08	0.295	0.153
Nonword Repetition	1.36	0.84 – 2.20	0.349	0.205
Real Word Blending	1.61	0.85 – 3.08	0.290	0.146
Nonword Blending	0.57	0.30 – 1.07	0.201	0.081
Real Word Parsing	1.19	0.68 – 2.10	0.696	0.544
Nonword Parsing	1.95	1.00 – 3.80	0.152	0.050
Random Effects				
σ^2	3.29			
$\tau_{00\ SubID}$	0.68			
$\tau_{00\ Item\_No}$	0.17			
ICC	0.21			
N_{SubID}	27			
N_{Item_No}	25			
Observations	1800			
Marginal R^2 / Conditional R^2	0.250 / 0.404			