

O0832-R

Dosage and Predictors of Naming Treatment Response in Aphasia

NCT02005016

Date: August 30, 2019

A. Background and Significance

Language and communicative impairments following stroke (aphasia) affect more than 30% of stroke survivors, with an incidence of over 180,000 new cases annually and a prevalence greater than that of Parkinson's Disease. The consequences of aphasia are far reaching and can affect psychosocial adjustment, family role participation, vocational opportunities and the ability to function independently in society. Recent estimates suggest that VHA outpatient clinics see 2000 new cases of aphasia each year^[26]. Based on estimates of stroke prevalence among veterans^[27] and the prevalence of aphasia in stroke^[28], it can be conservatively estimated that approximately 20,000 enrolled VHA patients and 100,000 US veterans are currently living with the condition. Despite the prevalence of aphasia and its broad social and economic consequences, access to effective aphasia rehabilitation services for community-dwelling veterans is limited.

In response to this need, the VA Pittsburgh Healthcare System GRECC established the Program for Intensive Residential Aphasia Treatment and Education (PIRATE). PIRATE employs a service delivery model that provides 4 weeks of intensive, evidenced-based aphasia treatments to community dwelling veterans while they reside in a residential setting on the VAPHS campus. PIRATE currently serves 18 veterans per year in bi-monthly sessions. Resource limitations associated with PIRATE and aphasia treatment in general require that treatments be offered to those veterans most likely to benefit from them, in the most cost-effective doses.

This proposed research directly addresses these inter-related issues by investigating response to intensive semantically-oriented naming treatment for veterans with aphasia. Naming is the most common domain of impairment among persons with aphasia.^[18] First, the research seeks to identify neural, cognitive, and psycholinguistic factors that predict positive treatment response. It builds on promising results indicating that individuals with a range of language-impairment profiles may benefit from semantically-oriented naming treatment^[12, 19], and that cognitive function^[10] as well as degree of cortical and subcortical left-hemisphere damage may be predictive of naming treatment outcomes^[21]. Identifying factors that are prognostic of positive treatment response is critical for effective allocation of limited healthcare resources. Second, the research examines the dose-response relationship for naming therapy when delivered on an intensive schedule. An intensive schedule is widely argued to be important to maximizing treatment gains^[29, 30] and it is increasingly being adopted as a treatment delivery model^[5], with at least 11 organizations in North America currently offering intensive aphasia therapy programs. Defining the dose-response relationship for aphasia therapy is a critical first step in determining optimum dosage levels. Providing answers to these two interlocking questions – for whom is aphasia therapy effective, and how much of it is needed? – can set transformative new standards for how aphasia treatment is delivered within VHA.

A.1 Rehabilitation of communicative and naming impairments in aphasia

A.1.1. Stimulation of semantic representations relevant to naming: Semantic Feature Analysis

As noted above, naming is the most common domain of impairment among persons with aphasia. Over 85% of PIRATE participants to date have received treatment focused on spoken word production, and word-finding difficulty is a salient and functionally significant disability for many veterans living with aphasia. Psycholinguistically-motivated treatment approaches targeting word-production deficits have shown significant evidence of efficacy. These approaches facilitate successful naming by stimulating one or more stages of the word-production process, engaging clients in activities designed to strengthen semantic or phonological processes through repeated structured practice. These treatment activities are intended to make effortful encoding processes more automatic for persons with aphasia, facilitating successful naming performance in therapy and broader communicative contexts. Remediating these individuals' word-production deficits is a critical component in the broader effort to improve communicative function among veterans with aphasia.

The treatment that will be administered to all participants in the proposed research is Semantic Feature Analysis (SFA)^[14, 31]. SFA is a naming treatment in which participants are first asked to name a target picture. Then, regardless of success, they are assisted by the clinician in generating six semantic features of the target. Following this, if the participant has not yet successfully produced the target word, the clinician provides it and requires the participant to repeat it and review its features. SFA is based on spreading-activation models of semantic processing^[32] and is consistent with current models of word production, which posit that word production begins with activation of conceptual semantic representations, which in turn activate corresponding lexical representations^[33, 34]. Under this view, early stages of naming an object involve activation of the object's semantic features. This activation spreads through the semantic network, activating concepts sharing those features and associated lexical representations. The representation with most activation is ultimately selected for phonological encoding. The rationale for SFA is that activation of the semantic network surrounding a target item should help increase the activation of the target above the threshold required for successful production^[14].

SFA was originally proposed as a method for remediating the processes underlying word retrieval^[35], i.e., a treatment that should not require conscious use of a compensatory strategy. However, it has also been noted that in some cases SFA may function as a word retrieval strategy^[15, 18, 36]. Regardless of whether SFA operates by a restorative or a compensatory mechanism in any given case, active generation of semantic features by participants appears to be an important component of the approach^[37]. SFA will be used as a remedial treatment in the proposed research, in line with the original formulation of SFA and with PIRATE's programmatic focus on impairment-oriented aphasia treatment.

SFA has several attributes that make it well suited for use in the proposed research. First, there is good preliminary evidence that SFA is associated with lasting improvement in the naming ability of persons with aphasia. SFA or closely related approaches have been applied to at least 21 participants with aphasia across 10 single-subject design studies reported in the peer-reviewed literature^[14-17, 36, 38-43]. Positive treatment effects on production of trained items were observed in 20 of these cases. Interestingly, and potentially relevant to Hypothesis 1b of the current proposal, the one individual that failed to respond to SFA had poor performance relative to other individuals with aphasia on (unspecified) nonverbal cognitive tests.^[38] Clear maintenance of improved naming was observed post-treatment in 13 of the 18 cases in which it was assessed. In two of the five cases where treatment gains were not clearly maintained, there was maintenance of some but not all trained behaviors.^[39, 40] Furthermore, as noted by Boyle^[37], in four of these five cases the participants were administered an atypical variant of SFA: participants were not required to generate semantic features but instead reviewed features provided by the clinician. This finding again points to the importance of the active generation of semantic features by participants.

Second, there is evidence that SFA treatment effects generalize to production of untreated words (i.e., response generalization) and to connected speech (i.e., stimulus generalization). Among the 20 individuals showing positive effects for trained items, response generalization was assessed in 19. In 15 of these cases there was positive evidence of response generalization. Three cases in which generalization to untreated items was not observed came from a single study^[41]. In this study, treatment items were selected from very specific closed-set narrative contexts (e.g., recounting how to change a tire), which may have contributed to poor response generalization. A distinguishing feature of the remaining case that failed to show response generalization was that it was the only one that received treatment on verb rather than noun targets^[42]. It is noteworthy that in 12 cases, improved naming generalized to items that were unrelated to any of the treated items. Also of note, generalization to untreated items was less consistent in 7 cases where bilingual participants reviewed but did not generate semantic features during treatment across two languages^[39, 40]. Stimulus generalization to connected speech was evaluated in 10 cases, with positive results in nine^[14-16, 36, 38, 41, 42]. These findings suggest that SFA shows promise for inducing broad improvement in naming performance.

A third factor that makes SFA suitable for the current proposal is that it appears to be appropriate for a relatively broad range of patients with aphasia, in terms of both language-impairment profile and severity. Patients responding positively to SFA have encompassed most of the commonly observed classical aphasia categories, including Broca's, Wernicke's, anomic, conduction, transcortical motor, and "fluent". These patients have also encompassed a wide range of severity, with Western Aphasia Battery Aphasia Quotients ranging from 27 to 90.6. In two cases, participants were also reported to have apraxia of speech^[39, 43]. Most importantly, the 21 participants discussed above included not only individuals with psycholinguistic profiles suggestive of lexical-semantic word production impairments^[16, 36, 41-43] but also with combined lexical-semantic and phonological impairments^[16, 41-43]. There is no specific evidence to date regarding the efficacy of SFA for patients with predominantly phonological impairments. However, studies using other semantically-oriented treatments suggest that they can result in improved naming in such patients in at least some cases^[12, 44, 45].

These latter findings are relevant to Hypothesis 1a, that semantically-oriented naming treatments will benefit individuals with lexical-semantic, phonological, and mixed naming impairments. Data from Wambaugh and from Abel^[12, 44, 45] suggest that semantic cueing hierarchies may benefit even individuals with primarily phonological naming deficits. In fact, close inspection of the data from two of these studies reveals that on average, participants with primarily phonological impairments ($n=4$) responded more favorably to semantically-oriented treatment than patients with primarily semantic impairments ($n=4$). The phonologically impaired participants showed an average improvement of 37 percentage points ($sd = 17$), while the semantically impaired participants showed an average improvement of 20 percentage points ($sd = 9$). An additional participant classified as having mixed semantic and phonological impairments demonstrated an intermediate response, with an average improvement of 33 percentage points across two sets of items. Although this comparison is based on limited data, it does suggest that psycholinguistic profile is unlikely to be a strong predictor of response to SFA treatment. Furthermore, there is reason to expect that improved access to lexical-

semantic representations may stimulate subsequent phonological processes in word production. In all models of word production, successful activation of a target lexical-semantic representation feeds forward to activate phonological representations needed to produce the associated word form. Successful stimulation of lexical-semantic processing by SFA may therefore ease later phonological processing, by alleviating bottlenecks occurring at earlier levels of processing.

While it is often assumed that individuals with lexical-semantic deficits will benefit most directly from semantically-oriented naming treatments, this assumption has rarely been directly tested in the literature. More generally, the question of which individuals are most likely to benefit from a specific course of aphasia treatment remains open. Comparing SFA treatment response across impairment-type groups – those with primarily lexical-semantic impairments, those with primarily phonological impairments, and those with mixed impairments – will provide important evidence regarding these questions in the aphasia treatment literature.

A.1.2 Dosage of SFA and facilitation of naming function

Existing studies of SFA, like most studies of impairment-focused aphasia therapy, have provided relatively low dosages of the therapy, e.g., 3 hours per week for four weeks^[14, 15]. Participants in the research proposed below will receive much higher doses of SFA: 20 hours per week for 4 weeks. This dosage of SFA is feasible for veterans enrolling in the study: as reported in B.2 below, previous PIRATE participants have responded well to similar dosages of impairment-focused treatment and of SFA specifically. Equally importantly, this high dosage level is critical for achieving the study aims above, in several respects. First, high dosages of behavioral therapy are important for inducing significant and lasting behavioral changes. Evidence from the educational literature indicates that extended training – to and even past criterion for a learned behavior – improves short-term and long-term retention^[46-48], results in greater endurance and resistance to distraction^[49, 50], and increases generalization to related behaviors^[47, 51]. Findings from the motor therapy literature also show that high-dosage therapy, particularly when administered on an intensive schedule, leads to improved function and retention^[52-54]. Consistent with these findings, work examining the effects of constraint-induced language therapy^[30, 55, 56] suggests that massed practice and multiple repetitions of specific training activities is important for maximizing gains. Furthermore, both constraint-induced and traditional model-based treatments appear to benefit from such high-dosage administration^[57]. These findings all indicate that high dosages of aphasia therapy – involving multiple repetitions of therapy activities, training targets to or past criterion – are crucial for maximizing improvements, retaining those gains, and generalizing to untrained stimuli and other behaviors.

Second, high treatment dosages like those proposed here are needed to induce neuroplasticity in treated individuals. The aim of aphasia therapy is to cause cortical reorganization, reflected in both grey- and white-matter changes. The neuroanatomical hypotheses above critically lean on the assumption that treatment will induce such neuroplasticity (viz. ^[11, 29, 57]). There is significant evidence from animal models that cortical reorganization of somatosensory areas may occur in response to extensive experience^[58, 59], and that increasing the amount of sensory input facilitates this reorganization^[58, 60, 61]. Furthermore, extended training of the sort described above (training to or beyond criterion) appears to be crucial for promoting such lasting cortical adaptation. Kleim and colleagues^[62] found that while adult rats' behavioral gains in response to motor training plateaued after 3 days, motor-cortex changes appeared only after 7 days or more of training. Relatedly, changes in MEG^[57] and EEG measures^[29] in response to constraint-induced aphasia therapy emerge after 30 hours or more of treatment. Extended repetition of therapy activities promotes lasting neural changes, of the sort assumed to underlie the structure-function correlations explored under Specific Aim 2.

Third, there is evidence that higher dosage of training activities specific to SFA yields greater improvement for patients receiving SFA. SFA facilitates naming through stimulation of the semantic system, by prompting participants to generate semantic features associated with treated items. As discussed above, SFA shows evidence of acquisition of treated items as well as generalization to related untreated items, as do other naming treatments that stimulate the semantic system^[63]. However, the speed of acquisition and the magnitude of generalization in SFA appear to depend on the number of repetitions of treatment activities patients complete. In two separate studies of SFA, patients reached criterion more quickly and/or showed stronger evidence of generalization when they were treated with fewer training exemplars (as few as 7) compared to more exemplars (up to 20)^[15, 16]. Interestingly, this trend is mirrored in other studies of naming treatment in aphasia. In a recent meta-analysis of published naming treatment studies, Snell and colleagues found that the magnitude of treatment response was moderately negatively correlated with the number of exemplars used in treatment^[64]. Both these findings point to the value of repetition in maximizing treatment gains: a larger number of treatment exemplars reduces the number of repetitions of each exemplar, as well as the number of treatment activities (such as feature generation) patients can complete in response to each exemplar. Parallel

to the educational-literature findings above, extended training using SFA -- involving many repetitions of training activities and exemplars -- appears to be important for effective acquisition and generalization to untreated stimuli.

A.1.3 Summary

Word-finding deficits represent a salient impairment for veterans with aphasia, and they are a common object of aphasia treatment in PIRATE and aphasia therapy more generally. Research examining response to treatment designed to remediate these deficits should employ well-defined naming treatment methods which target a single level of language function and are hypothesized to facilitate specific word-production processes^[37]. SFA is one such naming treatment, which has successfully been used with individuals representing a range of aphasia severities, classifications, and language-impairment profiles. Comparing the response of individuals with different language-impairment profiles to this single well-defined treatment will provide important evidence regarding whether its benefits are specific to one group of individuals or to impairments of one word-production process. Furthermore, providing therapy at high dosage levels is important for generating lasting and measurable improvements in naming function, in order to test study hypotheses regarding the dose-response relationship and predictors of the size of treatment-induced change.

A.2 Predictors of treatment response

In addition to determining whether groups of individuals (classified by language-impairment profile) will respond to a specific naming treatment, it is also important to identify factors that predict whether particular clients will respond to treatment. Identifying such factors can inform decisions about treatment candidacy and treatment type. A range of candidate factors have been identified, including demographic variables (age at onset of aphasia, education level, pre-morbid intelligence), medical variables (general health, presence of co-morbid conditions), and impairment-related factors (time post-onset of aphasia, aphasia type, aphasia severity, auditory comprehension). There is positive evidence of the effects of at least some of these factors: Robey^[1] reports that studies treating moderately severe aphasia patients exhibited larger treatment effect sizes than did studies treating more mildly impaired patients, and that treatment administered in the acute stage (<3 months post-onset) resulted in greater gains than treatment administered in the chronic stage. However, there are also counterexamples to these generalizations. For instance, the sentence-treatment studies exhibiting the largest effect sizes among those surveyed by Beeson and Robey^[65] all involved patients in the chronic phase of recovery. The question of how to identify the individuals aphasia treatment will benefit most thus remains open.

A.2.1 Cognitive factors predicting treatment response

One set of factors which have been shown to be potentially useful predictors of treatment response are non-language cognitive factors. There is a growing body of evidence suggesting that cognitive factors such as executive functioning, visuospatial skills, memory, and attention may carry significant predictive value for aphasia treatment outcomes^[10, 66, 67]. While the exact nature of the relationship between linguistic and nonlinguistic cognitive processes remains unclear, it is undisputed that multiple cognitive processes are important for learning. Encoding, storage, organization, and access of information are all aspects of cognition that are integral to the aphasia treatment process and ultimately impact treatment outcomes^[68].

A few studies have begun to identify specific cognitive factors and neuropsychological measures associated with positive naming-treatment outcomes among people with aphasia. Lambon Ralph and colleagues^[10] administered several neuropsychological and linguistic assessments to 33 individuals who received naming treatment using a phonemic and orthographic cueing hierarchy. Three pre-treatment language assessments (Boston Naming Test [BNT], three-picture version of the Pyramids and Palm Trees Test [PPT], and PALPA 31 word reading subtest) and three cognitive assessments (Test of Everyday Attention Elevator Counting with distraction subtest, Rey figure copy, and Rey Figure delayed recall) were found to correlate significantly with confrontation naming improvements post therapy. Additionally, principal component analysis identified a cognitive factor and a phonological factor that were independent predictors of immediate and long-term therapy gain. The cognitive factor was comprised of measures of executive function (Wisconsin Card Sorting Test), sustained and divided attention (elevator counting with and without distraction), and visuospatial skills/memory (Rey figure copy and recall). The PPT also loaded onto this cognitive factor. Further analysis revealed that the cognitive factor and the language severity rating from the BNT were the best predictors of treatment outcomes, and that phonological factors did not add to the model's predictive power.

Fillingham, Sage, and Lambon Ralph^[66] and Hinckley and Carr^[69] also found executive function abilities to have prognostic value for naming treatment outcomes in aphasia. Both studies found that the Wisconsin Card Sorting Test significantly correlated with language-performance outcomes. However, Hinckley and Carr found this relationship only for participants who received contextually-based treatment instruction (e.g. catalogue ordering), not for those receiving therapy using a skill-based treatment approach (like that used in

impairment-oriented naming therapy like SFA). Hinckley and Carr^[69] also found that performance on the Ravens Colored Progressive Matrices (RCPM) was related to some outcome measures. Interestingly, they also found additional predictive benefit of executive function abilities for the amount of treatment time needed for language skill acquisition. Scores on the WCST and the RCPM were significantly negatively correlated with the amount of treatment time required to reach performance criteria.

Recognition memory and attention skills have also been identified as important predictors of treatment success when contrasting errorful and errorless therapies for anomia. Fillingham and colleagues^[66] found that specific subtests on the Camden Memory Test (topographical and word subtests) were significantly correlated with language outcomes. The authors highlighted the importance of baseline working memory, delayed recall, and attention for the two participants that demonstrated larger treatment effects in the errorful learning condition, suggesting that errorful learning techniques may be most beneficial for patients with stronger memory and attention skills.

Interestingly, the predictive value of these cognitive factors may be stronger for treatment-related gains in specific language domains rather than for broad measures of communicative function. The relationship between nonlinguistic cognitive performance and language recovery as determined by improvements on the Boston Diagnostic Aphasia Examination post treatment was examined by Seniow, Litwin, and Lesniak^[70]. When using the RCPM as the primary measure for executive functioning/abstract thinking abilities in people with aphasia, the authors did not find significant correlations between post treatment scores on the BDAE and the RCPM. However, they did find a significant correlation between patient's performance on tests of nonverbal visuospatial working memory and naming and comprehension scores of the BDAE.

While these findings all point to the importance of cognitive factors in mediating treatment response, it is unknown how these factors relate to the different naming treatment effects highlighted above. In particular, it is unclear how they relate to the likelihood of learning the specific skills or language forms being treated in therapy (acquisition) versus generalizing to untreated forms and situations outside the therapy room (generalization). Acquisition of treated forms or skills is a necessary prerequisite for generalization^[71], but generalization is widely considered the ultimate goal of effective therapy. As noted above, there is good preliminary evidence that SFA is associated with successful acquisition of trained items, and with both generalization to untrained items (response generalization) and to word production in non-therapy context (stimulus generalization). Relatively little research to date has been directed at the question of what factors underlie successful acquisition versus generalization in aphasia therapy. However, there is emerging evidence suggesting that the neural and cognitive mechanisms responsible for the two types of learning may be distinct. Dickey and Yoo^[9] found that behavioral variables associated with positive treatment response were predictive of acquisition of treated forms, but not generalization to untreated but related forms. Furthermore, Meinzer and colleagues^[20] found that brain areas positively related to acquiring treated items in a naming treatment study were not correlated with generalization to untreated words. The relationship between individual differences in cognitive capacities and mastery of treated behaviors or generalization to untreated behaviors is unknown.

A.2.2 Neural correlates of treatment response

Complementing these findings regarding the importance of various cognitive factors in mediating treatment response, recent work using neuroimaging has identified some promising neural correlates of positive treatment outcomes^[21, 25, 72]. These results point to the importance of regions relevant for basic learning and memory (hippocampal regions) and regions relevant to functional reorganization in response to treatment (basal ganglia as well as contralesional right-hemisphere and intact left-hemisphere tissue).

One line of evidence suggests that patterns of damage that lead to greater recruitment of right-hemisphere regions during recovery are associated with positive response to treatment. Parkinson and colleagues^[21] studied the relationship between lesion location and extent and response to two different naming treatments. They collected structural measurements of cortical and subcortical tissue in 15 individuals with aphasia and correlated them with measures of treatment gains. (See Preliminary Studies, below, for a more detailed description of this study.) Parkinson and colleagues found that the extent of left anterior cortical lesion was positively correlated with improvement in action and object naming. This counterintuitive finding suggests that greater loss of left anterior tissue – which may force greater recruitment of homologous right-hemisphere tissue during recovery – was related to more positive treatment outcomes. Consistent with this, they also found that the extent of left basal ganglia lesion was negatively correlated with treatment effects. Given that inhibitory projections from the basal ganglia are helpful in suppressing recruitment of damaged left hemisphere and permit greater recruitment of right-hemisphere tissue^[22], this finding suggests that patterns of damage promoting right-hemisphere recruitment may yield more positive treatment response.

Consistent with these findings from structural imaging, Menke and colleagues^[24] found in a functional imaging study that greater naming improvement immediately post-treatment was associated with activations in the hippocampus and some right-hemisphere areas (precuneus and cingulate). They also found that naming improvements maintained at 8-month follow-up were associated with increased activity in the right Wernicke's area homologue, and less strongly in left peri-lesional areas. Also consistent with these findings, Crosson and colleagues^[25] found that a group of 4 individuals with aphasia whose naming function improved in response to treatment exhibited more right lateralized activation immediately post-treatment, compared to a patient who did not improve. Parallel functional imaging work by Fridriksson and colleagues has found that successful naming performance by individuals with aphasia is associated with greater activation of either peri-lesional left hemisphere tissue or unaffected right-hemisphere tissue^[73, 74]. These findings strongly suggest that right hemisphere and perilesional left hemisphere tissue can play a critical role in successful treatment response.

A second line of evidence suggests that the hippocampus may play a role in treatment-induced recovery from aphasia. As noted above, Menke and colleagues^[24] found that naming improvement was associated with increased activation in the hippocampus. Parallel to this finding, Meinzer and colleagues^[20] found that improvement on treated object naming items among 10 patients with mild to moderate chronic aphasia was negatively correlated with damage to the hippocampus. Volume loss to the left hippocampus was negatively correlated (-0.65) with improvement on treated items immediately after training, and even more strongly negatively correlated (-0.80) at 8-month follow-up. Loss of integrity of left hippocampus-adjacent white matter tracts (measured using fractional anisotropy maps) was negatively correlated with naming improvement after treatment (-.85) and at follow-up (-0.83). These relationships were specific to treated items: while there was some generalization of treatment effects to untreated items, these effects did not correlate with the structural measures of the hippocampus. This evidence suggests that the hippocampus, which is critical for formation and consolidation of episodic memory, is critical for supporting treatment-related aphasia recovery.

These two lines of evidence point to the importance of both cortical and subcortical tissue in facilitating recovery in response to treatment. However, there are conflicting findings regarding neuroanatomical predictors of aphasia recovery, particularly regarding the role of anterior left-hemisphere versus right-hemisphere tissue. Postman-Caucheteux and colleagues^[75] found in an fMRI study of three adults with chronic aphasia and naming impairments that trials with incorrect naming performance were associated with right-hemisphere activation (particularly in anterior regions such as inferior frontal gyrus, the same region recruited by patients showing improvement in a previous study by Crosson and colleagues^[21]). Successful naming by the patients and by a group of unimpaired controls did not elicit right-hemisphere activation, suggesting that right-hemisphere regions may play a maladaptive role in aphasic language performance. Converging with this finding is functional imaging evidence that right-hemisphere recruitment is associated with poorer outcomes in treatment-related recovery. In a study of 16 adults with chronic aphasia, Richter, Miller and Straube^[76] found that positive response to constraint-induced language therapy was associated with decreased activation in right-hemisphere anterior areas (inferior frontal gyrus and insula). Also, in a treatment and neuroimaging study of 26 individuals with aphasia, Fridriksson^[23] reported that improved naming in response to treatment was associated with increased left-hemisphere activation of medial and inferior frontal areas and more posterior parietal areas, rather than increased right-hemisphere activation. There is also fMRI evidence that better spontaneous recovery is associated with greater recruitment of peri-lesional left-hemisphere areas^[77]. Furthermore, Fridriksson^[23] found that lesions in posterior middle temporal regions were associated with poor response to naming treatment, suggesting that damage preventing recruitment of these areas reduces treatment response. Thus, the prognostic value of factors permitting functional reorganization to the right hemisphere is promising but not definitive. Furthermore, there is conflicting evidence that recruiting left-hemisphere regions may be more beneficial for aphasia recovery.

There are also a number of unanswered questions regarding how neuroanatomical factors may be associated with treatment response. First, it is unclear whether the previously identified neuroanatomical predictors of treatment response are predictive of improvement only for treated language behaviors, or whether they are predictive of more general improvements in language function. Studies showing a positive role for right-hemisphere recruitment have usually measured treatment-related improvements in specifically treated language behaviors^[21, 25]. In contrast, studies showing a negative role for right-hemisphere recruitment have typically employed broader measures of language function such as standardized assessments^[76]. Further data are required to settle this issue. Second, the neural correlates of the two types of treatment response (acquisition, generalization) are unknown. There is evidence that intact hippocampal tissue and structural factors promoting reorganization to the right hemisphere are predictive of improvements on treated items/behaviors^[20, 21, 24]. However, there is no evidence to date of regions that are associated with

generalization to untreated items/behaviors. Meinzer and colleagues^[20] found that intact hippocampal tissue was correlated with improvement on treated items, but not with improvement on untreated items. The neural correlates of successful generalization in response to treatment thus remain unknown^[9].

Interestingly, the neuroanatomical correlates of response to a particular form of naming treatment, or of response to a single treatment by different groups of individuals, have not been investigated to date. The fact that similar treatment-neuroanatomy relationships have been detected across studies suggest that these relationships may reflect global patterns of successful reorganization in response to treatment, rather than learning of therapy-specific processes^[22]. However, they may also reflect cortical/subcortical resources that are needed for improvements in specific processes associated with successful word production (e.g., activation or selection of a lexical-semantic representation). Research that delineates the neuroanatomical correlates of positive response to a specific and well-defined treatment will lay the foundation for future studies comparing the neural bases of response to different types of treatment.

A.2.3 Summary

Given that behavioral therapy for chronic aphasia is efficacious but resource-intensive, it is critical to identify factors that will be predictive of who is most likely to benefit from such therapy. Studies examining the role of cognitive factors in aphasia therapy provide evidence that a broad range of factors relevant to learning may be predictive of treatment response. Executive function, visuospatial memory, and attention (both divided and focused) have all been found to predict naming treatment outcomes. However, it is unclear whether these factors are more or less strongly predictive of outcomes for psycholinguistically-motivated naming treatments like SFA than for the cueing hierarchies used in previous studies. It is also unknown whether these cognitive factors are predictive of acquisition of trained behaviors, generalization (stimulus or response), or both.

Studies regarding potential neuroanatomical predictors of treatment response provide similarly encouraging but limited evidence for the potential prognostic role of neurological findings for aphasia therapy outcomes. Further research is required to determine the prognostic value of the extent of left-hemisphere lesion (and corresponding recruitment of right-hemisphere homologues) in forecasting likely treatment response. Further research is also required to determine whether neuroanatomical predictors are prognostic only of skills/items treated in therapy, or whether they are also prognostic of generalization to untreated items and to improvements in broader, standardized measures of naming and language function.

A.3 Dosage levels and treatment response in aphasia

There is abundant evidence for the efficacy of behavioral rehabilitation strategies rooted in cognitive neuropsychology and psycholinguistics for improving the language function of adults with aphasia^[1, 2], including naming treatments like those described above. Moreover, the literature also suggests that providing more treatment results in greater improvement in language function^[1, 6, 7, 78, 79]. However, these findings are preliminary, because the total amount of treatment has often been confounded with intensity of treatment. Research is therefore needed to address the effects of treatment amount independent of treatment intensity. Systematically examining the connection between the amount of treatment and the resulting gains is critical for the ultimate goal of identifying optimum dosages for aphasia therapy.

The general finding is that more treatment results in better treatment response. An early survey of aphasia treatment studies^[13] identified treatment duration as one factor relevant to positive treatment response, finding that studies in which patients received at least 3 hours per week for 5 or more months were likely to report improvements. A subsequent large-scale meta-analysis of aphasia treatment studies^[1] reported that the amount of treatment received (measured in weeks in treatment or number of treatment hours) was positively correlated with magnitude of improvement on study measures of language function ($r=.76, .64$, respectively). Similarly, one study of non-intensive, socially-focused group therapy found additional improvements on some performance-based outcome measures after 4 months (80 hours) as opposed to 2 months (40 hours) of treatment^[7]. These findings all suggest that there is a positive dose-response relationship in aphasia therapy: more treatment is likely to lead to better outcomes. However, there is considerable variation across these studies in the individuals treated and the outcome measures used, making comparison of treatment gains in response to varying amounts of treatment difficult. In addition, there are some findings suggesting that more time spent in treatment may not result in greater response. A study surveying treatment research using one sentence-level treatment approach^[80] found that auditory comprehension scores predicted treatment response, but that the number of treatment sessions was not correlated with the magnitude of treatment gains^[9].

Recent studies of intensive aphasia treatment (defined here as >3 hours per day for at least 2 weeks) have also drawn attention to the role that intensity along with treatment duration may play in maximizing treatment gains^[5, 6]. Intensive treatment is a tenet of many current aphasia treatment approaches, such as constrained-induced language therapy^[30]. Consistent with this emphasis on intensity, Robey^[1] reports that

studies delivering treatment at higher intensities exhibited larger effects. These findings are a significant part of the motivation for the intensive delivery schedule used in the PIRATE program at VAPHS. However, those studies which provided more intensive treatment have also provided more of it: the improvements seen in these studies could be due in part to greater treatment dosage^[7]. Furthermore, these studies have not systematically examined treatment amount, with most providing approximately 30 hours of treatment over a 2 week period^[29, 30, 55, 57, 81]. It is not known whether additional treatment on the same schedule would produce additional gains or would demonstrate diminishing returns.

Also unknown is how the amount of treatment received relates to acquisition and generalization in response to that treatment. As noted above, there is emerging evidence that the neural and cognitive mechanisms responsible for the two types of learning may be distinct^[9, 20]. In addition, there is evidence from the motor-learning literature suggesting that retention of trained behaviors and transfer to untrained behaviors may respond differently to different dosage levels^[82]. For non-speech motor learning, greater amounts of practice typically yield better retention of treated skills^[83, 84]. However, large amounts of practice may interfere with transfer to untreated skills, particularly under conditions of constant practice. Furthermore, the motor-learning literature suggests that the slope of performance curves for skill learning may vary widely even for a single behavior, depending on the choice of dependent measure^[85]. The relationship between treatment amount and mastery of treated skills or generalization to untreated skills thus remains unknown.

As evidence for the efficacy of language rehabilitation for adults with chronic aphasia accumulates, especially when delivered on an intensive schedule, it becomes increasingly important to determine in what amounts such treatment should be provided. However, current evidence does not permit researchers or policymakers to determine optimum dosage amount, or even whether the number of hours of treatment determines the size of treatment gains. Research studies that hold the delivery schedule constant while systematically measuring the effects found for different dosage levels are required to answer this question. Finally, such studies should ideally examine the dose-response relationship for not only acquisition but also generalization of treated skills, given the importance of generalization for broadly effective aphasia therapy.

B. Preliminary Studies

The proposed study builds on two parallel lines of work carried out by the investigators: a clinical program providing intensive impairment-based aphasia therapy to veterans with aphasia, and research programs investigating the relationship between neuroanatomical variables and aphasia treatment response.

B.1 The Program for Intensive Residential Aphasia Treatment and Education

In January 2009, the VAPHS GRECC initiated the Program for Intensive Residential Aphasia Treatment and Education (PIRATE), a novel service delivery program model designed to eliminate barriers preventing community-dwelling veterans with aphasia from accessing intensive language rehabilitation services. At its inception, the program provided approximately 80 hours of aphasia treatment over a 3-week period. Currently, the schedule has been expanded to approximately 110 hours over a 4-week period. After being evaluated for candidacy, veterans are enrolled in groups of three and reside in residential villas on the VAPHS H.J. Heinz Campus for the duration of the program.

Treatment provided within the context of PIRATE is grounded in current psycholinguistic and cognitive-neuropsychological approaches to aphasia therapy. Patients' language impairments are described in reference to current models of language processing, and treatments motivated by these models are applied to remediate these well-specified deficits. Treatments used in PIRATE also tend to rely heavily on practice and repeated drilling of tasks organized around relatively discrete units of language. While the particular stimuli used may have social relevance for a given patient, the focus of treatment is on improving the underlying cognitive-linguistic processes, rather than on language performance in a particular social context. The proposed research thus fits naturally with the treatment approach employed in PIRATE, with its core emphasis on repeated practice focused on basic cognitive-linguistic processes.

In complement to intensive impairment-focused treatment, PIRATE has also included elements of more socially-oriented approaches to aphasia treatment. These include group activities focused on conversation practice and conversational coaching with patients' significant others or volunteer partners. The purpose of these activities is to promote generalization of performance gains made in the context of psycholinguistically motivated treatments to other communicative contexts. Group activities occur once weekly during each PIRATE session, for approximately 1.5 hours. However, impairment-focused therapy constitutes the large majority of the 25 hours of weekly treatment routinely provided in PIRATE. Because this study focuses on how dosage levels and neuroanatomical and behavioral factors predict response to a well-motivated and specified treatment (SFA), and because of the potential confound posed by additional group treatment activities, we will deliver individual SFA treatment exclusively in the research proposed below.

B.1.1 PIRATE participants and outcomes to date

PIRATE employs both performance-based and patient-reported outcome measures. Performance-based measures are currently collected at the initial evaluation (time 1), program entry (time 2), program exit (time 3), and follow up (time 4), approximately four weeks post-exit. These outcome measures provide pilot data that are directly relevant to the current proposal in two ways. First, they suggest that the aphasia treatment provided in the context of PIRATE is effective. Second, the natural history of the program permits a preliminary evaluation of the dose-response question raised in Specific Aim 2, given that the first six PIRATE sessions were three weeks in duration, and the remaining ten sessions were four weeks in length.

To date, PIRATE has served 53 unique veterans. In one of these cases, the veteran ended program participation early due to medical issues. Descriptive demographic and clinical data are presented in Table 1 for the 52 unique cases in which the veteran completed PIRATE.

Months Post Onset at Program Entry, Median (IQR)	23 (29)
Age at Program Entry, Mean (SD)	58 (13)
Years of Education, Mean (SD)	14 (2.4)
Gender, N male	49
PICA Overall Score, %ile rank (SD) of mean score	51 (15)
Etiology of Aphasia, N	
Left Hemisphere Stroke	48
Bilateral Stroke	1
Left Penetrating Head Injury	1
Closed Head Injury	1
TBI + LH Stroke	1

Table 1. Demographic and clinical characteristics of veterans who have completed PIRATE (n =43).

B.1.2 Analysis of Preliminary Data 1: Effectiveness of PIRATE

The performance-based outcome measure most frequently used in PIRATE has been the Porch Index of Communicative Ability (PICA^[86]). For the PICA, we have complete data at times 1-3 for 39 cases. We excluded from the present analysis four patients who were less than 6 months post onset at the time of initial evaluation, in order to limit the influence of physiological recovery across the baseline interval. We converted PICA overall raw scores to standardized unit normal scores (z-scores) via the percentile tables published in the PICA manual. These z-scores were then analyzed using repeated-measures ANOVA, with planned contrasts between successive time points to test the hypotheses that there would be no difference between times 1 and 2, but significant positive change between times 2 and 3. The analysis demonstrated non-significant change between initial evaluation and program entry ($F[1,34] = 2.574$, $MSE = 0.009$, $p = 0.118$, $\eta_p^2 = 0.07$) and significant improvement between program entry and program exit ($F[1,34] = 56.364$, $MSE = 0.01$, $p < 0.001$, $\eta_p^2 = 0.624$). Mean values at each assessment point are displayed in Figure 1. To guard against bias resulting from exclusion of cases with missing data, we repeated these analyses with missing observations replaced by the mean of each patient's available observations. This strategy was conservative in that it is likely to overestimate positive change across the baseline interval (initial evaluation to program entry) and underestimate positive change across the treatment interval (program entry to program exit). The results of this re-analysis were similar to the original analysis, with the expected larger improvement across the baseline interval:

There was significant improvement from initial evaluation to program entry ($F[1,40] = 6.042$, $MSE = 0.014$, $p = 0.018$, $\eta_p^2 = 0.131$) and relatively greater improvement between program entry and program exit ($F[1,40] = 53.168$, $MSE = 0.013$, $p < 0.001$, $\eta_p^2 = 0.571$). The observed pattern of minimal change across the baseline interval coupled with significant improvement across the treatment interval supports the effectiveness of PIRATE in treating aphasia. In addition, these changes on this broad measure of language performance likely underestimate the magnitude of treatment-related changes in the proposed research. The PICA samples communicative functioning at a broad level, encompassing performance not only in targeted domains (such as naming)

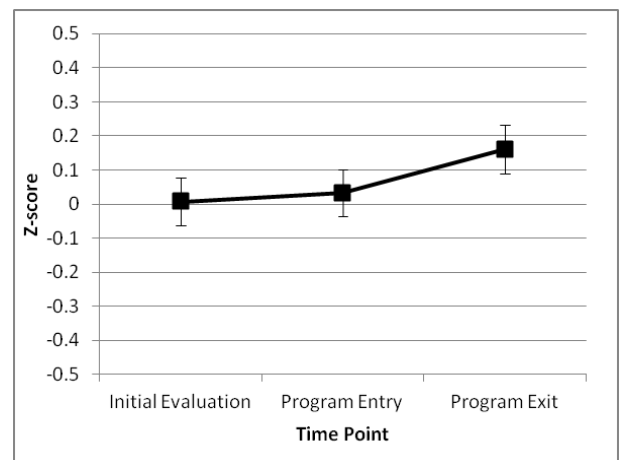


Figure 1. Mean PICA z-scores for n=35 patients with complete data available at three assessment points. Error bars indicate +/- 1 standard error.

but also in domains not targeted in therapy, which are unlikely to show improvement. The language performance assessments to be used in the proposed research measure performance on naming behaviors which are the object of treatment, rather than a broad index of both treated and untreated behaviors. Consequently, we may expect that the intensive delivery schedule followed in PIRATE and in the proposed research will result in even larger gains on study measures than those described here.

B.1.3 Analysis of Preliminary Data 2: Dose-Response Relationship

As noted above, the first six PIRATE sessions were three weeks in length, and the remaining sessions have been four weeks in length. Comparison of outcomes between groups receiving 3 versus 4 weeks of treatment permitted us to estimate the marginal benefit of an additional week of intensive aphasia treatment in terms of a broad performance-based outcome. To derive these estimates, we conducted ANCOVAs on the PICA overall z-scores, with the post-treatment scores as the dependent variable, treatment group (3-week vs. 4-week) as the independent variable, and pre-treatment scores collected at program entry as the covariate.

For the PICA, data at program entry and program exit were available for 15 veterans who received three weeks of treatment and 27 veterans who received four weeks. The demographic and clinical characteristics of each group are presented in Table 2.

	3-Week Group (n=15)	4-Week Group (n=27)
Months Post Onset at Program Entry, Median (IQR)	19 (38)	21 (31)
Age at Program Entry, Mean (SD)	57 (16)	55 (13)
Years of Education, Mean (SD)	14 (3)	14 (2)
Gender, % male	100	89
PICA Overall Score, %ile rank (SD) of mean score	47 (17)	58 (15)
Etiology of Aphasia, N		
Left Hemisphere Stroke	14	17
Bilateral Stroke	1	0
Left Penetrating Head Injury	0	1
Closed Head Injury	0	1

Table 2. Demographic and clinical characteristics of the 3-week and 4-week subgroups of PIRATE cases.

The analysis revealed a non-significant benefit for four weeks of treatment relative to three weeks ($F[1,39] = 0.917$, $MSE = 0.011$, $p = 0.344$, $\eta_p^2 = 0.023$, Cohen's $d = 0.37$). The difference in overall PICA scores at program exit, corrected for PICA score at program entry, was 0.035 standard score points. Interestingly, this estimated value of one additional week of treatment in terms of PICA overall z-score is approximately equal to 33% of the average change score for the three week group and 25% of the average change score for the four-week group. This estimate suggests that the relationship between improvement on broad measures of language performance and total time in treatment may be linear. It is unknown whether a similar trend may hold for measures of behaviors specifically targeted in therapy. However, as noted above, we may expect that treatment-related gains will be larger for measures that more narrowly sample the targeted behavior (naming and word production, in the proposed research).

B.2 SFA in the context of PIRATE

SFA is already one of the treatments administered on an intensive schedule within PIRATE to remediate participants' naming impairments. PIRATE participants receive 22.5 hours weekly of treatment focused on repeated drills targeting specific cognitive-linguistic processes. Veterans who choose to enroll in PIRATE (the population from which study participants will be drawn) are thus able to tolerate and remain engaged during high dosages of impairment-focused, drill-oriented treatments like SFA, at levels even higher than those proposed below.

Veterans who have been administered SFA as part of their treatment in PIRATE have received up to 37.25 hours of SFA (mean 16.2 hours), delivered in one or two 90-minute sessions per day, for up to four weeks. These veterans have responded positively to high levels of SFA: for those veterans for whom detailed probe information is available, mean probe accuracy for treated items increased from 30% at baseline to 72% at the time of their final probe. This positive treatment response is larger than that commonly reported in the SFA literature, and it suggests that extended SFA training administered intensively may result in significant improvement in naming function.

Furthermore, we present more detailed information for one PIRATE participant who recently received SFA as part of his overall treatment plan. His response to an intensive trial of SFA treatment provides pilot data

regarding the efficacy and the feasibility of the proposed design. The veteran presented with significant naming impairments on the naming portion of the CAT, as well as prominent phonological impairments. Demographic and language-testing data for this veteran are presented below.

Months Post Onset at Program Entry	37
Age at Program Entry	66
Years of Education	16
PICA Overall Score, %ile rank of mean score	72
Detailed Language Testing	
CAT Spoken Naming	25/58 (T-score = 50)*
CAT Repetition	43/74 (T-score = 49)*
CAT Comprehension of Spoken Words	25/30 (T-score = 51)*
Pyramids and Palm Trees (3-picture version)	51/52 (98%)
PALPA Same-Different Discrimination	63/72 (83%)*
PALPA Phonological Segmentation	33/45 (73%)*

Table 3. Demographic and clinical characteristics of veteran enrolled in SFA protocol

Consistent with the criteria for study participation (see section C.1 below), this veteran presented with naming deficits on the CAT (T-score of 50) with primarily phonological naming errors. His naming impairments would be classified as mixed according to participant classification measures described below (section C.2.3). He presented with impaired repetition (indicative of impaired phonological processing) and impaired performance on PALPA phonological processing subtests. He also exhibited some evidence of impaired lexical-semantic processing, as evidenced by his performance on spoken word-to-picture matching tasks. However, his conceptual semantic system appeared to be relatively intact, as suggested by his performance on Pyramids and Palm Trees.

This veteran received 13 hours of SFA on 10 training items over a six-day period. His response to treatment is summarized in Fig. 2. We probed this veteran’s performance daily on 10 treated items, 10 untreated related items, and 10 untreated unrelated items. He also received less-frequent assessment with short forms derived from Philadelphia Naming Test that are described below in section C.2.2. We note that because these data were collected in the context of the PIRATE clinical program, the veteran was receiving other treatments (Response Elaboration Training and Kendall’s phoneme-based treatment for anomia) during both the SFA baseline and SFA treatment phases presented in Figure 3. Despite the multiple-treatments confound, the probe data suggest a clear acquisition effect for the trained items as well as some generalization to untrained items. There was also improvement on the PNT short forms.

Figure 2. Pilot SFA participant’s treatment response, for acquisition of treatment exemplars (top panel), generalization to related and unrelated words (middle panel), and more general naming function (bottom panel).



Further evidence of this veteran’s improved function can be seen in the video Appendix A. The Appendix presents two representative treatment trials for the same treatment exemplar, drawn from the first and final SFA treatment session. As can be seen in the two samples, the veteran’s ability to name the treatment exemplar as well as to generate associated semantic features has improved during treatment. He has also remained engaged with the treatment and with the repeated treatment activities. These pilot data thus provide additional evidence suggesting that patients will be able to complete the extended course of SFA treatment proposed below, and that the multiple repetitions of SFA treatment activities can induce significant change in function (see section A.2 above). Furthermore, this veteran’s primarily

phonological impairments provide preliminary evidence in favor of Hypothesis 1a, that individuals with phonologically-based naming deficits will also benefit from SFA.

B.3 Correlates of language impairment and treatment response

B.3.1 Neuroanatomical predictors of treatment response

The investigators have also carried out work regarding the relationship between lesion site and treatment response, in particular investigating the role of recruitment of the non-dominant hemisphere during recovery. This work is directly relevant to the current proposal, since it provides preliminary evidence for the hypothesized role of selected brain regions in supporting recovery in response to aphasia therapy. In addition, we have available brain scan images for 6 veterans served by PIRATE, permitting a preliminary evaluation of the relationships between lesion patterns and treatment response relevant to the neuroanatomical hypotheses under Specific Aim 1.

As described briefly above, Parkinson, Raymer, Chang, FitzGerald, and Crosson^[21] studied the relationship between lesion location and extent and treatment effects in a sample of 15 patients who had received one of two semantically-oriented naming treatments. There were two major findings. First, when controlling for the extent of basal ganglia lesion, the extent of left anterior cortical lesion was positively correlated with improvement in action naming ($r = 0.821$, $p = 0.002$) and object naming ($r = 0.858$, $p < 0.0005$) on treated items (correlations with untreated items and/or standardized outcome measures were not reported). Second, when controlling for extent of left anterior cortical lesion, the extent of basal ganglia lesion was negatively correlated with treatment effects ($r = -0.785$ and -0.749 for actions and objects, respectively, both p 's ≤ 0.01). Based on these results, Parkinson, Crosson, and colleagues hypothesized that excessively “noisy” activity of the left anterior cortex may prevent adaptive re-organization from occurring in the non-dominant hemisphere (as well as in non-lesioned areas of the dominant hemisphere) to support recovery. They proposed that this interference may be reduced in the presence of more extensive lesions to the left anterior cortex and with sparing of the left basal ganglia, which has been argued to suppress peri-lesional activity during word production^[22]. These findings indicate that increased left-hemisphere involvement and intact left basal ganglia tissue should be associated with better treatment response, since both promote adaptive recruitment of other cortical resources (especially right-hemisphere tissue^[25]).

Preliminary data regarding the relationship between treatment response and lesioned and spared brain regions can also be obtained from the sample of PIRATE participants to date. Brain scan images are available for 6 of the veterans who have participated in PIRATE. Of these six veterans, we have pre/post picture-naming probe data on 4, all of whom had aphasia due to left hemisphere stroke of ≥ 19 months post onset. It was not possible to conduct voxel-based morphometry or voxel-based lesion symptom mapping analyses on these data, but they were descriptively analyzed with respect to the involvement of the regions of interest identified as potential predictors of treatment response above. A neuroradiologist who was blind to the behavioral data examined the scans and coded each case for involvement in each of four regions: the left anterior cortex (defined as the pre-central gyrus plus Brodmann areas 9 and 44-46^[21]), the left posterior cortex (defined as covering posterior medial temporal, inferior parietal, and anterior occipital areas, including Brodmann's areas 37, 39, and 19^[23]), the left hippocampus, and the left basal ganglia. The same scale used by Parkinson and colleagues^[21] was used for this analysis: 0 - no lesion; 1 - equivocal lesion or very small portion has patchy lesion; 2 - half of area has patchy or very small portion solid lesion; 2.5 – total area has patchy lesion or less than half of area has solid lesion; 3 – half of area has solid lesion; 4 – more than half of area has solid lesion; 5 – total area has solid lesion.

Pre-Tx Probe % Correct	Post-Tx Probe % Correct	Probe Change Score	L Ant. Cortex Lesion Rating	L Post Cortex Lesion Rating	L Basal Ganglia Lesion Rating	Hippocampus Lesion Rating
0.50	0.64	0.14	3	4	3	2
0.50	0.80	0.30	3	0	0	0
0.12	0.47	0.35	4	0	1	0
0.39	0.81	0.42	2	0	0	0

Table 4. Treatment response on naming probes and lesion ratings for regions of interest, for a sample of PIRATE participants (n=4).

These findings are from a small sample but are consistent with some of the neuroanatomical hypotheses above. The patient with the greatest degree of left posterior and subcortical involvement (with a 3 and 2 rating

for basal-ganglia and hippocampus involvement, respectively) showed the poorest response to treatment. While this patient had the greatest lesion extent overall, simple lesion size is unlikely to be responsible for his poor treatment outcomes: in separate studies, Meinzer and colleagues^[20] and Baker and colleagues^[87] have found that treatment gains were not correlated with lesion size. All three patients showing better treatment response (naming-probe change scores between .3 and .42) had some degree of left anterior involvement and little or no subcortical involvement, as well as little or no posterior cortical involvement. These findings suggest that greater basal-ganglia and hippocampal involvement^[20, 21] and/or greater left-posterior involvement^[23] may be associated with poorer response to treatment in PIRATE. There is less variability in left anterior cortical involvement in this sample, making conclusions regarding the role of left anterior cortical damage in treatment response difficult. However, the patterns do suggest that positive treatment response may be associated with left anterior cortical damage, provided that the subcortical structures needed for basic memory and learning and for promoting recruitment of non-damaged cortical areas are intact^[21, 22, 25].

These findings therefore provide preliminary evidence that veterans who have responded positively to intensive aphasia treatment in PIRATE have patterns of preserved and damaged brain areas which are in line with the hypotheses in Specific Aim 1, specifically Hypothesis 1c. In particular, they suggest that damage to left-anterior cortex may be associated with positive treatment outcomes. However, they leave open the question of the role of left posterior cortex and subcortical structures in recovery in response to treatment.

B.3.2 Voxel-based lesion symptom mapping for language impairment

The proposed research will use voxel-based lesion symptom mapping (VLSM) to test neuroanatomical hypotheses relevant to Specific Aim 1. VLSM is a method which has been used productively to examine correlations between behavioral variables, such as degree of language impairment or hemispatial neglect^[88], and brain lesion. In VLSM, continuous measures of performance are compared to the lesion status of cortical or subcortical tissue on a voxel-by-voxel basis. This method will be used to examine the relationship between treatment-related changes among PIRATE participants and lesions in brain regions which previous research has identified as being predictive of treatment response^[20, 21, 24].

The investigators have used this method to examine the relationship between verbal fluency and lesion site and extent in a group of 16 adults with aphasia^[89]. Participants' verbal descriptions of the "Cookie Theft" picture from the Boston Diagnostic Aphasia Exam^[90] were measured and coded for their relative fluency along five different dimensions: speech rate, degree of audible struggle, ratio of filler to meaningful material, type-token ratio, and ratio of spoken output to periods of silence (a simple global measure of speech productivity). These fluency measures were then correlated with these individuals' lesions, assessed via a high-resolution structural MRI which had been coded for lesion size and extent by a neuroradiologist. The relationship between speech fluency measures and patterns of brain damage is depicted in Figure 3.

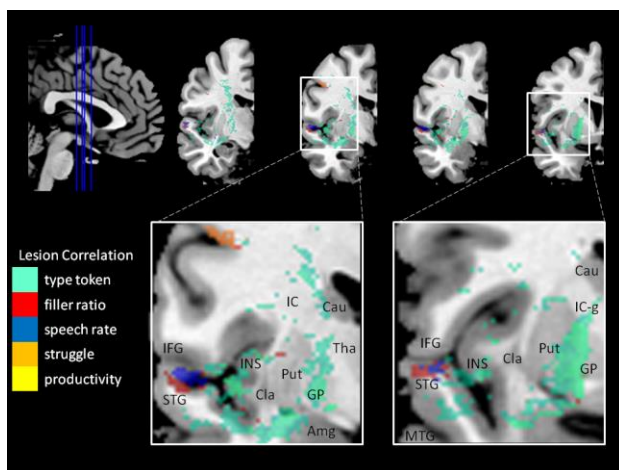


Figure 3: Parametric statistical map showing reliable relationships between speech fluency measures and lesion location in sample of participants with aphasia (Reilly, et al., in prep)

The correlations between continuous fluency measures and the likelihood of lesion in different brain regions revealed important connections between fluency and cortical and subcortical areas. Factors such as speech rate and audible struggle correlated with inferior frontal cortical damage, whereas superior temporal lobe damage predicted high rates of filler material. In contrast, the compromise of frontal subcortical structures was predictive of separate aspects of narrative dysfluency. Specifically, damage to the left insula, caudate, globus pallidus, and putamen were predictive of

impairments in type-token ratio and speech productivity. These findings illustrate the potential utility of VLSM for investigating the connection between treatment-induced changes in language performance measures (such as the PNT) and the degree of cortical and subcortical involvement, as hypothesized under Specific Aim 1.

C. Methods and Research Plan:

The proposed study will employ a within-subject repeated measures design to establish the functional relationship between amount of treatment and treatment response. We will also randomize half of the sample to a delayed treatment condition in order to control for the effects of social contact associated with PIRATE participation. Finally, we will supplement this design with between-subjects analyses designed to measure how different randomly-selected samples of participants respond to varying amounts of treatment. Study

participants will receive intensive semantically-oriented naming treatment in the context of the PIRATE program. Their response to treatment will be measured and compared across the 4-week duration of PIRATE using probes of treated and untreated stimuli as well as a standardized measure of naming performance. This comparison will address Specific Aim 2, to identify the dose-response relationship for aphasia therapy. Measures of naming gains in response to treatment will be compared across groups of participants defined in terms of their language-impairment profiles, to determine whether semantically-oriented naming treatment benefits all or only a subset of individuals with naming deficits. The study will also employ voxel-based lesion symptom mapping analyses to examine the relationship between lesion site and extent and treatment response, as well as regression analyses to examine the relationship between cognitive factors and treatment response. These analyses will together address Specific Aim 1, to identify psycholinguistic, cognitive, and neuroanatomical predictors of positive response to intensive semantically-based naming treatment.

This design avoids the confounds present in many previous studies of treatment response. First, it holds treatment intensity constant while systematically measuring the effects of varying amounts of treatment. Second, it uses both between and within-subject design components to control for a variety of potential confounding effects, including social contact, inter-subject variability, and the passage of time. Third, it examines the relationship between psycholinguistic, cognitive, and neuroanatomical factors and treatment response. Previous studies of treatment response have examined the effects of only one of these factors in accounting for differential response to treatment.

C.1 Study participants

Potential participants will be informed of the study on the PIRATE website as well as during PIRATE candidates' initial evaluation to determine eligibility for PIRATE participation. PIRATE candidates will be informed that they may be eligible to participate in the study if they meet additional study eligibility requirements. As described below, study participants will receive up to \$1500 to defray any travel expenses incurred as part of study participation, to be pro-rated if they fail to complete all study procedures (including follow-up visits).

All study participants will be community-dwelling veterans with a positive diagnosis of aphasia and the ability to live independently and carry out activities of daily living during their PIRATE program participation. Aphasia diagnosis will be provided by the referring provider, based on performance on standardized aphasia-assessment measures such as the Comprehensive Aphasia Test (CAT), the Western Aphasia Battery, Revised (WAB-R), or the Boston Diagnostic Aphasia Exam (BDAE). Aphasia diagnosis will be verified by PIRATE staff prior to enrollment based on clinical impression and performance on the CAT, with a language modality mean T-score below the 70 being required for aphasia diagnosis. Although this criterion is more liberal than the cut-off score of 62.8 described in the CAT manual, our experience suggests that patients may score above this cut-off but nevertheless present with clinically significant naming impairments. Additional exclusionary criteria for PIRATE participation are listed in Table 5 below.

Inability to carry out activities of daily living necessary for self-care
Lack of physical independence
Significant mood or behavioral disorders
Drug or alcohol dependence
Inability to tolerate intensive treatment
Medical conditions which would preclude independent living

Table 5. Exclusionary criteria for PIRATE participants.

Information regarding these exclusionary factors is obtained by PIRATE staff prior to program entry, based on review of veterans' medical charts, pre-entry applications, and interviews with family members and service providers. The absence of medical conditions which would preclude PIRATE participation is verified as part of a medical clearance exam provided by a VAPHS physician prior to program entry.

In addition to the criteria described above, participants in the proposed study must meet five additional criteria. First, only veterans with aphasia due to unilateral left-hemisphere stroke of greater than six months post-onset will be included in the study. Individuals with right-hemisphere involvement will be excluded because of the study's emphasis on testing hypotheses regarding potential non-dominant hemisphere recruitment in treatment response. Individuals less than 6 months post-onset will be excluded to ensure that participants' brain injuries are stable, and to reduce any potential effects of spontaneous recovery on treatment response. Second, only participants without progressive neurological disease or prior central nervous system injury or disorder will be enrolled in the study, since these comorbid factors could affect treatment response. Third, study participants must demonstrate less than or equal to 33% correct performance on at least 120 items contained in the 361-item baseline probe set (see section C.2.1. below), measured across three

administrations. This criterion is necessary to permit selection of the necessary treatment and generalization probe items as described below, and will minimize ceiling effects on the naming probes. Fourth, in order to avoid enrolling into the study participants who have profound naming impairments and are therefore unlikely to benefit from the study treatment (e.g., patients with global aphasia and/or verbal output limited to recurrent stereotypy), all participants must obtain a CAT Naming modality T-score greater than or equal to 40. Finally, we will exclude from participation potential participants with severe motor speech disorder (i.e., apraxia of speech and/or dysarthria). Motor speech diagnosis and severity judgments will be made by study investigators with experience in this area (Doyle, McNeil, Hula), based on the following samples of behavior: Connected speech samples (described below); three recorded attempts at laryngeal diadochokinesis at slow, medium, and fast rates; three trials of maximum sustained phonation; and recorded responses to the three subtests of the Apraxia Battery for Adults-2, including Speech Diadochokinesis, Repetition of Words of Increasing Length, and Word Repetition with Repeated Trials. Two judges will independently diagnose and rate the severity of motor speech disorder in each potential participant. A third judge will provide a diagnosis and rating to resolve any disagreements regarding presence or severity of motor-speech disorders. Chart review of PIRATE participants to date suggests that the above additional five criteria would have excluded approximately 20% of the unique veterans served by the program. Thus, out of 72 patients who will be seen in PIRATE over the four-year duration of the project, approximately 58 should be eligible for enrollment in the study. Of these 58 eligible veterans, it is reasonable to expect that at least 50 will enroll in the study and complete the treatment and all study visits (including follow-up visits). This sample size should be sufficient for detecting the relationships between treatment response, psycholinguistic profile, and cognitive performance relevant to Specific Aim 1, and to detect differences relevant to testing the dose-response hypotheses in Specific Aim 2.

Some participants may also have co-occurring conditions which would prevent them from participation in the imaging arm of the study (claustrophobia; presence of metal in the body). These participants will be enrolled in the study but will not complete imaging procedures. A chart review of PIRATE participants to date indicates that approximately 20% of veterans enrolled in PIRATE may have such exclusionary criteria: roughly 17% report pacemaker or other implanted metal objects or devices incompatible with MR scanning. Thus, the participant sample used to address the neuroanatomical hypotheses proposed under Specific Aim 1 may be smaller than the sample contributing data the other proposed analyses. Under the worst case scenario of no overlap between these imaging-related exclusionary criteria and those described in the preceding paragraph, we expect to obtain a sample size of 40 participants (50 minus 20%) to conduct the proposed imaging-related analyses. As described in section C.3.1.3 below, this sample size should be sufficient to detect the relationships relevant to the neuroanatomical hypotheses.

C.2 Procedures

The schedule of procedures for participants enrolling in the study is summarized in the flowchart in Figure 4. Participants enrolling in the research study will receive a single pre-treatment structural MRI and a battery of cognitive assessments at program entry (see section C.2.3 below). They will also have their performance on standardized measures of naming and other language functions assessed repeatedly throughout the protocol (see section C.2.2 below), and they will receive an intensive dose of a treatment (SFA) designed to target lexical-semantic processes (see section C.2.1 below). Participants' performance on treated and untreated lexical items will be probed three times weekly, to assess acquisition and generalization and to decide when participants will move to a novel set of treatment exemplars.

Half the participants will be randomly assigned to a delayed-treatment condition. The purpose of this condition is to provide an additional measure of experimental control, to ensure that any observed improvements are not due simply to the social contact afforded by the residential component of the treatment program. Previous work [7] has found that simple exposure or engagement in the absence of specific therapy

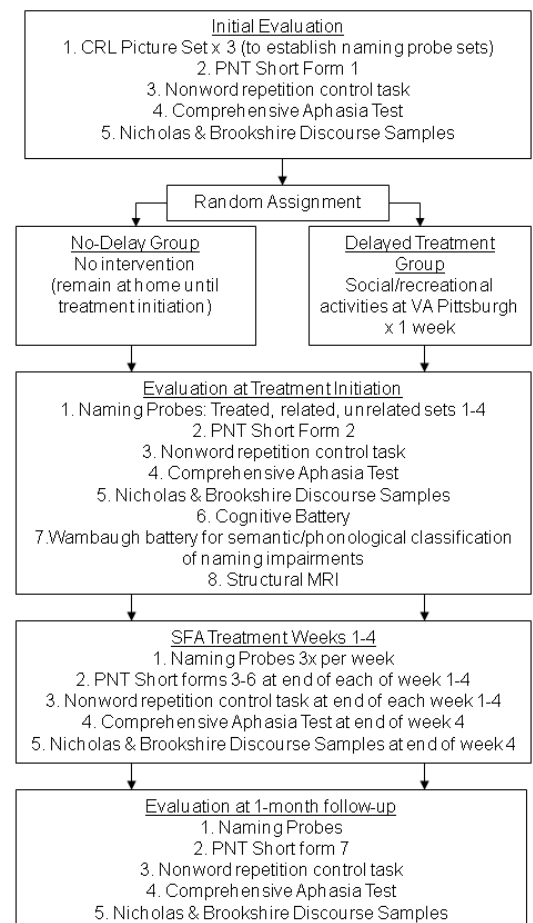


Figure 4: Testing and treatment procedures for study participants

activities for participants is insufficient to explain their gains in communicative function. However, inclusion of a delayed-treatment condition will help establish this more conclusively for the current study. The specifics of the delayed-treatment condition are given in section C.2.1.

PIRATE involves a single 2-day visit to VAPHS 4-6 weeks prior to entry for a medical clearance exam and an initial comprehensive speech-language evaluation and to initiate treatment planning. During this visit, the CAT and PALPA subtests (for confirmation of aphasia diagnosis and for participant classification and treatment planning) are administered. Accepted candidates return to VAPHS at program entry and reside in residential villas at the H.J. Heinz campus of VAPHS for 4 weeks (or 5 weeks in the delayed-treatment condition). At program entry, a second CAT is administered to participants, and PIRATE candidates enrolling in the research study will be given structural MRI scans and the battery of cognitive assessments described above. Therapy is provided in existing, dedicated PIRATE treatment rooms, also located on the Heinz campus. The treatment schedule is five days a week, approximately 5 hours a day, for a total of 25 hours of therapy each week. As described below, PIRATE participants enrolled in the proposed research will receive treatment five days a week, for 4 hours per day. All therapy will be provided by PIRATE-dedicated licensed clinical providers with multiple years of experience delivering aphasia therapy. Participants return to VAPHS for a single follow-up visit 4-6 weeks after program exit.

C.2.1 Treatment:

Over 85% of PIRATE participants to date have received treatment focused on spoken word production. This is also the unit of language function and the output modality targeted in the majority of treatment-neuroanatomy/neuroimaging studies reviewed in section A.2 above, as well as the majority of treatment-cognitive prediction studies reviewed in section A.2. SFA treatment will be implemented as described by Boyle & Coelho⁸ and Boyle.⁹ At the beginning of a treatment trial, the target picture will be placed in the center of a chart containing the following entries: Group, use, action, properties, location, and association. A treatment trial will proceed through the following steps:

1. The participant is asked to name the picture.
2. Regardless of accuracy in Step 1, the participant is asked to verbally produce at least six distinguishing semantic features of the target. For each feature, the clinician asks a question (e.g., What is it used for?). If the patient responds with an appropriate feature (or features), the clinician repeats the feature, provides optional verbal reinforcement, and writes the response(s) in the chart. The clinician may also prompt the participants to provide additional and/or novel responses before proceeding to the next feature. If the participant is unable to produce an appropriate response, the clinician provides a cloze sentence cue (e.g., This is a kind of _____). If the participant is still unable to provide an appropriate response, the clinician provides the feature verbally and in writing, and asks the participant to repeat the feature name.
3. After all six features are written on the chart, the clinician again asks the participant to name the target picture, regardless of whether the participant has named the picture successfully at any earlier point during the same trial.
4. If the participant is unable to name the picture, the clinician says the name aloud, requires repetition, and requires the participant to verbally produce the six features.

Treatment will be initiated with the first set of 10 treatment items for each patient. The pictures will be presented in a different pseudorandom order on each cycle through the set. If the patient names 90% of the treated items correctly on three of four consecutive probes, treatment on that set will be discontinued and a new set will be introduced. Since naming probes will be given 3 times per week, this criterion ensures that participants who show rapid acquisition will move on to a new treatment set as often as once per week. This in turn means that participants will receive treatment on up to 40 treatment exemplars during their time in the study. This number is greater than the number of treatment exemplars used in the majority of published naming treatment studies recently surveyed by Snell and colleagues^[64].

Treatment will be conducted five days a week (Monday through Friday) across two daily 120-minute intervals, one morning session and one afternoon session, for a total of 4 hours of treatment per day and 20 hours per week. A 20-minute break will be provided within each 120-minute treatment interval and a 120 minute break will be provided between each 120-minute treatment interval. Time in treatment will be recorded in daily logs, as will the number of exposures to each training exemplar, the number of treatment trials, and the number of features (both novel and recurring) generated for each item on a given treatment trial. In addition, participants will be engaged in planned social and recreational activities on the weekends, such as visits to

local sites and events (like a baseball game). These activities will not target or specifically stimulate language function or the skills trained in therapy.

As indicated above, half the participants will be enrolled in treatment immediately upon their arrival at VAPHS, and half will be entered into the delayed-treatment condition. Participants assigned to the delayed-treatment condition will arrive one week prior to treatment initiation. During this week, they will be in residence at the HJ Heinz campus and engage in non-language recreational activities in 2 two-hour sessions per day (the same schedule as will be followed during treatment). These activities will be led by a study staff member experienced in working with stroke survivors. The morning session will consist of recreational activities regularly conducted at the H.J. Heinz campus, listed in Table 6 below:

Billiards	Recreational Hall, H.J. Heinz campus
Physical exercise	Gym, H.J. Heinz campus
Ceramics, other visual arts	Art Studio, H.J. Heinz campus
Puzzles and board games	Recreational Hall, H.J. Heinz campus

Table 6. Non-language-focused recreational activities for participants enrolled in delayed-treatment condition

The afternoon session will consist of either complementary activities drawn from Table 6 or off-site visits to local destinations. While these sessions will afford multiple social contacts and interactions, typical of structured recreational programs, they will not target specific speech and language behaviors. Participants randomized to the delayed treatment group will receive probes (naming and PNT and other assessments [CAT, discourse measures]) on the same schedule as the no-delay group.

The following procedures will be implemented to monitor and ensure treatment fidelity. All treatment sessions will be video recorded. For each participant and each day of treatment, study staff not providing treatment will randomly select 15 minutes (~ 7% of total treatment time) of video and review it for adherence to the semantic feature analysis (SFA) protocol. All deviations will be discussed with the treating clinicians. Monitoring will occur and feedback will be provided to clinicians by the end of the business day following the monitored sessions. When monitoring of a given clinician-participant pair detects no more than one deviation on three consecutive days, the pair will be monitored weekly for the remainder of the 4-week session. If subsequent weekly monitoring reveals more than one deviation from the SFA protocol, daily monitoring will be re-instituted until the criterion for weekly monitoring is again reached.

Treatment and probe stimuli for study participants will consist of picturable nouns, consistent with methods used in existing SFA treatment studies. All pictures used as treatment and probe stimuli will be drawn from a subset of 361 of the 520 object pictures published by the UCSD Center for Research in Language International Picture Naming Project^[91]. These 361 pictures were selected by removing from the full set 159 picture stimuli that are also included in the Philadelphia Naming Test^[92], which will serve as a secondary outcome measure for Specific Aims 1 and 2. Participants will be asked to name this set of 361 pictures on three occasions across the two days of their initial evaluation, with all administrations separated by at least 5

hours. All treatment and probe pictures will be selected from the pictures named correctly on fewer than two of those occasions. Four sets of 10 items each will be selected as potential treatment items, along with four sets of 10 related generalization probe items and four sets of 10 unrelated generalization probe items. For each treatment item, a category coordinate will be selected as a generalization probe item. The unrelated generalization probes will be selected from remaining items belonging to categories not occurring within the treatment set. These items will also be selected to minimize associative and other semantic relationships with items in the treatment sets. Within the constraints of semantic relatedness and baseline performance level, we will balance the treatment and generalization sets for lexical frequency and length.

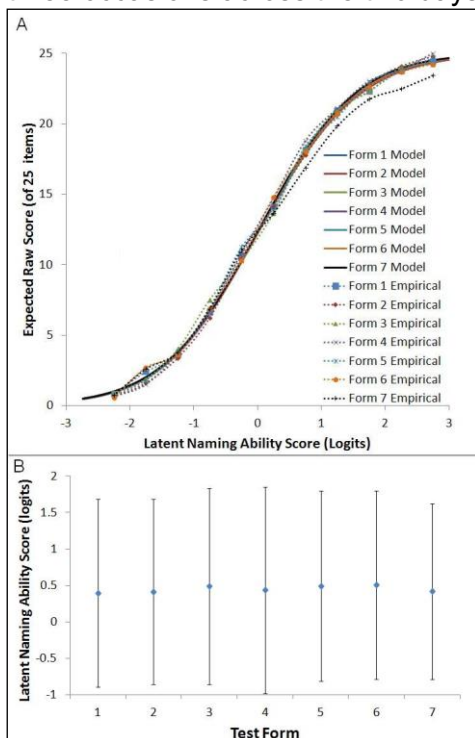


Figure 5. Panel A: Expected score for each of the seven 25-item PNT alternate forms, compared to empirical data of 254 aphasic patients. Panel B: Mean Rasch-scaled score for each of the seven alternate PNT forms, for a subsample of 57 aphasic patients.

C.2.2 Probe stimuli, procedures and schedule

Following baseline assessment of all 361 UCSD CRL items, the 40 items selected for training, 40 related items, and 40 unrelated items will be probed at treatment initiation, three times per week during

treatment, at program exit, and at one-month follow-up. As an additional measure of within-subject control, we will also obtain baseline measures and weekly probes on the non-word repetition subtest of the PALPA. Generalization of SFA treatment to improved performance on nonword repetition is theoretically possible, but unexpected. The Philadelphia Naming Test (PNT)^[92], a standardized measure of naming performance, will be administered weekly during baseline and treatment phases of the study. In order to eliminate the possibility that observed performance gains on the PNT could be attributed to repeated exposure to the same picture stimuli, we used methods based in item response theory^[93] (IRT) to organize the 175 PNT items into seven equivalent 25-item alternate forms. We accomplished this in the following steps, using data taken from the Moss Aphasia Psycholinguistics Project Database (MAPPD)^[94], which contains item-level PNT response data collected from 254 persons with aphasia. First, we confirmed that a single underlying factor can satisfactorily account for aphasic performance on the PNT. We fit a unidimensional model to the data using NOHARM^[95], and found that fit was excellent (root mean square of the model residuals = 0.010 vs. criterion of < 0.25; Tanaka Goodness of Fit Index = 0.985 vs. criterion of >0.95 for good fit). Next, we fit the data to a 1-parameter IRT model (Rasch) for dichotomous items using WINSTEPS^[96, 97]. We found that all items obtained information-weighted mean-square statistics <1.4, suggesting adequate fit to the model^[98]. We then used the model-estimated item difficulty values to construct the seven alternate forms such that they were of equivalent overall difficulty and gave equivalent expected proportion correct scores as a function of latent naming ability. Panel A of Figure 5 shows the model predicted and empirical raw score for each alternate form as a function of Rasch-scaled latent naming ability, estimated based on the responses to the full PNT. The model curves for the seven forms are essentially identical and well-approximated by the empirical data. We then cross-validated these forms using new data from a subsample of 57 patients who were administered the PNT a second time. For these 57 patients, we computed the latent naming ability score for each alternate form and found that they adequately fit a strictly parallel test model (i.e., had equivalent means, true score variances, and error variances), $\chi^2 = 37.757$ with 32 df, $p = 0.223$. The average correlation between forms was 0.87 (min-max = 0.82-0.93). The means and standard deviations for each of the seven forms are displayed in Panel B of Figure 5.

The naming probes and the PNT will be presented via computer by study staff not involved in treatment. The PNT alternate forms will be administered according to the procedures described by Roach et al. The naming probes will be administered in a different random order at each measurement point according to the same procedures as the PNT, with one modification: In order to make initial testing with a large number of potential treatment items feasible, participants will be given 15 rather than 30 seconds to respond to each item.

Responses to the naming probes and the PNT will be audio recorded for subsequent transcription and scoring according to the conventional scoring procedures described by Mirman et al.^[94] and available on the MAPPD website (www.mappd.org). Specifically, dichotomous correct/incorrect (1/0) scoring will be applied to the first complete attempt at each item. Minor sound distortions, additions, and omissions attributable to will be ignored for the purposes of dichotomous scoring. For the naming probes, the primary dependent variable will be the sum of these dichotomous item scores. For the PNT alternate forms, the primary dependent variable will be the logit-scaled latent ability score derived from the sum of the dichotomous item scores. The lexical and phonological error codes contained in the conventional and model coding system for the PNT responses will also be collected for all naming probe and PNT responses for use in secondary analyses. The nonword repetition control probe stimuli will be audio recorded by study staff and administered in pseudo-random order at each measurement point. Responses will be audio recorded for subsequent transcription and scoring. Scoring will be dichotomous correct/incorrect.

For all naming probes, nonword repetition control probes, and the PNT, initial dichotomous scoring will be completed online by the study staff administering the stimuli. This online scoring will be used for the purposes of establishing inter-rater reliability. Primary scoring of each of these variables will be completed by a different member of the study staff from the audio-recorded responses. Prior to primary scoring, the filenames of audio recordings will be scrambled with respect to administration date so that the scorer will be blind to the measurement point. Inter-rater reliability between primary scoring and online scoring will be assessed for each participant by percentage point-to-point agreement and across participants by intraclass correlation coefficient.

In addition to the probes described above, we will also collect data on two standardized assessments of language performance to examine the generalized effects of intensive lexical-semantic naming treatment. These include the Comprehensive Aphasia Test^[99] and the connected speech task described by Nicholas and Brookshire^[100]. The language battery of the CAT is comprised of several subtests designed to assess both word and sentence level performance in the areas of auditory and reading comprehension, repetition, naming, spoken and written picture description, oral reading and writing. T-scores based upon a well-described

normative sample of adults with aphasia (n = 146) may be derived for each of the eight modalities to provide a profile of the measured impairment, and a mean modality T-score is computed to provide an overall estimate of the severity of the impairment. We will use modality T-scores to examine changes in reading, writing and comprehension performance in response to treatment.

The Nicholas and Brookshire task employs 10 stimuli: 4 single pictures, 2 picture sequences, 2 requests for personal information and 2 requests for procedural information. The stimuli are presented in random order to elicit connected speech samples from which the following measures may be derived; number of words (#words) in the sample, number of words that are correct information units (#CIUs), CIUs per minute (CIUs/min), and percentage of words that are CIU (%CIUs). We will employ the measures of CIUs/min and %CIUs to assess the effects of intensive lexical-semantic naming treatment on the informativeness and efficiency of study participants' connected speech. These measures may be calculated with high inter- (words = 98%; CIUs = 90%) and intra-rater (words = 99%; CIUs = 95%) reliability, and have been reported to be stable across repeated assessments separated by 10 minutes, and by 7-10 days (mean absolute difference words = 6.82; mean absolute difference %CIUs = 3.50) in speech samples elicited by 5 stimuli, presented in random order, to adults with aphasia^[101].

Both the CAT and the connected speech task will be administered on 4 occasions: at baseline, study entry, study exit and follow-up. The connected speech task will employ a different, randomly selected set of 5 stimuli across repeated assessments, be audio recorded, orthographically transcribed and coded according to the conventions described by Nicholas and Brookshire^[100]. Both of these tasks will be administered by study staff in line with the procedures used for probe and PNT measures above (i.e., samples will be blinded to time point and coded and scored by study staff rather than PIRATE clinicians).

C.2.3 Participant classification and predictor measures: Psycholinguistic profile, cognitive assessment, and neuroimaging methods

Participants' naming-impairment profiles will be classified based on their performance on tests of lexical-semantic and phonological processing. These classification criteria will be the same as those used in Wambaugh and colleagues' previous study examining the effects of language-impairment profile on naming treatment response^[12]. Wambaugh and colleagues' criteria are summarized in Figure 6.

	Primary Semantic Deficit	Primary Phonological Deficit
Picture and word semantics		
Categoryization	= 90% accuracy	> 90% accuracy
Auditory word-to-picture matching	= 90% accuracy	> 90% accuracy
Written word-to-picture matching	= 90% accuracy	> 90% accuracy
Rhyming		
Pictorial input- word pairs	No cut-off	= 70% accuracy
Pictorial and verbal input	20% > pic. input-word pairs	= 70% accuracy
Pictorial input, field of 4 items	No cut-off; semantic foil errors predominate	= 70% accuracy
Picture naming		
Oral naming	= 60% of errors are semantic paraphasias, phonological accurate unrelated words, or no response; =40% are phonemic paraphasias	= 60% of errors are phonemic paraphasias, circumlocutions, or neologisms; = 40% are semantic paraphasias
Written naming	Same as above	Same as above
Phonemic cueing responsiveness		
First-sound cueing	No cut-off	Facilitative for = 15% of inaccurately named words Not responsive to miscues
Miscuing	May be responsive to miscues	
Oral reading and repetition		
Oral reading	= oral naming accuracy	No cut-off
Repetition	= 90% accuracy	< 90% accurate

Figure 6. Criteria for classification of patients' naming deficits, taken from Wambaugh et al.^[12]

The semantic categorization task will involve sorting of 70 pictures into seven different superordinate categories (e.g., transportation, animals, tools, etc.). Rhyming, repetition, word-picture matching, oral reading, and repetition will be assessed using appropriate PALPA subtests. Picture naming error type data will be collected from the initial presentation of the 361-picture set described below, and from the subset of 50 Philadelphia Naming Test items that each participant will receive by program entry, also as described below. Phonemic cueing responsiveness will be measured by presenting the pictures of the Boston Naming Test along with a single-segment phonemic cue. Those cues will either overlap with the first segment of the target (first sound cueing:

e.g., [b] for target 'broom') or will overlap with the first segment of a semantic competitor (miscueing: e.g., [m] for target 'broom,' overlapping with semantic coordinate 'mop')^[102].

All participants will also be administered a set of cognitive assessments immediately prior to program entry. These cognitive tasks will include measures of executive function abilities, visuospatial skills, recognition memory, and tasks of sustained and divided attention. The WCST and the picture and word versions of the Pyramids and Palm Trees Test (PPT) will be administered to assess executive function skills. The PPT was found to correlate with the cognitive factors identified in Lambon Ralph et al.'s study and performance has been shown to depend partially on reasoning and problem-solving abilities^[103]. Additionally, two of the studies found significant correlations between PPT performance and treatment response following naming therapy^[10].

^{66]}. The immediate and delayed recall of the Complex Rey Figure will be administered to evaluate visuospatial abilities and memory. This test has been found to be significantly correlated with naming treatment outcomes^[10]. Two of the short recognition subtests of the Camden Memory Test, the topographical and word subtests, will be administered to assess recognition memory^[66]. Additionally, study participants will be given the sustained and divided attention subtests of the TEA. As noted above, Lambon Ralph and colleagues^[10] found that the elevator counting with distraction subtest on the TEA was one of the significant cognitive factors that predicted naming treatment outcome. The TEA provides an aphasia-friendly, ecologically valid assessment of multiple forms of attention that has been standardized on a group of stroke patients^[104].

Finally, participants will receive structural magnetic resonance imaging scans immediately prior to program entry. Not all participants will be eligible to take part in imaging procedures (due to presence of exclusionary conditions such as claustrophobia or metal in the body). As noted above, current estimates are that approximately 20% of PIRATE participants to date have such exclusionary conditions.

At program entry, all participants eligible for imaging will have a single structural magnetic-resonance imaging scan. Scans will be collected in the scanning facility located at the University Drive campus of VAPHS. The facility includes a new 3T Siemens Magnetom Verio Scanner (70 cm open bore) with a 12 channel MATRIX head-coil. The gradient coil set achieves 45mT/m maximum amplitude and a slew rate of 200mT/m/s with up to FoV 50 cm. The scanner is operated by registered MR technologists and is equipped with MRI-compatible cardiac ECG equipment and blood oxygen monitors. All conventional and echo planar MR imaging functions are supported. The system is interfaced to a high-speed local area network for data transfer to workstations for analysis. All images will be collected under supervision of Dr. Forman.

For each patient, we will obtain whole brain, high resolution T1 and T2-weighted structural MR images on the VAPHS 3T scanner., using an MPRAGE sequence with an axial slice orientation [Repetition Time (TR)= 9.90ms, Echo Time (TE)=4.60ms, Voxel Size =.94mm x.94mm x 1.0 mm, Matrix Size=256mm x 256 mm, Flip Angle=8°]. Images will be reviewed by Co-Investigator Steven Forman, MD, Ph.D. and a staff neuro-radiologist. Any unanticipated findings from the MR scan will be shared with VAPHS medical staff and may be referred to participants' medical caregivers, as appropriate. These images will be housed on the VA network and will be viewed using ViSTA Image. We will manually trace all lesions in native space using the ITK-Snap image processing program^[105]. We will then normalize each scan and its associated lesion mask to a standard neurological template (i.e., Montreal Neurological Institute) using SPM-08's spatial normalization tool. Once all of the lesions have been oriented to a common template, we will conduct the VLSM analysis using the MRICron nonparametric analysis toolbox^[106].

C.3 Data analysis

Three separate sets of analyses will be carried out on study measures. The first set of analyses will examine differences between participants randomized to the delayed-treatment and immediate-treatment conditions. The second set of analyses will test hypotheses relevant to Specific Aim 1, identifying correlates of response to intensive semantically-oriented naming treatment. The third set will test the dose-response hypotheses relevant to Specific Aim 2.

The first analysis will evaluate whether the social contact associated with PIRATE participation results in improved performance on study measures, independent of any actual aphasia treatment. We will use a multivariate mixed model to test whether there are any differences between the immediate-treatment and delayed-treatment groups at the initiation of treatment, and whether there is any group-by-time interaction suggesting a positive effect of the delayed treatment week. The dependent variables will include performance on treated and untreated naming probes and PNT short forms.

If this analysis finds no significant differences between the groups, it will suggest that any observed treatment gains can be attributed to the SFA treatment rather than general effects of the social contact associated with PIRATE. In this case, the two groups will be collapsed for further analysis. In the unexpected event that this analysis finds a significant difference favoring the delayed treatment condition, it will suggest that observed treatment gains should be attributed to the combination of SFA and the socialization inherent in a residential treatment program, rather than to SFA alone. In this case, we will use the data to estimate the size of the socialization effect and statistically control for it in subsequent analyses, where possible.

C.3.1 Psycholinguistic, cognitive, and neuroanatomical predictors of treatment response

The primary dependent variable for analyses relevant to Specific Aim 1 will be the proportion of correctly-named treated stimuli at study exit (post-treatment), minus the proportion of correctly-named treated stimuli at study entry (pre-treatment). This difference score provides the simplest measure of naming treatment response. Supplementary analyses may be carried out using a related dependent variable, the percentage of possible naming improvement^[10]. This is a derived measure which divides the simple pre-post difference score

by the difference between pre-treatment naming performance and ceiling performance (100% accuracy minus observed pre-treatment accuracy). This measure helps correct for observed differences in baseline accuracy.

Secondary dependent variables for these analyses will include accuracy on untreated related stimuli (accuracy at study exit minus accuracy at study entry) and Rasch-scaled scores for PNT stimuli (scores at exit minus scores at entry). Previous studies have not found any predictive relationships between neuroanatomical or other variables and generalization to untrained stimuli^[9, 20]. However, those studies did not administer treatment in the amounts or intensities provided in PIRATE or in the proposed research. It is possible that more intensive and/or larger doses of aphasia treatment may uncover stronger generalization effects (as well as stronger predictive relationships) than have previously been reported.

C.3.1.1 Group effects on treatment response

Participants will be classified as having primarily phonological, primarily lexical-semantic, or mixed naming impairments, using the methods described in section C.2.3 above. A between-participants analysis of variance will be carried out on the primary and secondary dependent measures of treatment response, with language-impairment group as an independent variable. As predicted by Hypothesis 1a and consistent with previous findings in the naming-treatment literature^[12, 44], we anticipate that there will not be a main effect of language-impairment group on treatment response. The different groups of naming-impaired individuals will show comparable benefit from intensive, semantically-oriented naming treatment. If group differences are found in the ANOVA, we will explore their source using Tukey's multiple comparisons to control type-I error.

To determine whether the proposed sample size will be sufficient to detect these between-group differences, a power analysis was carried out based on data from Wambaugh et al.^[12] and Abel et al.^[44]. Together, these two studies provide pre-treatment and post-treatment naming probe data for four aphasic participants classified as having primarily semantic impairment, four participants classified as having primarily phonological impairment, and one patient with mixed impairments, all of whom received semantically-oriented naming treatment. As noted above in section A.1, the effects of psycholinguistic profile on response to semantically oriented treatment were potentially counter-intuitive: there was better response in phonologically impaired vs. semantically impaired participants. Nevertheless, we used the pre-treatment/post-treatment probe data from these 9 patients to estimate the variance in change scores. We assume that a 20 percentage-point difference in treatment effect (i.e., an average of 2 items, given the 10-item probe sets) is needed for the effect to be considered clinically meaningful. Assuming that 12 of the anticipated 50 participants will have primarily semantic impairments and 12 will have primarily phonological impairments, the proposed between-subjects ANOVA will have power of 0.82 to detect a 20-percentage point difference between these two groups.

If no group differences are found in treatment response, as predicted under Hypothesis 1a, the treatment-response data will be collapsed across language-impairment groups for subsequent analysis. If group differences in the magnitude of treatment response are found, separate analyses for the relevant groups will be carried out for the cognitive and neuroanatomical hypotheses, in the analyses described below.

C.3.1.2 Prediction of treatment effects by cognitive factors

Participants' performance on the set of cognitive measures identified above (section C.2.3) will be transformed into z-scores, to ensure that the scores may be compared across tasks. These transformed scores will be submitted to principal component analysis in order to identify one or more factors which account for performance across the measures. Based on previous findings^[10, 69], we anticipate that measures of executive function (WCST) and performance on the PPT will load on one factor, while measures of visuo-spatial memory (Rey Complex Figure) and other memory function (Camden Memory Test, topographical and word subtests) will load on a separate factor. Performance on tests of focused and divided attention (TEA) may also load on a separate factor.

These factors will be entered as predictors into separate linear regressions targeting primary and secondary dependent measures, to determine which cognitive factors predict these measures of treatment response. Based on previous results, we anticipate that the executive-function factor^[10, 71], the memory factor^[10, 66], and the attention measures^[10] will all be significant predictors of naming treatment response. If no reliable predictive relationships are found between these cognitive factors and treatment response, we will perform additional correlation analyses to examine the relationship between performance on individual cognitive tasks and treatment response. We may also subtract variance associated with aphasia severity (as assessed by the CAT mean language modality T-score) using stepwise linear regression, since different aphasia severities may be associated with different responses to treatment^[1]. Given that the studies referenced above have demonstrated the hypothesized relationships with sample sizes ranging from 7 to 33 participants, we expect that the present analyses will be sufficiently powered, given the proposed sample size of 50.

C.3.1.3 Voxel-based lesion symptom mapping (VLSM)

For each participant, lesion size and extent will be quantified and related to treatment response using voxel-based lesion symptom mapping (VLSM), to test the neuroanatomical hypotheses relevant to Specific Aim 1. These analyses will compare the structural MRI data collected for study participants with change scores based on their performance on outcome measures.

VLSM analysis will proceed in several steps. First, we will employ a multi-step lesion tracing procedure, first tracing lesions volume-by-volume in a descending axial plane while simultaneously referencing the coronal and sagittal views in the ITK-Snap image processing program^[105]. These initial tracings will be conducted via consensus among study staff. We will then subject these initial tracings to a second round of consensus with different raters and conduct minor adjustments as needed. Dr. Forman will have final say in any conflicts regarding lesion tracing. He will also be blind to each participant's corresponding behavioral data.

Upon reaching consensus on the lesion distributions, we will warp the original brains and their respective lesion masks into a standard neurological space (i.e., Montreal Neurological Institute template). We will first re-align each brain along its AC-PC axis and then segment it into distinct tissue types (i.e., gray matter, white matter, CSF, skull). After these initial pre-processing steps, we will normalize each brain using SPM-08. During normalization, we will apply the lesion tracings as masks with the goal of preventing tissue distortion in and around the lesion. We will then apply the original normalization parameters to the masks themselves in order to derive a set of normalized lesion patterns that may then be overlaid on a canonical brain and contrasted across subjects.

We will use the MRICron program's nonparametric lesion mapping (NPM) feature to complete the lesion symptom correlations^[106]. These data will be structured such that there is a binary classification at the voxel level (i.e., ±lesion) and continuous behavioral data (change score associated with the different outcome measures: PNT scores and scores on acquisition/generalization probes). Since these data are likely not to satisfy parametric statistical assumptions, we will use the NPM program's Brunner-Munzel nonparametric test to relate behavioral performance on each of the treatment-outcome predictors (change scores) to lesion distributions^[107-109]. We will apply a cluster threshold of 10 contiguous voxels and threshold for a false discovery rate of .05 in order to control for type I error while maintaining sufficient power to detect effects. We will also include voxels that have at least a 25% lesion overlap across patients. VLSM analyses will focus on four principal regions of interest (ROIs): left anterior cortex, left posterior cortex, left basal ganglia, and left hippocampus. Previous work (reviewed in section B.2 above) has shown that positive response to aphasia treatment is associated with intact left hippocampus and left basal ganglia. In addition, some previous work suggests that negative treatment response is associated with more intact left anterior cortex^[21], but other work suggests that greater recruitment of left anterior cortex is associated with positive treatment response^[23]. Additional work has also found that more intact left posterior cortex is associated with more positive treatment response^[23].

Change scores from program entry to program exit and/or follow-up will be computed for primary and secondary study measures: weekly PNT scores and acquisition/generalization probes. These scores will be correlated with lesion extent using VLSM, with analysis focusing on the four ROIs identified above. These correlations will serve to test the hypotheses relevant to Specific Aim 1 above. As predicted by Hypothesis 1c and consistent with the preliminary findings from the PIRATE participants (see B.2.1 above), we anticipate that there will be a significant negative correlation between change scores and lesion extent in left basal ganglia and hippocampus. Greater damage to structures required for learning and memory function (hippocampus) and to structures which can control maladaptive activation of damaged left-hemisphere tissue (basal ganglia) will result in poorer treatment outcomes. Also as predicted by Hypothesis 1c and consistent with preliminary data by Parkinson and colleagues (see B.2.1), we anticipate that extent of left-anterior hemisphere lesion will be positively correlated with change scores. Greater damage to left anterior cortex will prompt greater recruitment of right hemisphere cortex or other non-lesioned left hemisphere cortex, resulting in more positive treatment outcomes^[20, 25]. Furthermore, we anticipate that extent of left posterior cortex will be associated with poorer treatment outcomes, given Fridriksson's findings^[23].

Additional exploratory analyses will be carried out to examine how these neuroanatomical factors predict performance for the different outcome measures. We anticipate that brain regions which are positively correlated with change scores for acquisition probes will not be positively correlated with change scores for generalization probes, given the behavioral^[9] and neuroimaging evidence^[20] that treatment response for treated and untreated items/behaviors leans on different cognitive and neural mechanisms. We do not have any a priori predictions regarding what brain regions may be correlated with change scores for generalization probes.

To determine whether the proposed sample size will be sufficient to detect the lesion-behavioral correlations described above, a power analysis was carried out on the correlational data from Parkinson and colleagues^[21]. The correlation coefficients reported by Parkinson range from 0.65 to 0.85. Power analysis indicates that a sample size of 40 will provide power >0.99 for detecting effects of this magnitude and power of 0.80 for detecting correlations with r values of 0.38 or greater. Given the likely greater variability among the patients who participate in PIRATE, both in terms of their demographic characteristics and their language impairment profiles, we anticipate that any correlations between neuroanatomical markers and language performance and recovery may be somewhat weaker than those observed in the Parkinson and colleagues study^[21], which had a much more uniform sample. However, the projected sample size of 40 should be sufficient to detect weaker correlations than those observed by Parkinson and colleagues^[21].

If no significant correlations are found between the above ROIs and change scores for the outcome measures, we will perform additional VLSM analyses to examine the relationship between lesion site and extent and treatment response in subgroups of study participants. Since participants in this study will likely be more heterogeneous than those in previous studies, the data from the current study may be noisier due to greater inter-subject variability. To explore the effects of this variability on neural correlates of treatment response, participants will be divided by aphasia severity, their CAT T-scores, since different aphasia severities may be associated with different responses to treatment^[1]. Additional analyses may explore the lesion-treatment relationships found in different subgroups of veterans.

C.3.2 Dose-response analyses

Study measures will be subjected to two sets of analyses which test the dose-response hypotheses relevant to Specific Aim 2. These analyses will compare the PNT scores and acquisition and generalization probe scores across the 4 weeks of PIRATE treatment, to determine the dose-response relationship and to identify the points of largest change in response to treatment. The last of the three weekly naming probes will serve as a measure of weekly acquisition and generalization performance. This will ensure that performance on PNT and naming-probe measures may be directly compared, as they will be on the same schedule.

First, the data for the PNT scores and the acquisition/generalization probes will be analyzed with a multivariate mixed model with polynomial trend contrasts. To simplify interpretation, only time points 2 through 6 (program entry through program exit) will be included in the trend analyses. These trend analyses will first test whether there is a linear relationship between time in treatment and treatment response for study measures, as predicted by Hypothesis 2a. Second, they will test whether acquisition and generalization probes show additional curvilinear relationships to time in treatment, as predicted by Hypothesis 2b. We anticipate that there will be a positive linear trend between program entry and program exit for PNT scores and for acquisition and generalization probes. This finding would indicate that the four weeks of treatment results in significant score gains. We further predict that the trend for acquisition probes will also have a significant curvilinear (quadratic or higher-order) component, consistent with decelerating improvement across the four weeks of treatment. This finding would indicate that the largest gains for specifically treated skills/items occur early in treatment, with declining gains in later weeks. Finally, we predict that the generalization probes will also demonstrate a curvilinear relationship to time, due to accelerating improvement across the four week program. This finding would indicate that the largest gains for related but untreated items occur late in treatment, emerging later than the gains for specifically treated forms.

Second, six pairwise planned comparisons of scores across all seven consecutive measurement intervals will be carried out, using Bonferroni-corrected t-tests to maintain the familywise type-I error rate at 0.1. As predicted by Hypothesis 2a and consistent with the preliminary outcome data from PIRATE (see B.1.3 above), we anticipate that there will be a significant increase in both PNT score and acquisition/generalization probe performance for each additional one-week increment of treatment. We also predict non-significant changes between the initial evaluation and program entry and between program exit and follow-up.

In case a linear relationship is not found between time in treatment and outcome measures, additional analyses will be carried out examining the performance of different subsets of veterans. For example, veterans will be divided into groups based on their aphasia severity, since previous treatment findings have suggested that milder aphasia severity may be associated with less change on language-performance measures^[1]. Separate trend analyses will be carried out for the different severity groups, to examine whether different dose-response relationships may be observed for different aphasia groups.

In addition to these within-participant analyses examining dose-response relationships, a complementary set of between-participants analyses will be carried out, examining the effects of two versus four weeks of treatment for different randomly-selected groups of participants. It is standard practice in the dose-response literature to compare the response of groups exposed to different amounts of a therapeutic

agent, to see the incidence of positive or negative responses to the agent in the groups (e.g., frequency of remission or toxicity^[110]). These complementary between-participant analyses will simulate this comparison for the current study, providing an additional randomized-control measure of the effect of high dosage SFA on individuals with naming impairments. For these analyses, participants will be randomized to one of two groups. For half the participants, outcome measures (naming and PNT scores) will be sampled after two weeks' exposure, while for the other half, outcome measures will be sampled after the full four weeks of treatment. We anticipate that changes in study measures will be larger for the four-week group than for the two-week group.

C.3.3 Additional analyses of general language performance and connected speech

In order to evaluate whether the study treatment is associated with broad gains in communicative function beyond confrontation naming, we will analyze data obtained from the Comprehensive Aphasia Test (CAT) and the connected speech samples described above in section C.2.2. Specifically, we will conduct a multivariate linear mixed model analysis with measurement point (initial evaluation, program entry, program exit, follow-up) as a fixed factor, subjects as a random factor, and the following dependent variables: CAT reading, writing, auditory comprehension and mean modality T-scores, and, from the connected speech samples, the percentage of correct information units (%CIUs) and the number of correct information units per minute (CIUs/min). We anticipate that there will be significant increases in the CAT auditory comprehension and mean modality T-scores and in the two connected speech measures (%CIUs, CIUs/min) from program entry to program exit, showing generalized improvement in communicative function. We have no strong predictions regarding the generalized effects of SFA on reading and writing abilities as measured by the CAT.

C.4 Timetable

This proposal is for a four-year research project. Based on an October 1, 2012 start date, recruitment of participants should begin with the November 2012 PIRATE session of 2012. Enrollment will take place continuously through September 2016, with dissemination of results continuing through December 2016.

C.5 Investigators

Michael Walsh Dickey, Ph.D. and Patrick J. Doyle, Ph.D., CCC-SLP will serve as principal investigators. Dr. Dickey has expertise in the study and application of sentence processing treatments for aphasia. Dr. Doyle has expertise in aphasia assessment, treatment and clinical outcomes evaluation. The following individuals will serve as co-investigators: Steven D. Forman, M.D., Ph.D. has expertise in collection and analysis of brain imaging data and will oversee collection of structural imaging data and VLSM analyses; Bruce Crosson, Ph.D., has expertise in the application of brain imaging to aphasia treatment research and will advise regarding analysis and interpretation of imaging data and correlation analyses; Malcolm R. McNeil, Ph.D., CCC-SLP, has expertise in aphasia diagnosis, assessment, and treatment; William D. Hula, Ph.D., CCC-SLP, has expertise in analysis methods and measurement in aphasia; and Jamie J. Reilly, Ph.D., CCC-SLP, has expertise in language rehabilitation for aphasia and dementia, and in structural and functional neuroimaging and will advise regarding analysis and interpretation of imaging data.

C.6 Project resources

The VA Pittsburgh Healthcare System possesses the physical facilities and other resources necessary for the successful and timely completion of the proposed research project. The VAPHS Geriatric Research, Education and Clinical Center (GRECC) is a VHA supported Center of Excellence focused on stroke prevention, treatment, rehabilitation, and outcomes. Centralized GRECC funding supports a core staff of 12 FTEE including Drs. Dickey (2/8ths) and Doyle (8/8ths). Support for 2 FTEE PIRATE SLPs is provided by VISN 4 and GRECC Clinical Innovation Award Funds. The physical infrastructure of the GRECC includes office space and research labs equipped with VHA-networked secure computers and dedicated research servers. In addition, PIRATE has a dedicated residential villa located on the adjacent VAPHS Heinz campus. Daily PIRATE treatment sessions are also conducted on the VAPHS H.J. Heinz Campus in dedicated clinic space containing 3 individual treatment rooms, VHA-networked computers, and 6 dedicated non-VA-networked laptops used to design and manage treatment stimuli and deliver customized treatment programs. Dedicated PIRATE staff handle patient referrals, inquiries, scheduling, and logistical arrangements.

The VAPHS scanning facility includes a 3T Siemens Magnetom Verio Scanner (70 cm open bore) with a 12 channel MATRIX head-coil. The scanner is operated by registered MR technologists and is equipped with MRI-compatible cardiac ECG equipment and blood oxygen monitors. All conventional and echo planar MR imaging functions needed for the research are supported. The system is interfaced to the secure VHA network.

PIRATE receives enrollment applications from VA Medical Centers and DOD Hospitals nationwide and from key local referral sources. During its first 3 years of operation PIRATE received 136 applications for enrollment, with 19 veterans on the waiting list and several evaluations pending. This application rate is well in excess of the rates required to achieve study enrollment targets within the 4 years of the proposed project.

REFERENCES

1. Robey, R.R., *A meta-analysis of clinical outcomes in the treatment of aphasia*. Journal of Speech, Language, and Hearing Research, 1998. **41**(1): p. 172-187.
2. Holland, A.L., et al., *Treatment efficacy: aphasia*. Journal of Speech and Hearing Research, 1996. **39**(5): p. S27-S36.
3. Wertz, R.T., et al., *Comparison of clinic, home, and deferred language treatment for aphasia. A Veterans Administration Cooperative Study*. Archives of Neurology, 1986. **43**(7): p. 653-658.
4. Poeck, K., W. Huber, and K. Willmes, *Outcome of intensive language treatment in aphasia*. Journal of Speech and Hearing Disorders, 1989. **54**(3): p. 471-479.
5. Cherney, L.R., et al., *Evidence-based systematic review: effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia*. J Speech Lang Hear. Res., 2008. **51**(5): p. 1282-1299.
6. Bhogal, S.K., R. Teasell, and M. Speechley, *Intensity of aphasia therapy, impact on recovery*. Stroke, 2003. **34**(4): p. 987-993.
7. Elman, R.J. and E. Bernstein-Ellis, *The efficacy of group communication treatment in adults with chronic aphasia*. Journal of Speech, Language, and Hearing Research, 1999. **42** p. 411-419.
8. de Jong-Hagelstein, M., et al., *Efficacy of early cognitive-linguistic treatment and communicative treatment in aphasia after stroke: a randomised controlled trial (RATS-2)*. Journal of Neurology, Neurosurgery, and Psychiatry, 2010.
9. Dickey, M.W., Yoo, H., *Predicting outcomes for linguistically-specific sentence treatment protocols*. Aphasiology, 2010. **24**: p. 787-801.
10. Lambon Ralph, M.A., et al., *Predicting the outcome of anomia therapy for people with aphasia post CVA: Both language and cognitive status are key predictors*. Neuropsychological Rehabilitation, 2010. **20**: p. 289-305.
11. Parkinson, B.R., et al., *Lesion characteristics related to treatment improvement in object and action naming for patients with chronic aphasia*. Brain and Language, 2009. **110**(2): p. 61-70.
12. Wambaugh, J.L., et al., *Effects of two cueing treatments of lexical retrieval in aphasic speakers with different levels of deficit*. Aphasiology, 2001. **15**(10/11): p. 933-950.
13. Wertz, R.T., *Language treatment for aphasia is efficacious, but for whom?* Topics in Language Disorders, 1987. **8**(1): p. 1-10.
14. Boyle, M. and C.A. Coelho, *Application of semantic feature analysis as a treatment for aphasic dysnomia*. American Journal of Speech-Language Pathology, 1995. **4**(4): p. 94-98.
15. Coelho, C.A., R.E. McHugh, and M. Boyle, *Semantic feature analysis as a treatment for aphasic dysnomia: A replication*. Aphasiology, 2000. **14**(2): p. 133-142.
16. Boyle, M., *Semantic feature analysis treatment for anomia in two fluent aphasia syndromes*. American Journal of Speech-Language Pathology, 2004. **13**(3): p. 236-249.
17. Boyle, M. and C.A. Coelho, *Application of semantic feature analysis as a treatment for aphasic dysnomia*. American Journal of Speech-Language Pathology, 2008. **4**: p. 94-98.
18. Nickels, L., *Therapy for naming disorders: revisiting, revising, and reviewing*. Aphasiology, 2002. **16**: 935-979.
19. Wambaugh, J.L., *A comparison of the relative effects of phonologic and semantic cueing treatments*. Aphasiology, 2003. **17**(5): p. 433-441.
20. Meinzer, M., et al., *Integrity of the hippocampus and surrounding white matter is correlated with language training success in aphasia*. NeuroImage, 2010. **53**: p. 283-290.
21. Bruce Parkinson, R., et al., *Lesion characteristics related to treatment improvement in object and action naming for patients with chronic aphasia*. Brain and Language, 2009. **110**(2): p. 61-70.
22. Crosson, B., et al., *Role of the right and left hemispheres in recovery of function during treatment of intention in aphasia*. J Cogn Neurosci., 2005. **17**(3): p. 392-406.
23. Fridriksson, J., *Preservation and modulation of specific left hemisphere regions is vital for treated recovery from anomia in stroke*. J Neurosci., 2010. **30**(35): p. 11558-11564.
24. Menke, R., et al., *Imaging short- and long-term training success in chronic aphasia*. BMC.Neurosci., 2009. **10**: p. 118.
25. Crosson, B., et al., *Regional changes in word-production laterality after a naming treatment designed to produce a rightward shift in frontal activity*. Brain and Language, 2009. **111**(2): p. 73-85.

26. Basvaraju, A. and S.R. Pratt, *Aphasia data extraction report*. Center for Health Equity Research & Promotion and Geriatric Research Education & Clinical Center, VA Pittsburgh Healthcare, 2008.
27. Yu, W., et al., *Prevalence and costs of chronic conditions in the VA health care system*. Med.Care Res.Rev., 2003. **60**(3 Suppl): p. 146S-167S.
28. Kauhanen, M.L., et al., *Aphasia, depression, and non-verbal cognitive impairment in ischaemic stroke*. Cerebrovascular Diseases, 2000. **10**(6): p. 455-461.
29. Pulvermuller, F., et al., *Therapy-related reorganization of language in both hemispheres of patients with chronic aphasia*. NeuroImage, 2005. **28**(2): p. 481-489.
30. Pulvermuller, F., et al., *Constraint-induced therapy of chronic aphasia after stroke*. Stroke, 2001. **32**(7): p. 1621-1626.
31. Massaro, M. and C.A. Tompkins, *Feature Analysis for Treatment of Communication Disorders in Traumatically Brain-Injured Patients: An Efficacy Study*. Clinical Aphasiology, 1994. **22**: p. 245-256.
32. Collins, A.M. and E.F. Loftus, *A Spreading-Activation Theory of Semantic Processing*. Psychological Review, 1975. **82**(6): p. 407-428.
33. Levelt, W.J., A. Roelofs, and A.S. Meyer, *A theory of lexical access in speech production*. Behavioral and Brain Sciences, 1999. **22**(1): p. 1-38.
34. Dell, G.S., et al., *Lexical access in aphasic and nonaphasic speakers*. Psychological Review, 1997. **104**(4): p. 801-838.
35. Haarbauer-Krupa, J., et al., *Cognitive rehabilitation therapy: Late stages of recovery*, in *Head injury rehabilitation: Children and adolescents*, M. Ylvisaker, Editor 1985, College Hill Press: San Diego, CA. p. 311-340.
36. Peach, R.K. and K.A. Reuter, *A discourse-based approach to semantic feature analysis for the treatment of aphasic word retrieval failures*. Aphasiology, 2010. **24**(9): p. 971-990.
37. Boyle, M., *Semantic feature analysis treatment for aphasic word retrieval impairments: what's in a name?* Topics in Stroke Rehabilitation, 2010. **17**: p. 411-422.
38. Lowell, S., P.M. Beeson, and A.L. Holland, *The efficacy of a semantic cueing procedure on naming performance of adults with aphasia*. American Journal of Speech-Language Pathology, 1995. **4**:109-114.
39. Edmonds, L.A. and S. Kiran, *Effect of semantic naming treatment on crosslinguistic generalization in bilingual aphasia*. Journal of Speech, Language, and Hearing Research, 2006. **49**(4): p. 729-748.
40. Kiran, S. and P.M. Roberts, *Semantic feature analysis treatment in Spanish-English and French-English bilingual aphasia*. Aphasiology, 2010. **24**: p. 231-261.
41. Rider, J.D., et al., *Using semantic feature analysis to improve contextual discourse in adults with aphasia*. American Journal of Speech Language Pathology, 2008. **17**: p. 161-172.
42. Wambaugh, J. and M. Ferguson, *Application of semantic feature analysis to retrieval of action names in aphasia*. Journal of Rehabilitation and Research Development, 2007. **44**: p. 381-394.
43. Hashimoto, N. and A. Frome, *The use of a modified semantic features analysis approach in aphasia*. Journal of Communication Disorders, 2011. **44**: p. 459-469.
44. Abel, S., K. Willmes, and W. Huber, *Model-oriented naming therapy: Testing predictions of a connectionist model*. Aphasiology, 2007. **21**: p. 411-447.
45. Wambaugh, J.L., et al., *Effects of two lexical retrieval cueing treatments on action naming in aphasia*. Journal of Rehabilitation and Research Development, 2002. **39**: p. 455-466.
46. Ivarie, J.J., *Effects of proficiency rates on later performance of a recall and writing behavior*. Remedial and Special Education, 1986. **7**(5): p. 25-30.
47. Bloom, B.S., *The hands and feet of genius*. Educational Leadership, 1986. **43**: p. 70-77.
48. Olander, C.P., et al., *Retention among college students: A comparison of traditional versus precision teaching*. Journal of Precision Teaching, 1986. **6**(4): p. 80-82.
49. Binder, C., *Behavioral fluency: A new paradigm*. Educational Technology, 1993. **33**(10): p. 8-14.
50. LaBerge, D. and S.J. Samuels, *Toward a theory of automatic information-processing in reading*. Cognitive Psychology, 1974. **6**: p. 293-323.
51. White, O.R., *Aim*star wars (setting aims that compete)*. Journal of Precision Teaching, 1984. **5**: 55-63.
52. Kunkel, A., et al., *Constraint-induced movement therapy for motor recovery in chronic stroke patients*. Archives of Physical Medicine and Rehabilitation, 1999. **80**(624-628).
53. Taub, E., G. Uswatte, and R. Pidikiti, *Constraint induced movement therapy: A new family of techniques with broad application to physical rehabilitation—A clinical review*. Journal of Rehabilitation and Research Development, 1999. **36**: p. 237–251.

54. Taub, E., G. Uswatte, and T. Elbert, *New treatments in rehabilitation founded on basic research*. Nature Reviews: Neuroscience, 2002. **3**: p. 228-236.
55. Meinzer, M., et al., *Long-term stability of improved language functions in chronic aphasia after constraint-induced aphasia therapy*. Stroke, 2005. **36**(7): p. 1462-1466.
56. Meinzer, M., et al., *Extending the Constraint-Induced Movement Therapy (CIMT) approach to cognitive functions: Constraint-Induced Aphasia Therapy (CIAT) of chronic aphasia*. NeuroRehabilitation., 2007. **22**(4): p. 311-318.
57. Meinzer, M., et al., *Intensive language training enhances brain plasticity in chronic aphasia*. BMC Biology, 2004. **2**.
58. Kilgard, M.P. and M.M. Merzenich, *Cortical map reorganization enabled by nucleus basalis activity*. Science, 1998. **279**(5357): p. 1714-1718.
59. Merzenich, M.M., et al., *Topographic reorganization of somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation*. Neuroscience, 1983. **8**: p. 33-55.
60. Bach-y-Rita, P. and S.W. Kercel, *Sensory substitution and the human-machine interface*. Trends Cogn Sci., 2003. **7**(12): p. 541-546.
61. Jenkins, W.M., *Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation*. Journal of Neurophysiology, 1990. **63**: p. 82-104.
62. Kleim, J.A., T.A. Jones, and T. Schallert, *Motor enrichment and the induction of plasticity before or after brain injury*. Neurochemistry Research, 2003. **28**(1757-1769).
63. Kiran, S. and C.K. Thompson, *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, 2003. **46**(4): p. 773-787.
64. Snell, C., K. Sage, and M.A.L. Ralph, *How many words should we provide in anomia therapy? A meta-analysis and a case series study*. Aphasiology, 2010. **24**(9): p. 1064-1094.
65. Beeson, P., Robey, R., *Meta-analysis of aphasia treatment outcomes: Examining the evidence*, in *Clinical Aphasiology Conference 2008*: Jackson Hole, Wy.
66. Fillingham, J.K., K. Sage, and M.A. Lambon Ralph, *The treatment of anomia using errorless learning*. Neuropsychol.Rehabil., 2006. **16**(2): p. 129-154.
67. Van de Sandt-Koenderman, W.M., et al., *MAAS (Multi-Axial Aphasia System): Realistic goal setting in aphasia rehabilitation*. International Journal of Rehabilitation Research, 2008. **31**: p. 314-320.
68. Helm-Estabrooks, N., *Cognition and aphasia: A discussion and a study*. Journal of Communication Disorders, 2002. **35**: p. 171-186.
69. Hinckley, J.J. and J. Carr, *Differential contributions of cognitive abilities to success in skill-based versus context-based aphasia treatment*. Brain and Language, 2001. **79**: p. 3-9.
70. Seniow, J., M. Litwin, and M. Lesniak, *The relationship between non-linguistic cognitive deficits and language recovery in patients with aphasia*. Journal of the Neurological Sciences, 2009. **283**: p. 91-94.
71. Hinckley, J.J., J.P. Patterson, and T.H. Carr, *Differential effects of context- and skill-based treatment approaches: Preliminary findings*. Aphasiology, 2001. **15**(5): p. 463-476.
72. Meinzer, M., et al., *Recent developments in functional and structural imaging of aphasia recovery after stroke*. Aphasiology, 2011. **25**(3): p. 271-290.
73. Fridriksson, J., Baker, J. M., & Moser, D., *Cortical mapping of naming errors in aphasia*. Human Brain Mapping, 2009. **30**: p. 2487-2498.
74. Fridriksson, J., Bonilha, L., Baker, J. M., Moser, D., & Rorden, C., *Activity in preserved left hemisphere regions predicts anomia severity in aphasia*. Cerebral Cortex, 2010. **20**: p. 1013-1019.
75. Postman-Caucheteux, W.A., et al., *Single-trial fMRI shows contralesional activity linked to overt naming errors in chronic aphasic patients*. J Cogn Neurosci., 2010. **22**(6): p. 1299-1318.
76. Richter, M., W.H.R. Miltner, and T. Straube, *Association between therapy outcome and right-hemispheric activation in chronic aphasia*. Brain, 2008. **131**: p. 1391-1401.
77. Saur, D., et al., *Dynamics of language reorganization after stroke*. Brain, 2006. **129**: p. 1371-1384.
78. Basso, A. and A. Caporali, *Aphasia therapy or the importance of being earnest*. Aphasiology, 2001. **15**(4): p. 307-332.
79. Basso, A., *How intensive/prolonged should an intensive prolonged treatment be?* Aphasiology, 2005. **19**(10/11): p. 975-984.
80. Thompson, C.K. and L. Shapiro, *Treating agrammatic aphasia within a linguistic framework: Treatment of underlying forms*. Aphasiology, 2005. **19**(10/11): p. 1021-1036.

81. Barthel, G., et al., *Intensive language therapy in chronic aphasia: Which aspects contribute most?* *Aphasiology*, 2007. **22**(4): p. 408-421.
82. Maas, E., Robin, D., Austermann Hula, S., Freedman, S., Wulf, B., Ballard, K., Schmidt, R., *Principles of Motor Learning in Treatment of Motor Speech Disorders*. *American Journal of Speech-Language Pathology*, 2008. **17**: p. 277–298.
83. Park, J.H., & Shea, C. H., *Effect of practice on effector independence*. *J. of Motor Behav*, 2003. **35**:33-40.
84. Park, J.H., & Shea, C. H., *Sequence learning: Response structure and effector transfer*. *Q.J.Exp.Psychol.A*, 2005. **58**: p. 387-419.
85. Schmidt, R.A. and T.D. Lee, *Motor Control and Learning: A Behavioral Emphasis* 2005, Champaign, IL: Human Kinetics.
86. Porch, B., *Porch index of communicative ability* 2001, Albuquerque, NM PICA Programs.
87. Baker, J., Rorden, C., Fridriksson, J., *Using transcranial direct current stimulation to treat stroke patients with aphasia*. *Stroke*, 2010. **41**: p. 1229-1236.
88. Verdon, V., Schwartz, S., Lovblad, K-O., Hauert, C-A., Vuilleumier, P., *Nueroanatomy of hemispatial neglect and its functional components: a study using voxel-based lesion-symptom mapping*. *Brain*, 2010. **133**: p. 880-894.
89. Reilly, J., Park. H., Bennet, J., Towler, S. Benjamin, M., Crosson, B., *Broca and beyond: Decomposing the Lesion Correlates of Narrative Speech Fluency in Aphasia*. in preparation.
90. Goodglass, H., E. Kaplan, and B. Barresi, *The assessment of aphasia and related disorders*. Vol. 3rd 2001, Baltimore: Lippincott, Williams, & Wilkins.
91. Szekely, A., et al., *A new on-line resource for psycholinguistic studies*. [References] 2004: *Journal of Memory and Language*. Vol.51(2), Aug 2004, pp. 247-250.
92. Roach, A., M.F. Schwartz, and N. Martin, *The Philadelphia Naming Test: Scoring and rationale*, in *Clinical Aphasiology*, M.L. Lemme, Editor 1996, Pro-Ed: Austin, TX. p. 121-133.
93. De Ayala, R.J., *The theory and practice of item response theory*, 2009, New York: The Guilford Press.
94. Mirman, D., et al., *A large, searchable, web-based database of aphasic performance on picture naming and other tests of cognitive function*. *Cognitive Neuropsychology*, 2010. **27**(495-504).
95. Fraser, C. and R.P. McDonald, *NOHARM: A windows program for fitting both unidimensional and multidimensional normal ogive models of latent trait theory*, 2003, Niagra College: Welland, ON.
96. Linacre, J.M. and B.D. Wright, *WINSTEPS: Multiple-choice, rating scale, and partial credit Rasch analysis*, 2003, MESA Press Chicago.
97. Linacre, J.M., *WINSTEPS*, 2009, Winsteps.com: Chicago.
98. Wright, B.D. and J.M. Linacre. *Reasonable mean-square fit values*. 1994 2009/11/30/ [cited 8; 3:[370]. Available from: <http://rasch.org/rmt/rmt83.htm>.
99. Swinburn, K., G. Porter, and D. Howard, *Comprehensive aphasia test*. 2004, NY: Psychology Press.
100. Nicholas, L.E. and R.H. Brookshire, *A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia*. *Journal of Speech and Hearing Research*, 1993. **36**: 338-350.
101. Brookshire, R.H. and L.E. Nicholas, *Speech sample size and test-retest stability of connected speech measures for adults with aphasia*. *Journal of Speech and Hearing Research*, 1994. **37**(2): p. 399-407.
102. Howard, D. and C. Gatehouse, *Distinguishing semantic and lexical word retrieval deficits in people with aphasia*. *Aphasiology*, 2006. **20** (9/10/11): p. 921-950.
103. Jefferies, E. and M.A. Lambon Ralph, *Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison*. *Brain*, 2006. **129** p. 2132-2147.
104. Murray, L.L., *Attention deficits in aphasia: presence, nature, assessment, and treatment*. *Seminars in Speech and Language*, 2002. **23**(2): p. 107-116.
105. Yushkevich, P., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G., *User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability*. *NeuroImage*, 2006. **31**: p. 1116-1128.
106. Rorden, C., *MRICro: Imaging analysis toolbox.*, 2007.
107. Brunner, E., Munzel, U., Puri, M.L., *Rank-score tests in factorial designs with repeated measures*. *Journal of Multivariate Analysis*, 1999. **70**: p. 286-317.
108. Rorden, C., Bonilha, L., Nichols, T., *Rank-order versus mean based statistics for neuroimaging*. *NeuroImage*, 2007. **35**: p. 1531-1537.
109. Rorden, C., Karnath, H-O., Bonilha, L., *Improving lesion-symptom mapping*. *Journal of Cognitive Neuroscience*, 2007. **19**: p. 1081-1088.
110. Ting, N., *Dose Finding in Drug Development*. 2006, New York: Springer.