DEVELOPMENT AND VALIDATION OF A COMPUTER-BASED APHASIA THERAPY FEATURING AUTOMATIC SPEECH RECOGNITION

by

Francesco Usai

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
July 2024

Dalhousie University is located in Mi'kma'ki, the
ancestral and unceded territory of the Mi'kmaq.
We are all Treaty people.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Throughout the Western world, the systems devised to provide treatment for aphasia limit support to the acute phase, and only for a fraction of the recommended sessions. This goes against solid evidence showing that key to a successful recovery of language function is intensive, long-term therapy. The high cost of private therapy, lack of personnel, and barriers to transportation further limit accessibility to adequate services. Computer-based therapies (CBATs) can address these problems, serving as an adjunct to traditional therapy provided by professionals or as an alternative for those people who have no access to this service. However, while enormous progress has been made in designing CBATs, recent works have highlighted key factors that limit their effectiveness: first, people often find them to be boring to use, which hinders motivation to adhere with the recommended training regime; second, they do not provide feedback on how well people speak, a crucial factor for therapy to be effective. This dissertation is part of a broader project whose mission is to develop a CBAT that improves currently available applications, by focusing on usability and — critically — feedback based on automated speech recognition (ASR). The research presented here focused on the initial user-centered design of a CBAT prototype, usability testing, and preliminary testing of the efficacy of the ASR-based feedback. To this end, this dissertation set out to: conduct a needs-finding analysis to identify the most common unmet needs of individuals with aphasia, along with barriers and facilitators in the use of technologies that can inform the design of a prototype (Chapter 2); develop and test the *usability* of a prototype of home-based therapy, featuring an ASR engine, geared for treatment of word-finding skills (Chapter 3); conduct a preliminary test of the efficacy of delivering feedback via an ASR (Chapter 4). The work conducted as part of this dissertation will hopefully provide the basis upon which to integrate additional improvements (i.e., novel game-mechanics, wider variety of tasks and stimuli, adaptive level of difficulty, optimized delivery of feedback etc.) and to make the training more enjoyable and, ultimately, effective.

LIST OF ABBREVIATIONS USED

AAC — Augmented Alternative Communication

AATD — Adapted Alternating Treatment Design

ADHD — Attention Deficit Hyperactivity Disorder

AG — Angular Gyrus

API — Application Programming Interface

AQ — Aphasia Quotient

ASR — Automatic Speech Recognition

CBAT — Computer-Based Aphasia Therapy

IFG — Inferior Frontal Gyrus

IPNP — International Picture Naming Project

IQR — Inter-Quartile Range

IRR — Inter-Rater Reliability

ISO — International Standard Organization

ITS — Inferior Temporal Sulcus

MS — Multiple Sclerosis

MTG — Middle Temporal Gyrus

MVP — Minimum Viable Product

OS — Operative System

PAND — Percentage of All Non-Overlapping Data

PCA — Phonological Component Analysis

PMC — Premotor Cortex

PND — Percentage of Non-Overlapping Data

PwA — People with Aphasia

RCT — Randomized Controlled Trial

RT — Reaction Time

SDT — Self-Determination Theory

SFA — Semantic Feature Analysis

SLT — Speech-Language Therapy

SMG — Superior Marginal Gyrus

SSH — Secure Shell

SSL — Secure Socket Layer

STG — Superior Temporal Gyrus

STS — Superior Temporal Sulcus

SUS — System Usability Scale

SVZ — Subventricular Zone

WAB-R — Western Aphasia Battery-Revised

WWC — What Works Clearinghouse

ACKNOWLEDGEMENTS

Going through the PhD program was a long (it couldn't get any longer!) and very though journey. Thinking back about all that has happened to me during the last ten years, about how I have changed both personally and professionally, I cannot but be extremely grateful for having had the chance to embark on this journey which allowed me to learn vital lessons that reach far beyond the narrow area that I expected to develop in when I decided to pursue my graduate studies. Needless to say, the outcome of this journey was strongly — and positively — influenced by many people, to whom I can finally express my deepest gratitude for their support, guidance, and company.

First, this journey started thanks to my supervisor, Aaron, who granted me the opportunity to join his laboratory to pursue my research interests. It could continue only thanks to your unending patience during the long erratic and frustrating phases of this program, and to the support that I received from you when I had to find the right focus. Thank you also for the numerous opportunities to continue growing as a student, for giving me the opportunity to join your lab, and to learn from you.

Thank you to professors Gail Eskes and Michael Kiefte, for being regularly available to put your expertise in service of my work, providing valuable guidance throughout these many years.

Thank you to all the members of Aphasia Nova Scotia, which I had the pleasure to serve as a volunteer during the early years of my PhD. I am extremely grateful to have had the opportunity to be part of your community. This, among other invaluable things, allowed me to take a privileged perspective and understanding of what it means to live with aphasia. Special thanks to Judy Arbique, founder of Aphasia NS, for her immense support and friendship, for sharing and pursuing her vision to ameliorate the conditions of those living with aphasia, and for being such an inspiring model of perseverance and optimism in the face of seemingly unsurmountable challenges.

Thanks to the staff of the Aphasia Institute of Ontario, of the Cowichan Brain Injury community, and to Dr. Lisa Harris, for the immense, extremely valuable support recruiting participants for my studies throughout Canada.

Thanks to the several students who supported me in the process of data collection and analysis with dedication.

Thanks to the staff of the library at Richterswil (where I lived for the last three years). You invited me to use the spaces and resources at your disposal during the last, most challenging part of my dissertation. I did greatly enjoy the peace I found at the library, which was essential to get through the long, intense phases of writing most chapters of this dissertation.

Thanks to my friends, old and new. You blessed me with the chance to be part of your life, we shared many important experiences and beautiful memories over these years. You played a big part in making me fit to complete this journey!

Thanks to my parents and relatives, who have always been a firm source of grounding, peace and joy.

Finally, and most importantly, thanks to my partner Lorenza, for embarking in so many important exciting journeys together, especially that of starting a family. The immense love that I received from you and Antonello kept me afloat during the difficult moments, and made the entire experience much more enjoyable. Your patience and self-sacrifice, all while building a career for yourself and raising a son, allowed me to put in the extra effort required to get to the final step. I am extremely grateful and feel very fortunate for sharing my life with you.

Chapter 1 – INTRODUCTION

Aphasia is an acquired neurological disorder that disrupts a person's ability to communicate. It can manifest as a selective impairment in speech production, speech comprehension, reading, writing, or as a combination of these deficits (Chapey, 2008a; Simmons-Mackie, 2017).

Aphasia has a shocking effect on a person's life. As it frequently results from stroke (Heart & Stroke, 2017) or traumatic brain injury (Simmons-Mackie, 2017), people with aphasia (PwA) have to deal with a rather abrupt transition from a life in which communication, in all its sophisticated facets, is taken for granted, to one in which even the easiest acts of communication feel like challenges that are impossible to take on. The immense frustration that is caused by this reality is compounded by the fact that a person's struggle using language to communicate is frequently attributed to reduced levels of intelligence among those who know little about aphasia (National Aphasia Association, 2020, 2022; Simmons-Mackie et al., 2002). This is a very common state of affairs: surveys on this topic showed that, in Canada, as of 2015, about 94.3% of interviewees reported not knowing about aphasia (Patterson et al., 2015), while in 2020 in the United States this statistic was about 86% (National Aphasia Association, 2020; Simmons-Mackie, 2017).

Many PwA have to endure this situation for the rest of their lives, as it is common for members of their family and friends to instinctively bypass them in all matters that are of direct concern (Hartwell, 2015; Parr, 2007; Simmons-Mackie et al., 2007). This might explain why many PwA question their intelligence, feel inadequate in front of other people, and are frustrated when experiencing the communication barriers that seem to act as an invisible wall standing between them and everybody else (Simmons-Mackie, 2017).

Aphasia forces people to completely reboot their lives, as most of them can hardly maintain their old job or find a new one. In the relatively rare cases when a new position is offered, it is not as economically remunerative as the previous one (Caporali & Basso, 2003; Graham et al., 2011). Economic hardship is only the most obvious consequence of the loss of a job. To most people, their job provides most opportunities to establish and nurture social connections, acts as a powerful source of meaning, and it is key to maintaining a healthy sense of worth and a positive image of themselves (Vestling et al., 2003). The lack of self-confidence, economic hardship and a reduced sense of self-worth induces many PwA to retreat away from what it feels to them being a rather unfriendly environment, leading to social isolation, exacerbating the sense of loneliness and helplessness, and favoring the onset of depression, which for most people becomes a chronic condition (Simmons-Mackie, 2017).

Aphasia goes as far as to undermine even the strongest, most intimate relationships. Couples where one of the partners has aphasia are more likely than others to report lower quality of life, and present higher rates of divorce (Christensen & Anderson, 1989; Dorze & Brassard, 1995; Williams & Freer, 1986), further aggravating the phenomenon of social isolation. Therefore, it comes as no surprise that, when compared with other stroke survivors (Hilari, 2011) or with people with other disabilities (Lam & Wodchis, 2010), PwA report lower quality of life.

Incidence, Prevalence, and Costs

Aphasia is most frequently caused by stroke (Heart & Stroke, 2017)—a phenomenon that in Canada occurs to someone every 9 minutes, largely affecting people over age 65 (Dickey et al., 2010; Heart & Stroke, 2017; Mozzafarian, 2016). Approximately 34% of stroke survivors are diagnosed with aphasia (Flowers et al., 2016), with recent estimates reporting approximately 165,000 individuals living with aphasia in Canada alone, and 2.5 million in North America (Simmons-Mackie, 2017). Code and Petheram estimated that, throughout the developed world, the incidence of aphasia ranges between 0.02-0.06%, while prevalence ranges between 0.1-0.4% (Code & Petheram, 2011).

Efforts to persuade citizens to embrace healthier lifestyle choices have recently ramped up, holding the promise to help curb the incidence of stroke, and aphasia, in the future (Heart & Stroke, 2020). Nonetheless, the prevalence of aphasia is expected to rise dramatically due to the conjunction between the increase in life expectancy and the improvement in medical practices that result in lower rates of deaths from stroke (Mozzafarian, 2016). Therefore, in the future, governments (and society writ large) will be required to exert ever increasing efforts to face the human and financial burden of stroke-induced aphasia (G. Howard & Goff, 2012; Wafa et al., 2020). Indeed, studies have estimated that by the year 2030, in the US alone, the total costs due to management of stroke cases will amount to 184 billion, up by a factor of three relative to the year 2012 (Ovbiagele et al., 2013). As discussed by Simmons-Mackie (2017), management of stroke-induced aphasia will claim a disproportionate amount of this budget, as it is well documented that PwA require more physician visits (Kagan et al., 2017), make greater use of rehabilitation services (Dickey et al., 2010), have longer inpatient stays (Bersano et al., 2009; Ellis et al., 2012; Flowers et al., 2016), and are more likely to be readmitted to hospital after having being discharged (Bersano et al., 2009) than stroke survivors without aphasia.

The Context for this Dissertation

Throughout the Western world, the systems devised to provide treatment for aphasia limit support prevalently to the acute and subacute phases (i.e., up to six months post onset), and only for a fraction of the recommended sessions (Simmons-Mackie, 2017; Code & Petheram, 2011). This goes against solid evidence showing that key to a successful recovery of language function is intensive, long-term therapy (Allen et al, 2012; Brady et al; 2022). Currently, the limited access to services devoted to language therapy is one of the biggest barriers that PwA experience, thwarting their chances to recover communication skills, and, ultimately, to realizing their full potential as individuals, members of their family and of society: in short, to living a successful life with aphasia (Simmons-Mackie, 2017; Arcury et al., 2005).

As it will be discussed in depth in subsequent sections of this chapter, work is underway to transition from a service delivery model that is heavily focused on acute/subacute care, where access to therapy is limited by physical and financial barriers, and that is predominantly provider-centered, to one that is geared towards the management of aphasia as a chronic condition, that puts an emphasis on promoting clients' autonomy by prioritizing their needs, interests and goals, and that provides them with resources to manage their condition. Computer-based therapies for aphasia (CBATs) can help address these problems as they provide people with necessary levels of access to a potentially cost-effective therapy from the comfort of their home (Des Roches & Kiran, 2017; Lavoie et al., 2017). This dissertation fits within a wider project whose mission is to harness the immense potential of CBATs to remove any physical and financial barrier that currently puts therapy out of reach to most PwA, extending access to anyone with an internet connection and a computer or tablet.

However, while over the last decade enormous progress has been made in designing CBATs, recent works have highlighted some key factors that might limit their effectiveness (Palmer et al., 2019, 2020; Harrison et al., 2020): first, people often find them to be boring to use, which can hinder motivation to put in the necessary hours of training; second, they do not provide feedback on how well people speak, a crucial factor for therapy to be effective (see section titled 'Limitations of Computer-Based Aphasia Therapies', at page 51, for a more detailed discussion on this topic). The project that this thesis is related to was motivated by the need to transcend these limitations by: harnessing principles of game design to maximize engagement, fun, and in turn promote intrinsic motivation in adhering to therapy; integrating an automatic speech recognition (ASR) system into the CBAT.

As a first step towards this goal, the scope of this thesis is to develop a prototype that can deliver therapy for naming and word-finding via a confrontation naming task. This choice was motivated by three reasons: a) word naming is the most frequent symptom of aphasia (Goodglass & Wingfield, 1997; Pedersen et al., 2003), making even this preliminary solution appealing to a significant portion of PwA; b) confrontation naming is a very simple exercise to implement on computer, making it ideal for a first prototype; c) the task requires participants speaking, therefore it is an ideal ground to test the ASR. The prototype is intended to integrate an ASR system, and to carefully assess the feasibility of using it to deliver feedback on naming accuracy during self-managed therapy by a) quantifying its accuracy; b) evaluating the opinions of PwA on its accuracy and usefulness; c) quantifying the extent to which PwA can benefit from ASR-based feedback, relative to a person's self-monitoring abilities.

The following sections present an overview of the cognitive and neural bases of word production, of the key principles underpinning recovery from aphasia, and of the main approaches to treatment of anomia. This will be useful to discuss the mechanisms behind the specific task that was implemented in the prototype whose design and preliminary testing this thesis is about.

Anomia

Anomia is defined as impaired word retrieval, and manifests as a difficulty in finding the right words, and/or as word production that is laden with errors (Chapey, 2008a; Damasio, 1998; Goodglass & Wingfield, 1997; Laine & Martin, 2006). At stroke onset, anomia frequently co-occurs with other, more severe communication deficits which could be due to: the presence of telegraphic speech; an inability to correctly process function words, which is key to produce and comprehend sentences (agrammatism); severe deficits in speech comprehension accompanied by fluent, yet meaningless speech production (Wernicke's aphasia) (Baker et al., 2008). Furthermore, while spontaneous recovery frequently lifts most PwA from the effects of the most severe syndromes, anomic symptoms are often present in those who have not recovered completely (Pedersen et al., 2003), making it the most common type of aphasia among chronic stroke survivors.

Although considered in isolation anomia is arguably the mildest form of aphasia, still difficulty with word retrieval does often result in communication breakdowns (Grossman, 2014), which are a cause of frustration among those who struggle with it.

Behavioral Symptoms of Anomia

Saying a word aloud entails converting thoughts into speech. We currently understand how this feat is accomplished in terms of a series of cognitive processes operating on different types of mental representations which are distributed over a wide, bilateral network in the brain (Fridriksson et al., 2018; Hickok & Poeppel, 2007; Kiran & Thompson, 2019). Each cognitive process itself relies on the coordinated activity of specific areas within this network. As a result, every person with anomia shows a characteristic error profile, as the locus and size of the damage determine which aspects of word production are disrupted.

The most frequent types of errors that result from impairments in word-retrieval processes are: semantic errors, wherein the output is semantically related to the intended word (i.e. *dog* instead of *cat*); phonological errors, wherein the output is phonologically related to the intended word (i.e. *mat* instead of *cat*); mixed errors, wherein the output is both semantically and phonologically related to the intended word (i.e. *rat*, instead of *cat*); unrelated errors, wherein there is no obvious relationship between output and target word (i.e. *house* instead of *cat*); and nonwords, namely utterances that do not belong to any vocabulary (i.e. *zat* instead of *cat*).

The Cognitive Underpinnings of Anomia

Modern accounts of speech production (and repetition) explain a wide range of behavioral variants of anomia (Dell et al., 1997, 2004; Foygel & Dell, 2000; N. Martin & Dell, 2019; Rapp & Goldrick, 2000), of aphasia more in general (Chang & Lambon Ralph, 2020; Roelofs, 2014; Ueno et al., 2011), provide detailed insights into the temporal dynamics of the various cognitive processes involved in word production in both healthy people and PwA (Levelt, 1992; Levelt et al., 1999; Roelofs, 1992, 1997), or tackle a specific aspect of speech processing (i.e. how sensory processing of syllables and phonemes drives motor programming during speech production; Hickok, 2012).

The goal of this section is not to provide a detailed review and comparison between these models, nor to attempt to synthesize and reconcile their findings (to this end, a number of recent reviews are useful; Dell et al., 2014; Kerr et al., 2022; Roelofs & Ferreira, 2019). Rather, the goal is to provide a broad view of the mechanisms underpinning word production, and how their breakdown explains the breadth of symptoms that characterize anomia. This is key to understanding the rationale behind therapy targeting anomia. To the best of my knowledge, the cognitive model that currently provides the most detailed account of the wide variety of symptoms of anomia is the *interactive 2-step model* that was originally proposed, and subsequently refined over the last three decades, by Dell and colleagues.

The Interactive 2-Step Model

One of the first models of speech production that was computationally instantiated and tested is the interactive 2-step model, first proposed in the context of sentence production (Dell, 1986; Dell & O'Seaghdha, 1992), then narrowed down to explain how single word production occurs (Dell et al., 1997, 2004; Dell & O'Seaghdha, 1992; Foygel & Dell, 2000; N. Martin & Dell, 2019; Schwartz et al., 2006).

A key assumption of this model (which is shared by most models of word-production) is that information about words is broken down into distinct, cognitively and neurally independent yet tightly connected systems: semantics, which encodes the meaning of words; lexical, which encodes grammatical properties of a word (i.e. noun/verb status, gender); and phonology, which encodes information about the sequence of phonemes and syllables that represent the auditory makeup of a word. This assumption is backed by a wealth of converging evidence from studies adopting complementary approaches such as analysis of error patterns (Fromkin, 1971; Garrett, 1975), lesion-symptom mapping (Dell et al., 2013; Halai et al., 2017, 2018), task-based activation mapping (Saur et al., 2008; see also reviews by Hickok and Poeppel 2007 and Indefrey 2011), and electrophysiology (Indefrey, 2011; Indefrey & Levelt, 2004; Levelt et al., 1999), showing that semantic, lexical and phonological processing engage different areas of the brain at different times. In the model, each level of representation is assigned a dedicated layer of nodes (computational units), with each node encoding a specific feature — a semantic feature (i.e. being an animal; having four legs; having a white fur etc.), a lemma (i.e. noun/verb status; gender, etc.), or a phoneme (i.e., /p/, /d/, etc.).

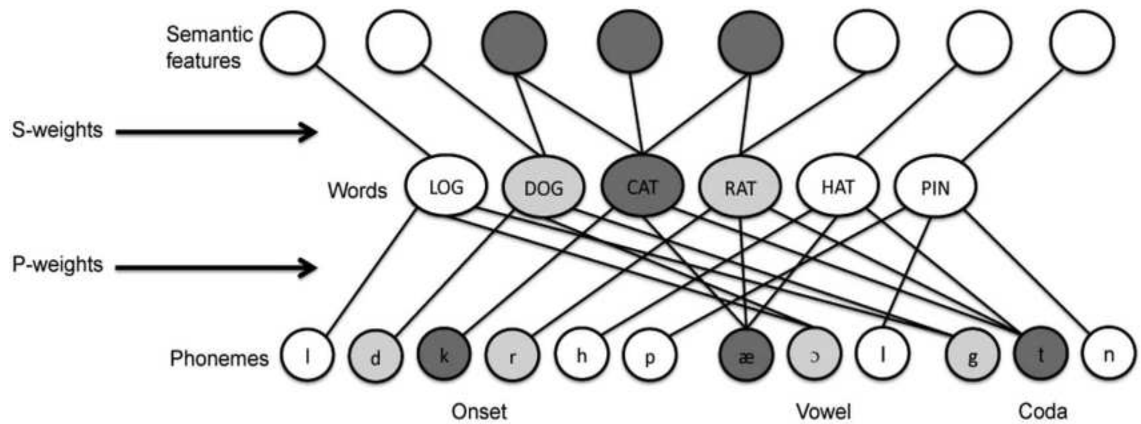

Figure 1. 1 - Architecture of the interactive 2-step model of word-production. Computational units encoding semantic features are connected with associated lexical features, which in turn are connected to associated phonological features. The strength of the connections, which reflects the balance between signal and noise that is transmitted, is encoded in the s- and p-weights, respectively. Copyright © 2024 by APA. Reproduced with permission. (Middleton et al., 2015).

The model takes its name from the fact that word retrieval is decomposed in two sequential steps: lemma retrieval, followed by phonological retrieval. Three computational principles regulate how information is processed: spreading activation, cascading, and interactivity.

As mentioned earlier, word production entails converting thoughts into speech. In the interactive 2-step model, a specific thought that needs to be expressed verbally (which might or might not be triggered by external stimuli, like a picture, sound etc.) induces an increase in the activation level of all the semantic features/units that are associated with it. This activation is then transmitted to the lexical level of representation, and from here, to the phonological level. The transmission of information is governed by the principle of spreading activation, which, computationally, consists of: spreading, summation, and decay (Dell, 1986; Roelofs, 1992). Spreading entails the transmission of activity from a node at one level of representation to all nodes at adjacent levels that encode related information (see Figure 1.1 for a graphical representation of the model). The strength of the connections between semantic and lexical levels of representation (i.e., the amount of activity of a semantic node that is transmitted to connected lexical nodes) is encoded in the *s-weights*, while the *p-weights* encode the strength of the connections between lexical and phonological levels. To better understand the dynamics of spreading activation, let's consider the example wherein one intends to name the word 'cat'. In this case, the unique combination of nodes encoding the semantic features associated with this word will get active, transmitting a signal that is proportional to their level of activity to all the nodes encoding lexical information that they are connected with. This means that not only the lemma associated with the word 'cat' will receive input from semantic nodes, but also all those sharing at least one semantic feature with it. At this point, summation of all the inputs will determine the total activation level of a given lexical node. Normally, this will be highest for the target lemma since it will receive more input than nodes sharing semantic features. Finally, the process of decay determines how fast a node's level of activation fades off. Ultimately, only one of the possible words is uttered. This poses the problem of which word, among those that are concurrently activated at the lemma and phonological levels, is selected. In the interactive 2-step model, this problem is addressed by pre-specifying a time interval during which lemmas receive input from higher (semantic) and lower (phonological) level nodes. At the end of this interval, the lemma with the highest level of activation is selected by receiving an extra boost of activation, which in turn will amplify the activation of its associated phonemes (and semantic features). Then, the process of phonological encoding will continue for another fixed amount of time, at the end of which the sequence of phonemes with the highest level of activation will be selected.

The concepts of cascading and interactivity were introduced to explain a set of empirical findings, such as the lexical bias (the tendency for errors to be words that do belong to a person's vocabulary, as opposed to turn out as nonwords) and mixed errors, that were difficult to account for by other popular models of word production that hypothesized that phonological processing could start only when lexical selection had occurred (Levelt, 1992; Levelt et al., 1999; Roelofs, 1992). Cascading refers to the fact that information processing at different levels of representation can overlap in time (i.e. activated units at higher levels transmit information to lower levels before selection is completed). Interactivity refers to the fact that processing at lower levels can influence processing at higher levels.

Cascading and interactivity are not only essential concepts to explain the spectrum of errors that are frequently observed in people with anomia (see next subsection for a detailed account of how the model explains these errors), but also other behavioral (Cutting & Ferreira, 1999; Dell et al., 2014; Goldrick & Blumstein, 2006; Peterson & Savoy, 1998) and neurophysiological (Kerr et al., 2022; Miozzo et al., 2015; Nozari & Pinet, 2020; Strijkers et al., 2017) phenomena, further corroborating the theoretical validity of such mechanisms.

To illustrate how cascading and interactivity influence semantic, lexical and phonological processing, we consider the example in which the target word is 'cat'. Initially, the set of semantic features associated with the concept of 'cat' get activated, with activation immediately spreading to the lexical network encoding lemmas, and from there down to the phonological network. In such a situation, the lemma associated with the target word will be the most activated. Due to the dynamics of spreading activation, lemmas that are semantically related to the target (e.g., dog, rat, lion) will also be activated in a manner that is proportional to the degree of semantic similarity to the target word. In turn, due to cascading, activated lexical nodes will transmit activation down to their associated phonological nodes before lexical selection has occurred, with the phonemes associated with the target word receiving the highest activation from the lexical level.

The presence of feedback connection (interactivity) further amplifies selection of the correct word, as the target lemma and semantic features get further reinforced from activation at the lower level. The amplification stems from the fact that when, for example, lexical nodes receive feedback from the phonological nodes, this will summate with the residual activation that lexical nodes carried on from the previous step, which will be still higher for the target lemmas. Importantly, feedback connections from the phonological levels have the effect of expanding the pool of lemmas that compete for lexical selection, due to activation of lemmas that are phonologically related to the target, but did not receive earlier activation due to lack of semantic relatedness. However, in normal conditions the novel competitors will not pose a significant challenge, as they do not benefit from the summation effects that were just discussed. Lexical selection occurs after a predefined number of steps, providing a further boost of activation to the lemma with the highest level of activity. This will be followed by the selection of the sequence of phonemes that will be uttered.

In sum, in a normal, unimpaired cognitive system the principles of cascading and interactivity tend to boost the activation of target nodes at every level of processing, which might explain why errors are very rare in healthy speakers (Laine & Martin, 2006).

The Relationship Between Errors and Impaired Cognitive Processing

In the interactive 2-step model, transmission of information entails both a signal and noise component (Dell et al; 1986). Indeed, the s- and p-weights encode both these components, reflecting the sum of a constant value (the signal) with one selected at random, at any given time step of processing, from a range of values with mean equal to zero and standard deviation equal to a fraction of the signal (the noise). In normal conditions, the noise portion is a negligible fraction of the signal, and therefore the s- and p-weights reflect predominantly the signal component — that is, a high signal-to-noise ratio (Dell et al., 1986, 1997, 2004; Schwartz et al., 2006). However, by selectively "lesioning" the s- or p-weights, which entails reducing the signal relative to the noise component, it is possible to simulate the effect of lesions to lexical-semantic or phonological processing on naming performance. In practice, lesioning s-weights amounts to reducing the differences in levels of activation between the target lemma (or units associated with the target) and those of its competitors, in a manner that is proportional to the level of severity.

Following this line of thinking, one approach that has been used to assess how well the interactive 2-step model accounts for the range of errors that are frequently observed in people with anomia is to fit the s- and p-weights to individual patients' data collected during naming tasks. In other words, for each patient, the task is to find the combination of s- and p-weight that best reflects the patient's naming performance in terms of both overall accuracy, and proportion of errors of different kinds (Dell et al., 1997, 2004; Nozari et al., 2011; Schwartz et al., 2006). Several studies of this kind have documented not only that the model can provide satisfactory fits to a wide range of patients' data, but also that the above-mentioned approach can provide insights into the cognitive mechanisms underlying specific types of errors.

For example, impaired transmission of information from semantic to lexical nodes can result in several distinct types of errors, depending on how severe the lesion is. For weak impairment/lesion to s-weights, the advantage that the target lemma has with respect to its strongest semantically-related competitors would be reduced, increasing the chances that a competitor is selected, which would lead to a semantic error. In case a target lemma (e.g., <cat>) has competitors that are both semantically- and phonologically-related (e.g., <rat>), lesions to s-weights would likely result in mixed error. This is because throughout the phase wherein lexical processing is underway, the level of activity of this class of competitors will be always similar to that of the target lemmas, as they would receive comparable activation from semantic features and, due to interactivity, further activation from the phonological units. By the same token, if the target lemma has only or mostly competitors that are phonologically-related, the combined effect of lesions to s-weights and intact p-weights is likely to result in phonological errors (e.g. due to the selection of the lemma <mat> instead of <cat>), as during the process of lexical selection only phonologically-related competitors would have comparable levels of activation to that of the target lemma. Importantly, the presence of mixed and phonological errors hinges strongly on the interactivity process: impaired p-weights would result in weak activation of phonological units, and even weaker feedback to lexical units, making such errors less likely to occur.

A moderate lesion to s-weights will further expand the pool of competitors to also include lemmas that are not semantically related to the target (i.e., lemmas of words that are phonologically related to the target's semantic competitors), and that therefore have no relationship, be it semantic or phonological, with the target itself (unrelated errors). Finally, severe lesions selectively affecting s-weights might lead to no responses (omissions), as the signal is so low that noise would make activation across lemmas quite uniform, with none of the lemmas reaching a level of activation required to trigger the process of selection.

The dynamics of interactivity also explain the lexical bias effect. Strong feedback from phonological to lexical units would likely boost the level of activity of lemmas to such an extent that one of them (no matter how distant from the target in semantic/phonological space) will be ultimately selected, phonologically encoded, and uttered, resulting in a lexical (as opposed to nonword) error. Conversely, while severe lesions in p-weights are likely to have little effect on the process of lexical retrieval (due to reduced interaction), they certainly have a disruptive effect on the process of phonological encoding, wherein a word form is converted into the associated sequence of phonemes. Flaws at this stage can lead to addition, deletion, substitution of phonemes which often results in the production of nonwords (i.e., ‘katr’, ‘ka’, or ‘zat’ instead of ‘kat’).

Word Repetition

While originally set to account for the mechanism underpinning naming, the 2-step interactive model also provides a satisfactory account of how word repetition is carried out. However, the architecture that was described in the previous section needs to be upgraded with two elements. First, a distinction between input and output phonology, to explain empirical evidence from patients showing a double dissociation between these two (i.e., performance within normal range in auditory phoneme discrimination tasks, and impaired performance in speech production tasks such as nonword repetition requiring intact phonological encoding, or vice versa; R. C. Martin et al., 1999). Within this model, auditory stimuli would map directly to input phonology first, then to lexical nodes, and then, via the same route employed in word production (but in reverse direction) onto semantic features, to extract word meaning. While such a model could theoretically support word repetition via the input phonology –> lexical nodes –> output phonology route (also called lexical route), it could hardly explain the ease with which healthy subjects can repeat nonwords. In fact, if the lexical level of representation takes part in the processing, there would be strong lexical effects in nonword repetition (i.e., the tendency to convert the nonword into words that sound similar) even in healthy subjects, which is not the case (Nozari et al., 2010).

Therefore, a second, direct route connecting input and output phonology directly is needed — a nonlexical route, which is primarily responsible for the mapping between sounds and the motor commands required to drive the articulatory system when reproducing them.

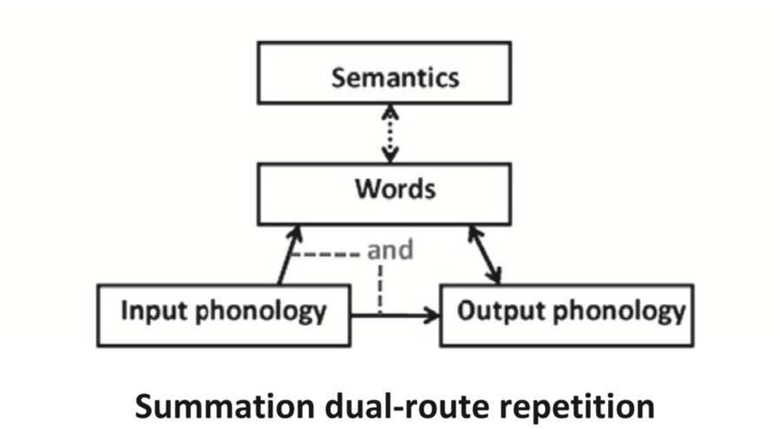

Figure 1. 2 Simplified depiction of the interactive 2-step model integrating a layer to map auditory input into phonology, which connects directly (via the nonlexical route) and indirectly (via the lexical route) to output phonology. Figure taken from Nozari (2010). Reproduced with permission from Elsevier.

While nonword repetition is carried out solely via the nonlexical route, word repetition might rely predominantly on processing occurring in the lexical or nonlexical routes, depending on language and cognitive skills. A study conducted by Nozari and Dell (2013) set out to identify the conditions under which the lexical and nonlexical routes are recruited. The study had three hypotheses: 1) damage to the lexical-phonological route is the key factor determining whether the nonlexical route is engaged for support or not; 2) damage to the nonlexical route precludes its recruitment; 3) failure to access meaning from sound leads to the recruitment of the nonlexical route. To test these hypotheses, the authors first analyzed naming and repetition data taken from a sample of 103 PwA, with the aim to estimate (at the subject-level) the level of impairment of each of the three processes that were hypothesized to play a role in word repetition. To do so, s- and p-weights were estimated by fitting an interactive 2-step model to naming data, while the nl-weights — which encode the strength of the direct connection between input and output phonology — were estimated by fitting data from nonword repetition tasks using a model with only the nonlexical route. Then, two variants of the interactive 2-step model were fitted to the naming and repetition data: a variant implementing only the lexical-route, and a dual-route model implementing both the lexical and nonlexical routes. The authors then divided patients into two groups, based on which of the two models provided a better fit, and compared the s-, p- and nl-weights characteristic of each group to ascertain whether any of these parameters predicted which model fitted best. Results showed that, surprisingly, the level of impairment in the lexical-phonological and nonlexical routes (as measured by p- and nl-weights) did not come up as strong predictors of which model fitted best. Therefore, the first two hypotheses of the study were not corroborated. Instead, the patients for whom the dual-route model provided a better fit had lower s-weights and better phonological working memory, relative to those for whom the lexical-route provided a better fit. In other words, the higher the impairment in accessing word meaning from input phonology, the higher the likelihood that the nonlexical route was recruited, regardless of whether or not this route, or the lexical-phonological process, were themselves damaged. Furthermore, key for the successful recruitment of the nonlexical route was that information stored in phonological working memory could be held for as long as needed for the conversion into phonological output to be carried out.

These results have important implications for therapy planning, as they suggest that a task such as word repetition (which is frequently employed for the treatment of anomia) might not always be useful. Specifically, according to the results presented in this section, word-repetition tasks are expected to yield benefits only when they engage the lexical route — which occurs when access to semantics works — and in so doing improving the p-weights via practice. However, when access to meaning from input phonology is impaired (i.e., low s-weights) other tasks should be preferred, like naming.

Neural Bases of Anomia

The neuroanatomical circuitry underpinning word production and comprehension is deeply embedded into an extensive, bilateral network wherein information is channeled in two, highly specialized, yet interdependent streams: the ventral and dorsal stream. Below is an overview of the areas of the brain that fall within each stream, and about their specific role in language processing.

The Dual-Stream Model of Language Processing

Currently, the most influential neuroanatomical model of speech perception and production is the dual-stream model (Hickok & Poeppel, 2007; Stefaniak et al., 2020), which has been informed by a wealth of converging data from lesion-mapping studies (Kümmerer et al., 2013; Mirman et al., 2015), tractography (Saur et al., 2010), computational modeling (Chang & Lambon Ralph, 2020; Ueno et al., 2011), task-based activation mapping (Saur et al., 2008). As the name suggests, this model breaks down the circuitry underpinning language processing in two main streams: the ventral stream, and the dorsal stream.

The ventral stream is primarily responsible for extracting conceptual and semantic information from the auditory input stream. As such, it is heavily responsible for auditory comprehension. As shown in Figure 1.3, starting from primary auditory areas (dorsal superior temporal gyrus, (STG)), where basic spectrotemporal features of the signal are extracted, the ventral stream runs inferiorly to adjacent areas located in the mid-posterior superior temporal sulcus (pSTS), which map the aforementioned features to associated abstract phonological representations, then to the posterior parts of the middle temporal gyrus (pMTG) and inferior temporal sulcus (pITS), to attain lexical access by mapping phonological information onto word forms (lemmas).

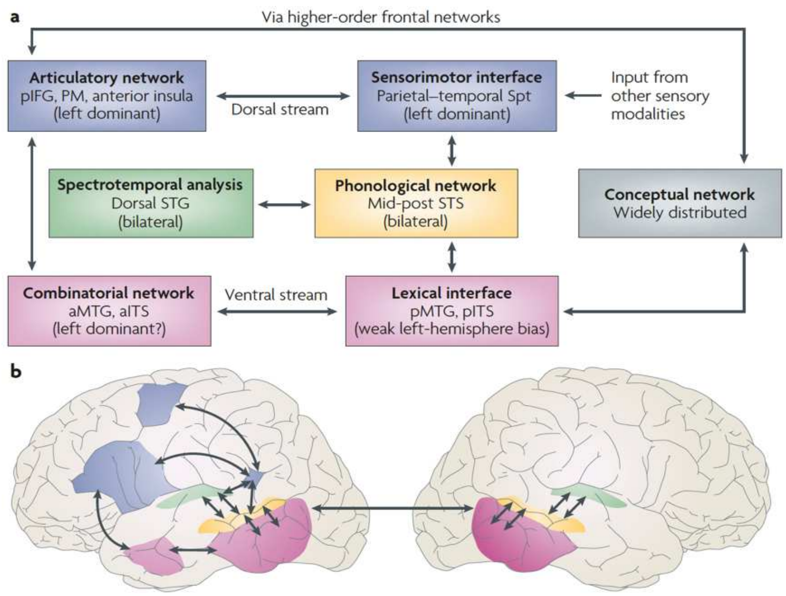

Figure 1. 3 - Depiction of the dual-stream model of language processing. Figure taken from Hickok (2007). Reproduced with permission from Springer Nature.

From here, information processing runs towards the anterior parts of these regions (aMTG and aITS) via the middle longitudinal fasciculus (Ueno et al., 2011) to map lexical entities onto the associated meaning, achieving speech comprehension. The ventral stream extends over a set of bilateral areas which, with the exception for the absence of the right aITS and aMTG, are homologues of those present in the left hemisphere. The ventral stream connects directly with frontal areas like the posterior inferior frontal gyrus (pIFG) via a temporofrontal fasciculus (extreme capsule) and via the uncinate fasciculus (Hickok & Poeppel, 2007), and indirectly via the dorsal stream.

The dorsal stream is strongly left lateralized, and is responsible for the mapping between sounds and the motor commands required to drive the articulatory system when reproducing them. Processing in the dorsal stream starts in primary auditory areas (dorsal STG), from which the input goes directly to the area Spt (located in the Sylvian fissure in the parietotemporal boundary), or indirectly via the mid-post STS involved in phonological processing. In this model, the Spt plays the key role of intermediary hub involved in sensorimotor integration, namely the computational task of mapping a given sensory target onto the motor program required to produce it. This information is then passed on to a cluster of frontal areas, such as posterior inferior frontal gyrus (pIFG), premotor cortex (PMC), and anterior insula, which encode information about the motor plans required to produce specific syllables (Hickok, 2012; Peeva et al., 2010). These areas are connected to the Spt via the arcuate fasciculus (Geschwind, 1970; Ueno et al., 2011).

A Higher Resolution View of the Networks Supporting Normal and Impaired Naming

The dual-stream model hints to the fact that the cognitive processes required for successful word production rely on the interplay between areas roughly embedded in the ventral (mapping of semantic features onto lexical items) and dorsal streams (encoding of phonological forms and planning of the articulatory commands).

However, lesion-symptom mapping studies — which inform on the role of a cortical area on a specific behavior by analyzing the statistical relationship between the integrity of such area and the severity of the deficits of the behavior under investigation — suggest that a higher resolution view is required to explain the fact that patients with similar language deficits (e.g., word-finding difficulties with predominant phonological errors) might present focal lesions in different, distant areas that might not all be confined within the boundaries of a specific stream (DeLeon et al., 2007; Yourganov et al., 2015). Currently, these findings are explained by considering networks of cortical areas, as opposed to single areas, as the fundamental unit of analysis when it comes to mapping the neural bases of a cognitive process. According to this view, specific cognitive processes emerge from the coordinated activity within a cluster of brain regions that are tightly connected, wherein some might play a more important role than others. Importantly, this view places an emphasis not only on the integrity of a cortical/subcortical region, but also on the integrity of white matter tracts that connect regions subserving the same cognitive function.

In order to shed light on the networks responsible for various aspects of word production, Fridriksson and colleagues (2018) combined the traditional lesion-symptom mapping approach with an approach that associates lesions to white matter tracts with language deficits, called connectome lesion-symptom mapping. In this study, several different dimensions of language were assessed, from fluency to sentence comprehension, reading and writing, semantic processing and speech rate. Of relevance to this section, naming was also included among the behaviors that were tested. By and large, results showed that semantic errors were associated with impairment to areas spanning the entire ventral stream, including both those involved in semantic processing (i.e., angular gyrus; AG), and in lexical access (pMTG). Furthermore, impaired connections within the temporal gyrus — a hub for lexical access (Hickok & Poeppel, 2007; Indefrey, 2011) — were all strongly associated with the frequency of semantic errors.

Phonological errors were associated with lesions in areas located in the ventral and dorsal streams: within the former stream are white matter connections between areas involved in semantic processing (AG) and lexical access (pMTG), within areas engaged in lexical access (pMTG and ITG), and lesions to areas engaged in phonological encoding (pSTG; Hickok & Poeppel, 2007; Indefrey, 2011); within the latter stream are connections between areas involved in lexical access (ITG) and areas responsible for articulatory planning (insula). These results are in line with the mechanics of the interactive 2-step model (Schwartz et al., 2006), which can generate phonological errors by either disrupting semantic and lexical processing while keeping phonological encoding intact (e.g., mild lesion to s-weights and intact p-weigths), or by solely disrupting phonological encoding (e.g., intact s-weights and severely lesioned p-weights).

Repetition errors were associated with a wide network including areas in both ventral (primarily areas involved in lexical access (STG, MTG, ITG) and associated connections) and dorsal streams (areas involved in articulatory planning, as the posterior insula); and with connections between hubs in the two streams being strongly associated with performance in repetition tasks.

In short, a relatively simple task such as naming relies on the coordinated activity of a wide network of areas, and insult to any of them might result in temporary or permanent behavioral deficits, depending on the size of the injury and on the behavioral interventions that are adopted to revert its effects. The next section discusses the sequelae of neurobiological events underpinning spontaneous recovery, behavioral interventions informed by cognitive models of anomia presented in previous sections, and the key principles of neuroplasticity that such interventions can exploit to maximize the extent to which people can recover from aphasia.

Recovery from Aphasia

Neurobiological Bases of Recovery from Stroke

Decades of research on animals and humans have elucidated several key aspects of the dynamics underpinning the brain response to stroke. A solid finding stemming from this research program is that the brain response to stroke takes several months before it stabilizes, and it unfolds in distinct phases — each characterized by unique molecular and cellular processes and by specific patterns of large-scale neural activity and behavioral outcomes in response to a task (Cramer, 2008; Grefkes & Fink, 2020). The next sections provide a brief account of the processes characterizing these phases.

Recovery in The Acute Phase

Stroke occurs when a blood vessel is ruptured (hemorrhage) or obstructed (ischemia). This event triggers a drastic reduction of blood flow in the neural tissue that is directly subserved by the affected vessel. In this area, called infarct core, the loss of blood supply triggers immediate excitotoxic and inflammatory processes, which will lead to complete neuronal death within hours from the onset of stroke (Krakauer & Carmichael, 2022). Moreover, the inflammatory response will not only radiate to the intact, adjacent tissue (peri-infarct or perilesional area), but its effects will also be relayed to areas of the brain that are distant, yet structurally or functionally connected to the infarct core (Carmichael, 2016). This is due to the fact that, as the axons of cells within the infarct core die, neurons in distant areas connecting to neurons in the infarct core lose a source of input, which leads to the loss of dendritic spines and to cellular dysfunction (diaschisis) or death (Carrera & Tononi, 2014; Kiran et al., 2019; Krakauer & Carmichael, 2022).

Functional neuroimaging studies conducted in PwA have provided results that are consistent with the molecular and cellular dynamics characterized in animal studies. Indeed, during the acute phase (1-week after the onset of stroke), there is very limited activation throughout the bilateral network that is known to be involved in language processing in healthy subject (Saur et al., 2006; Stefaniak et al., 2020; Stockert et al., 2016; Wilson & Schneck, 2020), reflecting the hypometabolism induced by cellular death, the predominant inhibitory effect due to the abnormal levels of extracellular GABA, and diaschisis (Carmichael, 2016; Krakauer & Carmichael, 2022).

The acute phase is when the most severe behavioral symptoms are observed, with severity of aphasia measured at this point being strongly correlated with the size of the lesion (Laska et al., 2001; Pedersen et al., 1995, 2003; Saur et al., 2006). In the case of ischemic stroke, if appropriate medical procedures (i.e. administration of anticoagulants; Canadian Stroke Best Practices, 2024) are undertaken quickly enough, complete recovery of function can be observed immediately.

Recovery in The Subacute Phase

Starting approximately one week after the onset of stroke, the very same inflammatory signals that had initially caused neuronal dysfunction and death play a pivotal role in kickstarting a long, complex process that leads to tissue repair, functional reorganization of large-scale neural networks, and ultimately to total or partial recovery of function. This process is called spontaneous recovery.

Some of the molecular agents that were called into action as part of the initial inflammatory response, like free radicals (Le Belle et al., 2011) and inflammatory cytokines (Gleichman & Carmichael, 2014), activate intracellular cascades that induce angiogenesis, the production of neural growth factors that is conducive to neurogenesis, guide the migration of pluripotent stem cells from the region where neurogenesis takes place (i.e. subventricular zone (SVZ), (Krakauer & Carmichael, 2022)) to the perilesional areas, and stimulate and guide the process of axonal sprouting towards functionally-linked adjacent and distal areas (Krakauer & Carmichael, 2022; Li et al., 2015). Furthermore, following an initial decrease in dendritic spine density, neurons in perilesional areas exhibit recovery to pre-stroke levels, or even an increase relative to such levels (Joy & Carmichael, 2021; Krakauer & Carmichael, 2022; Mostany et al., 2010).

Recruitment of new cells in perilesional and distal areas, re-establishment of old connections and/or the formation of new ones, and increased dendritic spine density contribute to the increase in cortical excitability observed throughout the language network. Functional neuroimaging studies adopting a longitudinal design have revealed that, during the subacute phase, there is a strong increase in activation of perilesional and contralesional areas of the language network, relative to the acute phase (Saur et al., 2006; Stockert et al., 2016; Wilson & Schneck, 2020). However, the balance of activation between the two hemispheres is rarely as observed in healthy subjects, and varies considerably across patients, with lesion size being the strongest predictor of the observed pattern. Generally, as the size of the lesion affecting left hemisphere language areas increases, the pattern of activation observed in aphasic stroke patients diverges more and more from that of healthy subjects, progressively moving from a strongly left-lateralized network to one where activation levels are comparable, or higher in the right hemisphere (Kiran et al., 2019; Kiran & Thompson, 2019; Stefaniak et al., 2020; Wilson & Schneck, 2020).

In most cases, a drastic amelioration in the symptoms of aphasia ensues over a period of six months after the onset of stroke, with initial severity of impairment consistently emerging as the strongest predictor of final outcome — intuitively, the lower the initial severity score, the lower the severity score measured at later points in time (Laska et al., 2001; Lazar et al., 2010; Pedersen et al., 1995, 2003). Studies following patients with aphasia from stroke onset until they entered the chronic phase showed that, while it is common for people starting with mild symptoms to attain full recovery during the subacute phase (Pedersen et al., 1995, 2003), a significant portion of those initially presenting with severe symptoms do not attain complete recovery of function, with anomic symptoms being one of the most common complaints for people living with chronic aphasia (Pedersen et al., 2003).

In addition to the beneficial effects of spontaneous recovery, studies have shown that engaging in therapy as early as a person can tolerate during the subacute phase yields benefits that go over and above those that are brought about by spontaneous recovery. While intuitive, this has proven very challenging to back up empirically. In fact, demonstrating the effectiveness of early treatment does not entail solely a between-group comparison between the effects of therapy versus no-therapy, but also that the groups are adequately matched for variables that are known to affect the potential for recovery, such as lesion size and location, initial severity, and age (Fridriksson & Hillis, 2021; Kiran & Thompson, 2019; Kristinsson, 2021; Roberts et al., 2022). To the best of my knowledge, only one study, conducted by Roberts and colleagues, has succeeded in controlling for these factors, demonstrating that receiving early therapy had favorable effects on a number of language tasks, over and above those that were attributable to spontaneous recovery, and that these effects could be measured up to one year after the onset of stroke (Roberts et al., 2022).

However, it is currently not clear whether, in light of the heightened state of plasticity that is observed during the subacute phase, all controlled therapy delivered during this period is more effective than when delivered during the chronic phase (when the neural mechanisms underpinning spontaneous recovery have subsided). A meta-analysis of 55 studies has attempted to address this question by pooling together results from studies contrasting treatment during the subacute phase with no treatment (hence contrasting therapy with the effects of spontaneous recovery) and from studies contrasting the effects of treatment versus no-treatment in the chronic phase (Robey, 1998). While results showed that effect sizes associated with early therapy were twice as large as those attributable solely to spontaneous recovery and to treatment delivered in the chronic phase, several of the biographical and neuroanatomical variables that affect recovery mentioned above were not controlled for, and therefore it is not possible to confidently conclude that this is the case (Fridriksson & Hillis, 2021; Roberts et al., 2022).

Recovery in The Chronic Phase

The stroke-induced molecular and cellular mechanisms that drive neural tissue repair, structural and functional reorganization, and recovery of function throughout the acute and subacute phases taper off starting approximately six months after the onset of stroke, and along with them the potential for spontaneous recovery that ensued as a result of this period of heightened brain plasticity. From this point, absent engagement with targeted speech-language therapy or with regular, structured conversations in social contexts (such as that offered in aphasia centers (Holland et al., 2017)), behavioral symptoms remain largely stable throughout the chronic phase of aphasia (Davis, 2013). However, a wealth of empirical evidence shows that, even during the chronic phase — and in general throughout the rest of one's life — the brain retains the ability to respond to specific kinds of experiences in ways that drive desirable and lasting changes in behaviors (Berthier & Pulvermüller, 2011; Kiran & Thompson, 2019; A. M. Raymer et al., 2008; Stefaniak et al., 2020). In other words, brain plasticity can be harnessed throughout the lifespan with adequate behavioral interventions. Key to exploit brain plasticity to drive recovery of function is to identify the key ingredients that are conducive to it. In this regard, few solid principles can be distilled from the literature on experience-induced stroke recovery in both animals and humans (Kiran & Thompson, 2019; Kleim, 2011; Kleim & Jones, 2008; A. M. Raymer et al., 2008).

Key Principles of Neuroplasticity to Inform Aphasia Therapy

Learning (or relearning) hinges on practicing the target behavior. With time, absent periodic rehearsals, the memories underpinning knowledge of facts or of how to perform a given task decay. In the context of rehabilitation, practice is thought to yield its restoring effects by selectively modulating neural activity in the networks subserving the target behavior (Kiran & Thompson, 2019; Krakauer & Carmichael, 2022). This, in turn, promotes the strengthening of existing connections and the formation of new ones in order to encode the task-related information underpinning performance improvements. This principle is currently a central tenet of the impairment-based approach to aphasia therapy (Chapey, 2008a; Nickels, 2002).

Once the right circuit/behavior is targeted, high levels of dosage (total amount) — both in terms of intensity (amount of practice per training session) and frequency (number of sessions per week) — of practice are paramount to ensure not only that performance improvements take place, but also that they are encoded in neural circuits in a way that promotes long-term maintenance of the gains (Kiran & Thompson, 2019; Kleim & Jones, 2008; Krakauer & Carmichael, 2022; A. M. Raymer et al., 2008). In fact, animal studies show that inducing changes in neural structure and function (i.e., increased number of synapses and motor map reorganization) that underlie learning require repeated exposures to the target behavior, with every act of repetition reinforcing the relevant memory (see Krakauer & Charmichael (2022) for a review).

Another factor that promotes the acquisition and maintenance of gains is the salience of the experience that people have. Whether salience manifests as the expectation of a highly valued extrinsic (i.e., relevance to one's own life goals) or intrinsic (i.e., fun, curiosity) reward, it acts by boosting motivation and helping with directing attention towards the task at hand (Biel et al., 2022; Biel & Haley, 2023; Kiran & Thompson, 2019; Ryan & Deci, 2017), both of which in turn have a positive influence on the depth of acquisition and consolidation of information.

Behavioral Approaches to the Treatment of Anomia

Treatment of naming and word-finding disorders is typically guided by an impairment-based approach, which rests firmly on the principles described in the previous section: adoption of tasks that engage the impaired cognitive process will stimulate the associated neural circuits and, by means of high dosage (repetition) and intensity, ensure that any changes in activity underpinning behavioral improvements are encoded in such a way to promote long-lasting maintenance. This will ultimately reduce impairment and, by virtue of this, restore the functioning of a given cognitive process (Brady et al., 2022; Chapey, 2008a; Kiran & Thompson, 2019; A. M. Raymer et al., 2008).

Given the long-standing consensus on the centrality of semantic and phonological processing in word and speech production, the most prominent impairment-based approaches for the remediation of anomia target these two domains of cognition (Nickels, 2002; Wisenburn & Mahoney, 2009).

Semantic-Based Approaches

As mentioned in the section describing the interactive 2-step model, the spreading activation dynamics lead to competition for lexical selection among items that share semantic features, with the strength of competition being proportional to semantic similarity. When the target item has a very close semantic competitor (i.e. horse, donkey; eagle, osprey) successful retrieval relies on optimal connection between the semantic feature that distinguishes the two items and the target lexical unit. In this situation, even small impairments in the connections between semantic and lexical nodes, which result in proportional increases in the level of noise in the system, might level out the contribution to target lexical activation due to the distinguishing feature, leading to the erroneous selection of a semantic competitor. As the severity of impairment increases, so does the range of items that compete for lexical selection expand towards items that are less and less semantically similar to the target. To counter this situation, semantic-based approaches are designed to strengthen semantic representations and their connection to lexical ones (Efstratiadou et al., 2018).

Several tasks fall within the category of semantic-based approaches. One of the most common is picture-word matching, wherein clients are tasked with selecting, among a pool of candidates, the picture that is associated with a given word. In the presence of pictures that share ever increasingly similar semantic features, this task challenges people to identify ever subtler features that differentiate the target picture from its competitors, and in so doing inducing a strengthening of these representations. Engaging in this process of categorization at ever increasing resolution is meant to progressively shrink the range of viable competitors to the target, thus increasing the chances of retrieving the correct items. The odd-one out task —

identifying the common category among a number of items, be them words or pictures, and selecting the one that does not fit in that category — has a similar logic.

Semantic feature analysis (SFA) is another commonly used task. Here, typically clients are asked to either provide a description of specific features of the target item (Boyle, 2010), or to select the right description among a set of competing ones (Kiran & Thompson, 2003). For example, for every picture to be named, clients might be asked to provide a specific description (i.e. for 'cat', which group it belongs to, what it is commonly associated with, what it looks like, where it can be found, what actions it can perform). This exercise is meant to exploit the principle of spreading activation within the semantic network, whereby inducing co-activation of semantic features associated with the target item, which in turn facilitates target retrieval (Boyle, 2010; Efstratiadou et al., 2018; Kiran & Thompson, 2003).

Phonologically-Based Approaches

Phonologically-based approaches are conceptually similar to semantically-based ones, but aiming at restoring the connections between lexical and phonological representations, or strengthening phonological representations per se (Madden et al., 2017). Examples of tasks falling in this category are: picture naming with the support of phonological (or orthographic) cues such as first phoneme, first syllable, a word that rhymes with target, and auditory/written input of target word to repeat/read (Abel et al., 2005; Nickels, 2002; Wambaugh, 2003; Wisenburn & Mahoney, 2009); picture naming along with analysis of the phonological makeup of the target word (phonological component analysis, PCA; Leonard et al., 2008), which entails a prompt to spontaneously break down a word form into its first sound, first syllable, a rhyme, final sound, number of syllables.

Mechanisms of Action of Semantic and Phonologically-Based Approaches

Both semantically- and phonologically-based approaches employ cueing strategies to facilitate semantic, lexical and phonological processing. The facilitatory effect of cues is attributed to the fact that, by directly raising the level of activation of a given semantic/lexical/phonological unit associated with the target item, they boost activation of the target word relative to that of its competitors (priming; Nardo et al., 2017; Nickels, 2002). For example, when presented with the picture of a cat along with the auditory cue /k/, the active phoneme activates (via feedback connections) all lemmas starting with it, hence boosting <cat> among its semantic competitors.

Semantically- and phonologically-based approaches have proven successful in driving improvements in naming and word-finding abilities, in some cases leading to long-term maintenance of gains (Nickels, 2002; Sze et al., 2020; van Hees et al., 2013; Wisenburn & Mahoney, 2009). However, while originally conceived to restore the processing capability at a given level of processing (Nickels, 2002), the literature is laden with conflicting results (Kristinsson et al., 2021; van Hees et al., 2013). Some studies contrasting the two approaches reported results in line with the theoretical rationale (i.e. people with predominantly semantic deficits respond better to semantic approach, and vice versa; Wambaugh, 2003), while others found the opposite pattern (Nickels & Best, 1996; A. Raymer et al., 1993; van Hees et al., 2013), or found no clear relationship between the locus of breakdown and effectiveness of a given approach (i.e. improvements observed after having received both treatments, irrespective of locus of breakdown; Doesborgh et al., 2004; D. Howard et al., 1985; Lorenz & Ziegler, 2009; Nickels, 2002). These conflicting findings led some authors to question whether recovery of naming skills is best attained by drilling directly on the impaired process, as opposed to exploiting residual abilities (van Hees et al., 2013).

This apparent paradox could be resolved by considering three important factors about the cognitive mechanisms engaged during word production, and about the nature of the semantically- and phonologically-based tasks. First of all, in the context of the interactive 2-step model, phonological processing can affect and facilitate lexical processing, even (and most obviously) in the presence of semantic impairments. In this case, when presented with a picture to name, spreading activation from relevant semantic units reaches associated lexical items. As mentioned in the previous section, the presence of impairment results in high levels of competition for selection, drastically increasing the chance of selecting the wrong item. In this context, exposure to phonological cues would prime, via feedback connections, those among the competing lexical items that embed the cue itself (most likely only the target), facilitating correct retrieval. In time, with practice, successful attempts at retrieving words directly from semantics, with the key support of phonological cues, would lead to the reinforcement of the (item-specific) connections between semantic features and lexical items (s-weights). Second, and strictly related to the latter point, this could work only if the supporting mechanism (lexical-to-phonological mapping, and feedback from phonological to lexical units, in this case) is spared – and this explains why drilling on residual skills might be more effective than targeting the impaired ones. Third, as it is obvious from the example made above, while the semantically- and phonologically-based tasks discussed above place an emphasis on semantic and phonological processing, respectively, in practice they engage the entire chain of processes, albeit to different extents (Nickels, 2002).

Importantly, the fact that in some cases people with predominantly semantic/phonological deficits benefit more from phonological/semantic therapy, respectively, does not mean that targeting impaired processes is doomed to fail. Rather, it might be simply the case that for this approach to yield benefit it takes much more time than when targeting residual skills.

Neuroimaging studies lend support to the idea that exploiting residual capabilities might be a faster route to recovery. One study designed to map the areas engaged by semantic and phonological therapy (van Hees et al., 2014) showed that damage to the left caudate - an area engaged in selecting among semantic candidates (Copland, 2003; Crosson et al., 2003) – limited the effectiveness of semantic-based treatment. Similarly, the effectiveness of PCA was associated with the integrity of the left supramarginal gyrus (SMG), which has been shown to be involved in phonological processing (van Hees et al., 2014).

While the debate over the optimal treatment for anomia given a person's language profile (Kristinsson, 2021; Kristinsson et al., 2021) is still open, a number of studies suggest that integrating both semantically- and phonologically-based approaches is often beneficial, especially considering that most PwA do manifest deficits in both semantic and phonological processing (Madden et al., 2017; Nickels, 2002; Wisenburn & Mahoney, 2009).

The Role of Effort and Success in Word Learning

A task that is commonly used to restore the mapping between conceptual and phonological levels of representation is confrontation naming, wherein PwA are presented with a picture of objects and asked to name the item in the picture. The literature on aphasia rehabilitation presents three main strategies to implement confrontational naming. These strategies, referred to as errorful learning, errorless learning, and retrieval practice, are informed by theories of rehabilitation that see *effort* and *success* as key ingredients for optimal recovery (Nunn et al., 2023).

Errorful learning places an emphasis on effort, under the assumption that experiencing the right level of challenge when attempting to retrieve a word drives engagement and attentional mechanisms that in turn boost memory encoding (Nunn et al., 2023; Nunn & Vallila-Rohter, 2023). The challenge derives from the fact that PwA are prompted to retrieve the correct word from long-term memory, typically without support, except for feedback on the accuracy of vocal production to reinforce the correct association between stimulus and vocal output or to signal the need for correction. In case the trial and error procedure turns out to take too long to lead to the correct vocal output, semantic and/or phonological/orthographic cues are made available to support word retrieval naming. This strategy is referred to as an errorful approach because in the pursuit of retrieving the correct word PwA tend to make mistakes before getting the word right. Retrieval practice largely overlaps with the approach discussed above, placing a strong emphasis on the beneficial effect of effort on learning, but in addition it allows subjects to be first exposed to the correct mapping (Middleton, Schwartz, et al., 2015; Nunn et al., 2023).

Errorless learning puts a premium on minimizing the occurrence of errors (J. K. Fillingham et al., 2003). This is achieved by pairing a picture with its written and/or spoken form — thus enabling a person to produce the correct word by simply reading or repeating it, without the need for feedback or cues. Errorless learning is believed to exert its positive effects via Hebbian learning mechanisms — whereby the strength of a memory that associates two stimuli (i.e. picture and word) is directly proportional to the extent to which the stimuli are temporally correlated (Baddeley & Wilson, 1994; J. Fillingham et al., 2005; J. K. Fillingham et al., 2003, 2006; Nunn et al., 2023; Ralph & Fillingham, 2007). Therefore, the more a person is capable of naming a picture correctly, the stronger the memory encoding will get. Similarly, making mistakes would reinforce the wrong memory, absent the ability to learn from feedback. This approach was originally applied with success to patients with amnesia on the premise that, given their marked deficits in declarative memory —which limit the ability to learn from feedback on naming accuracy — it would prove to be more effective than approaches relying on trial-and-error guided by feedback (Baddeley & Wilson, 1994). The success of errorless learning on people with amnesia, along with the fact that some PwA often present similar cognitive impairments, encouraged testing errorless learning on PwA (J. Fillingham et al., 2005; J. K. Fillingham et al., 2006).

A line of studies have directly tested the efficacy of these methods on PwA, showing that while errorless learning has consistently proven to yield similar or higher gains in naming abilities when people were tested shortly after the training session (up to 1 day after the end of training), methods that couple encouraging participants to retrieve words from memory with the provision of corrective feedback have proven more effective at consolidating gains over longer terms (as measured one week and one month after the end of training) (Middleton, Schwartz, et al., 2015; Schuchard & Middleton, 2018b, 2018a).

Currently, the best explanation for this pattern of results indicates two key differences between approaches prioritizing success versus effort. The first difference is that, while errorless learning fundamentally entails word repetition, which relies on retrieving words from short-term memory, errorful learning forces subjects to retrieve words from long-term memory. This requires more effort, which in turn induces higher levels of engagement and depth of processing. These same mechanisms are thought to underpin the advantage for distributed over massed practice, especially when comparing long-term gains. In fact, massed practice entails intensive rehearsal of the same item over very short periods of time (i.e. several times within the same session), and therefore it is likely to rely on retrieval from short-term memory. Instead, distributed practice entails spreading bouts of practice over longer periods of time (i.e. days), forcing more effortful retrieval from long-term memory. A line of studies contrasting these two approaches in PwA have consistently shown that, while massing practice leads to faster acquisition of behavior and higher immediate gains in naming performance, distributing practice over longer time intervals yields advantages in maintenance (Middleton et al., 2016, 2019, 2020; Schuchard et al., 2020).

The second key difference between errorful and errorless methods is that retrieving from long-term memory requires mapping from conceptual/semantics to lexical forms, then from lexical to phonological forms, while word repetition engages only the second step (Nozari & Dell, 2013). Therefore, while the former strategy is thought to reinforce both steps, the latter is thought to reinforce only the second. This hypothesis is supported by empirical studies showing that PwA showing predominantly semantic deficits benefited more from engaging with tasks that emphasize effort rather than success, while those with predominantly phonological deficits benefitted from both tasks to a similar extent (Schuchard & Middleton, 2018a, 2018b).

The Quest for Personalized Treatment and the Role of Recent Methodological Innovations

Aphasia is a complex disorder, and as such it requires complex interventions that have to account for several parameters in order to yield optimal results: in short, it requires a personalized approach. Therapy outcomes are likely the result of the interplay between several factors, such as biographical (age, sex, handedness, SES), neuropsychological (severity of aphasia, language and cognitive deficits), anatomical (lesion location and size), and therapy parameters (type of therapy, dosage, frequency, intensity, timing; Brady et al., 2022; Kristinsson, 2021).

Up until recently, the gold standard for generating evidence on the role of a specific factor on therapy outcomes was to contrast group-level summary statistics related to different levels of the factor under investigation (i.e., SLT vs no-SLT; early SLT vs late SLT; intensive SLT vs not-intensive SLT, etc.) while averaging across all the other factors (Brady et al., 2022; The RELEASE Collaborators et al., 2022). There are several key limitations to this approach. First, characterizing the role of each factor individually is practically infeasible, as measuring their effect while controlling for several possible confounding variables requires the contrast of groups of participants that are homogenous in all but the variable of interest, which in turn requires sample sizes that no randomized-controlled trial (RCT) so far has been able to achieve. Otherwise, it is likely that, even using randomized allocation, there will be confounding effects (Brady et al., 2022; Debray et al., 2018; The RELEASE Collaborators et al., 2022). Second, it only allows for investigation of the main effect of a variable, and precludes the investigation of interactions among factors. Third, traditional pairwise meta-analysis of RCTs take as input the mean standardized difference between the effect of two treatments, allowing one to make conclusions only about the direct contrast between treatment or factor levels — for example, if a study investigates the effect of therapy A vs B, and one of A vs C, it is not possible to infer the effect of B vs C. Given these limitations, it comes as no surprise that the best available meta-analysis of all RCTs conducted to assess the effect of specific therapy parameters on language outcomes has failed to provide deep insights, except for the non-controversial finding that engaging with therapy leads to significant improvements relative to no therapy (Brady et al., 2016).

The failure of such methods to illuminate the dynamics of the interplay between relevant factors has pushed the research on complex disorders to adopt novel approaches to data analysis. One such approach is network meta-analysis (Brady et al., 2022; Debray et al., 2018). This approach solves (or alleviates) the problems described above by gathering information about individual participants' raw data (i.e., biographical, neuropsychological, therapy parameters) from multiple sources, including datasets from RCTs, case-series studies, observational studies etc., and exploiting individual participants' data, at their full resolution, to treat all factors as covariates whose effect on treatment outcome can be estimated. This approach allows contrasting the effects of treatments that were not directly compared in a single study.

Employing NMA, Brady et al. (2022) were capable of putting together data from 959 participants, from 25 RCTs. The results of this analysis largely confirmed past conclusions regarding the benefits of engaging with therapy, but importantly provided detailed insights into the roles of timing of therapy, age, aphasia severity, and therapy parameters.

For example, the effect of therapy intensity on gains in functional communication was modulated by age, with people younger than 55 years benefiting from higher levels (4 h/w) than those older than 65, whose gains peaked at 2 h/w. In contrast, gains in auditory comprehension were associated with more than 9h/w of training for both groups, which also shared the same optimal levels of frequency (4-5 d/w) and dosage (more than 50 h). Interestingly, people engaging with therapy during the first three months after onset of stroke showed optimal gains for low intensity (2 h/w) and moderate dosage (20-50 h), while those engaging with therapy later required higher intensity (4 h/w) and dosage (more than 50 h). The effects of age and time post-stroke on the optimal level of intensity is, possibly, due to the fact that elderly people, and those dealing with the immediate aftermath of stroke, cannot tolerate higher levels of intensity due to proneness to fatigue. Highest gains were also associated with total duration of therapy being more than 20 weeks.

Despite the unprecedented amount of insights, some facets could not be explored adequately due to lack of data. For example, there were not enough data to make conclusions about the ideal set of parameters that lead to optimal recovery of naming skills, nor to quantify the effect of spontaneous recovery, as there were not enough individual datasets reporting changes in language performance following periods where therapy was not administered during the first months from stroke onset. In the future, additional — and potentially high volumes — of data contributed by CBATs could prove important to illuminate these issues.

The Continuum of Care for People with Aphasia

Aphasia is a chronic condition that requires long-term management (Chapey et al., 2000; Nichol et al., 2019; Simmons-Mackie, 2017). Broadly speaking, the continuum of care that is currently in place to support the management of this condition unfolds in three major phases: acute hospitalization; inpatient/outpatient rehabilitation; long-term care. While all phases share some objectives (i.e. delivering speech therapy, educating PwA and their families on how to best manage aphasia), they have different priorities. Unfortunately, as currently implemented, the service delivery model underpinning the continuum of care has severe limitations when it comes to addressing the needs of PwA and their families.

Acute Hospitalization

Upon the onset of stroke or traumatic brain injury leading to aphasia, PwA are hospitalized for about 4-7 days (Winstein et al., 2016). During this period, the priority is to deliver treatment that can limit or reverse the effects of stroke, and monitoring the situation to ensure that it stabilizes. If deemed possible it is also recommended that speech-language therapy is delivered, and that patients, along with their family members, are engaged in the process of planning for the management of aphasia post-discharge (Winstein et al., 2016).

While acute hospitalization has been proven very successful in achieving its declared priority, as demonstrated by the ever-declining rates of death from stroke (Mozzafarian, 2016), much work is needed to improve on the other two objectives (Simmons-Mackie, 2017). In fact, therapists operating in the acute care units spend about half (52.8%) of their time delivering speech therapy, and only 10% of this (~5% of their total time) on patients receiving more than 3 hours per week of therapy (Code & Heron, 2003). Also, as discussed in detail in different studies (Avent et al., 2005; Foster et al., 2016; Simmons-Mackie et al., 2007), PwA and their family members are not usually adequately informed about aphasia, its consequences, ways to promote recovery, and strategies that can be employed to facilitate communication. This issue amplifies various problems. For one, people that do not know how to support communication for a person with aphasia tend to not engage them in conversation, even when it comes to discussing issues that are of personal relevance. Indeed, plans for discharge from the hospital, for starting therapy, and the statement of therapy's goals are often discussed by therapists and caregivers, bypassing PwA and their carers (Godecke et al., 2014). Given that topics such as the nature of aphasia and its consequences, dynamics of recovery, and strategies to facilitate communication are commonly at top of the list of things that they want to know about before being discharged from acute care (Ganzfried & Hinckley, 2012; Howe et al., 2012; Wallace et al., 2017; Worrall et al., 2011), it comes as no surprise that both PwA (Simmons-Mackie et al., 2007) and their carers (Simmons-Mackie, 2017) complain about this state of affairs.

Inpatient and Outpatient Rehabilitation

Following discharge from the acute hospitalization, patients might be referred to an inpatient rehabilitation facility and, if not, they can be referred to outpatient services. Briefly, the former kind of institution is intended to deliver "*hospital-level care to patients needing intensive, interdisciplinary rehabilitation programs to upgrade their ability to function*" (Winstein et al., 2016), while the latter typically consists of home health care or community aphasia programs that focus on facilitating the transition from the acute hospitalization phase to life at home, while providing therapy that can help patients maximizing recovery of function.

Inpatient and outpatient rehabilitation share similar limitations. For example, within inpatient rehabilitation facilities, patients are expected to receive up to 3 hours of therapy per day, for 5 days per week, for about 2 weeks in the US (Winstein et al., 2016) and 5-6 weeks in Canada (Canadian Stroke Network, 2011). However, on average PwA receive 13 and 24 minutes of therapy per day when enrolled in inpatient and outpatient rehabilitation facilities, respectively (Foley et al., 2012; Simmons-Mackie, 2017), a dosage that is far below the 3 hours per day that is recommended by best practices (Hebert et al., 2016), and that is not nearly enough to attain optimal recovery (Brady et al., 2022). Contrary to best practices is also the inadequate importance given to understanding clients' goals and interests, and on targeting participation-level outcomes by engaging with real-life scenarios that are relevant to them (Simmons-Mackie, 2017).

Another problem that has been documented is that rates of patient referral (i.e. patients who can get a certain number of therapy sessions reimbursed by their private insurance company, or by the public healthcare system) to both inpatient and outpatient facilities are lower than expected if current recommendations were followed. Studies suggest that, in Canada, only about 37% of stroke patients with a moderate-severe diagnosis are referred to a rehabilitation facility (Canadian Stroke Network, 2011). For those who do not benefit from subsidies, or that have exhausted their benefits, there are several barriers to accessing adequate levels of therapy, due to high costs, insufficient number of therapists, and difficulties with transportation, which are particularly accentuated in rural areas (Arcury et al., 2005; Nichol et al., 2019; Nichol, Rodriguez, et al., 2023).

Long-Term Care

One of the most commonly declared objectives of PwA is to find a way that allows them to be active members of their family, community, and society. Being able to communicate with others is paramount to attain this goal. Since recovery from aphasia typically occurs over months or years (Chapey, 2008a; Flowers et al., 2016; Nichol et al., 2019; Pedersen et al., 2003; Simmons-Mackie, 2017), and it is often incomplete (Laska et al., 2001; Lazar et al., 2010; Pedersen et al., 1995, 2003), reintegration in society can be successful only if there are opportunities to engage with a community wherein the specific ways that PwA can contribute are valued and encouraged. This can hardly occur solely as the result of a person's motivation to find a new place in society, rather it requires a system of support for PwA and their families. A common concern expressed by therapists is that the current service delivery model provides very little in the way of support to PwA once they are discharged from inpatient or outpatient rehabilitation (Chapey et al., 2000; Nichol et al., 2019; Nichol, Pitt, et al., 2023; Nichol, Rodriguez, et al., 2023; Nichol, Wallace, Pitt, Rodriguez, Diong, et al., 2022; Nichol, Wallace, Pitt, Rodriguez, & Hill, 2022; Simmons-Mackie, 2017). Currently, aphasia centers, defined by Elman as *"a service delivery model that provides an interactive community for persons with aphasia*" (Elman, 2016, p. 154), represent a promising solution to this problem: their success, measured by the positive changes in quality of life and satisfaction expressed by PwA and their caregivers (Mayo et al., 2015), is attributed to their guiding philosophy which conceives them as environments where people can engage with activities that are calibrated to their needs and ambitions; where there is a supportive community of PwA, professionals, caregivers and trained volunteers that have a deep understanding of the challenges posed by aphasia; and where one feels to be a valued member of the community and is encouraged to contribute to the well-being of others (Elman, 2016; Simmons-Mackie & Holland, 2011). Nonetheless, it has been estimated that, as of 2016, throughout North America there were only about 26 such centers (Simmons-Mackie, 2017). Therefore, despite their promise, they currently serve only a fraction of the people in need.

Towards a Self-Management Approach to Aphasia

A vast literature portrays a service delivery system that fails to address the needs of PwA on multiple fronts: on the one hand, understaffing, high costs and difficulties with transportation drastically limit accessibility to therapy services, a problem that is particularly accentuated in the chronic phase; on the other hand, little emphasis is placed on understanding the goals and interests of PwA, the unique view of their carers, and on providing them with adequate education and resources on how to manage aphasia, an approach that has long been endorsed by the international community of researchers and therapists (Chapey et al., 2000; Hebert et al., 2016; Simmons-Mackie, 2017; Winstein et al., 2016).

When it comes to speech-language therapy, the gap between demand and supply is staggering and, realistically, it cannot be closed by doing "more of the same", like simply hiring more therapists. While this would certainly help, there are various reasons to believe that it is neither realistic, nor the most adequate solution to address the problems at hand. First of all, over the last decades there has been a clear trend towards tightening the budget to fund the healthcare system (Chapey et al., 2000; Chapey, 2008a; Rose et al., 2013; Simmons-Mackie, 2017), and this trend is not expected to reverse in the near future. Second, while increasing the size of the therapists' workforce could lead to a reduction in the tariffs that PwA would have to pay, it would probably still not be enough to be accessible to most people. Third, it does not guarantee that the barriers related to transportation would be removed: accessing outpatient services would remain as difficult to people living in rural areas, and to those with motor disorders living in metropolitan areas and relying on public transport.

There is currently a consensus regarding the fact that meeting the challenges faced by PwA requires a rethinking of how the service can be delivered (Chapey et al., 2000; Nichol et al., 2019; Nichol, Pitt, et al., 2023; Nichol, Rodriguez, et al., 2023; Nichol, Wallace, Pitt, Rodriguez, Diong, et al., 2022; Nichol, Wallace, Pitt, Rodriguez, & Hill, 2022; Palmer et al., 2019; Simmons-Mackie, 2017). In this regard, one idea that is gaining traction is that of *patient self-management*. The self-management philosophy moves away from a provider-centered service delivery model, wherein clients are mostly passive recipients of therapy, to a client-centered one, where they are encouraged and supported in regaining their autonomy, to the extent that this is possible. The emphasis is placed on ensuring that clients take an active role in communicating their priorities, that they contribute in setting realistic goals and planning of therapy, that they are educated on how to monitor and communicate their symptoms and needs, and in self-administering therapy (de Silva, 2011; Nichol et al., 2019). Key for this approach to be effective is that clients understand the nature of their condition and of the key factors that play a role in promoting optimal recovery (Warner et al., 2015); that they embrace a positive mindset that holds the target of achieving therapy goals center and front, while trusting in their capacity to do so despite the challenges that they will inevitably face (Jones & Riazi, 2011); that they are instructed in how to recognize and overcome challenges, and that they receive adequate external support when needed (Nichol, Rodriguez, et al., 2023; Nichol, Wallace, Pitt, Rodriguez, Diong, et al., 2022; Nichol, Wallace, Pitt, Rodriguez, & Hill, 2022).

Importantly, while it might seem that within the context of patient self-management the role of the therapist is diminished, this is far from being the case: the reduction in time spent delivering face-to-face therapy that is brought about by the (partial) shift to self-administered therapy is balanced by taking on new responsibilities, such as monitoring progress and updating therapy plans, providing psychological support when needed, education on aphasia and on self-management (Nichol, Pitt, et al., 2023; Nichol, Rodriguez, et al., 2023). Adopting a self-management approach, therapists can embrace all these roles by spending less time on an individual patient case than they would, had they also the duty to deliver therapy directly, hence allowing them to serve more clients. The combined effect of empowering therapists to reach more clients, decreasing the costs per client and shifting towards a client-centered approach hold the promise to transcend the limitations of the current service delivery model (Nichol et al., 2019; Nichol, Rodriguez, et al., 2023; Palmer et al., 2019).

The Role of Motivation in Patient Self-Management

As mentioned, a central tenet of the self-management approach is to promote, to the maximum extent that is possible, a person's autonomy in managing their condition, in particular when it comes to self-administering therapy — a task that PwA are expected to spend a significant amount of their time on. As such, it is crucial to ensure that they adhere to the levels of intensity that are required to maximize the chances of recovery. This is far from trivial. Aphasia therapy is a long process that requires high levels of intensity and dosage to attain optimal recovery (Brady et al., 2022), and that often entails drilling on specific language processes, which can feel very monotonous and unpleasant to endure (Biel et al., 2022). Therapists often admit that, when clients are tasked with self-administering specific exercises at home in-between face-to-face therapy sessions, they frequently do not do their homework, mentioning as the most likely causes the lack, or loss, of motivation due to the feeling that practice is not effective, the lack of support, and depression (Harmon et al., 2018).

The role of motivation as a factor that can influence therapy outcomes is widely recognized in the field of aphasia rehabilitation. In this regard, recent works have used a particular strand of motivation theory — self-determination theory (SDT) — to better understand how to inform aphasia interventions and management in such a way as to maximize motivation (Biel et al., 2022; Biel & Haley, 2023). Briefly, SDT posits that, in order for people to be motivated to do something challenging, there are three fundamental psychological needs that must be satisfied: the need for competence, relatedness, and autonomy (Ryan & Deci, 2017).

The *need for competence* refers to the fact that people are more likely to engage with a certain task, and to work towards a goal, if they feel confident in their own abilities to do so (Biel & Haley, 2023; Ryan & Deci, 2017). In line with one of the tenets of the self-management approach, this need is satisfied when people have a good understanding of their own condition, and of the challenges and potential rewards to be expected when embarking on the rehabilitation journey. Also, feedback provided throughout therapy plays an important role, as information on performance helps clients keeping track of their progress, which boosts the belief in one's capacity, and understanding what aspects of their behavior they need to work more on.

The *need for relatedness* refers to the fact that, for clients, it is important to perceive that those who support them have a genuine, unconditional interest in their own wellbeing, as opposed to being motivated to help for pure material gains.

Finally, the *need for autonomy* reflects the fact that people are more likely to engage in a given behavior, and to do so more fruitfully, if they perceive that the actions that they have to take align with their values and goals, as opposed to choosing to undertake them solely to satisfy an external constraint (i.e. regulation, reward or punishment) whose value they do not appreciate. Within the SDT framework, these two categories of motivations are termed autonomous and controlled, respectively. Both can be promoted by different sources, whose effectiveness in driving behavior varies along a continuum based on the degree to which they align with a person's values, interests, and goals. For example, in the context of self-management of aphasia, autonomous motivation can be promoted when a person understands that engaging with therapy will lead to improvements on a specific aspect of communication, which in turn could lead to improving one's condition by facilitating reintegration in society, and/or the pursuit of their goals. The strongest, most effective kind of motivation is *intrinsic motivation*, which people experience when they engage with an activity for its own sake (i.e., a hobby, a passion, or an activity that is pleasant and fun; Biel & Haley, 2023; Ryan & Deci, 2017).

Current Research on Self-Management of Aphasia

While currently little has been done in the way of implementing a comprehensive self-management approach to aphasia that follows a client from the acute to the chronic phase, the field of aphasia rehabilitation has embraced many of the tenets of the self-management philosophy. A recent scoping review on this topic underscored the presence of three strands of research, each focusing on a key aspect of the self-management philosophy (Nichol et al., 2019).

The first strand is concerned with group and community-based interventions to: promote life participation and social interaction by delivering speech-language therapy to groups of PwA; providing education on the nature of aphasia, along with the challenges and opportunities presented during the process of recovery; providing counseling on mental health issues. The second strand is concerned with training communication partners on how to support PwA in communicating their needs, their ambitions and specific goals, and on how to provide a context in which to practice and improve their communication in a safe environment. The third strand of research is mostly concerned with developing and testing computer technologies that enable self-administration of therapy in a cost-effective, accessible fashion.

This dissertation is concerned primarily with the last strand of research. Therefore, below I provide an overview of the potential roles that CBAT can play in the management of aphasia, of the state-of-the-art in the field, and of the main factors that currently prevent CBATs from reaching their full potential, which ultimately this thesis and the wider project it is embedded in seek to overcome.

The Potential of Computer-Based Therapies: Self-Management and Personalized Therapy

Modern technologies have long been considered as a possible way to improve accessibility of therapy via tele-rehabilitation (Caughlin et al., 2020; Cherney & Van Vuuren, 2012; Choi et al., 2016; Meltzer et al., 2018) or by enabling PwA self-administering therapy remotely, from the comfort of their home (Chapey, 2008b; Lavoie et al., 2017). In addition to this, CBATs allow therapists to shift their focus away from delivering intensive drill-based therapies to providing their clients with all the tools they need to be as autonomous as they can. This can include providing education on aphasia, counseling on the best course of action based on clients' specific needs and goals, adjusting their recommendations on the fly by remote monitoring of performance, providing encouragement, feedback and support when required (Nichol, Pitt, et al., 2023; Nichol, Rodriguez, et al., 2023). This can allow therapists to provide more services in a designated amount of time, and/or service more clients. As such, the adoption of CBATs empowers therapists as well, boosting their productivity and allowing them to serve more clients.

Aside from the obvious advantage of expanding the reach of this service to anyone with a computer and an internet connection, the large-scale deployment and adoption of CBATs holds the promise to break new ground on matters that are relevant to clinical practice. As discussed in the section on the quest for understanding how to personalize treatment, several biographical, neuropsychological, and therapy parameters are likely to affect therapy outcomes, and therefore delivering therapy that is tailored around the strengths and weaknesses of each person requires an understanding of how these parameters interact. Key to achieve this feat is to build datasets that are big enough to capture a wide range of variability within each of the factors of interest, which could then be analyzed using modern statistical or machine learning techniques that are known to excel at uncovering relevant patterns in the data. In this regard, CBATs, when deployed at large scale, unlock the possibility to easily and quickly build up such datasets by collecting data from hundreds of thousands of people throughout their entire rehabilitation journey, a resolution that no research study has ever possessed. This effectively brings the prospect of delivering personalized therapy — the "holy grail" of any medical endeavor that is concerned with complex, multidimensional disorders such as aphasia — within reach for the first time after decades of unsuccessful efforts.

State of the Art of Computer-Based Therapies

The interest in CBATs dates back to the dawn of the era of personal computers (Linebarger et al., 2007), and ever since then research output has increased steadily. Recent systematic reviews and meta-analysis (Brady et al., 2016; Lavoie et al., 2017; Zheng et al., 2016), along with large-scale studies investigating the efficacy of specific applications (Braley et al., 2021; DesRoches et al., 2015; DesRoches & Kiran, 2017; Liu et al., 2023; Palmer et al., 2019, 2020), have largely corroborated not only the feasibility of using CBATs to enable self-administered, home-based therapy, but also their non-inferiority relative to face-to-face therapy when it comes to measure the gains across a wide range of language functions (Brady et al., 2016, 2022).

However, to the best of my knowledge, currently only few of the applications whose efficacy has been investigated in research studies have been commercialized (*Constant Therapy*, 2022; *StepByStep Aphasia Therapy*, n.d.), or else made accessible to the public. Commercially available applications have already proven their utility not just in supporting self-administration of language therapy, but also in building the empirical evidence required to better characterize, with a high level of granularity, the effect of specific therapy parameters (such as dosage; Liu et al., 2023, and frequency; Cordella et al., 2022) on therapy outcomes, something that could not be adequately illuminated by traditional research studies.

Limitations of Computer-Based Therapies for Aphasia

As discussed earlier, rehabilitation from aphasia poses several challenges to PwA. Aside from the issues associated with accessibility, for therapy to be effective patients have to engage with high amounts of speech therapy, every day, several days per week. As progress is typically slow and hinges on drilling on tasks that are very monotonous, maintaining the motivation required to adhere to the necessary levels of therapy represent a significant challenge for most PwA.

A recent study assessing whether computerized therapy can serve as a cost-effective tool to complement usual care suggested that the extent of this challenge is significant, with about 75% of participants (out of a pool of 85) engaging with computer-therapy for half or less of the recommended time (Palmer et al., 2019, 2020). In order to understand how to induce desirable changes in behavior in those who showed inadequate levels of adherence with therapy, a follow-up study set out to identify which factors set apart participants with low versus high levels of adherence (Harrison et al., 2020). The findings showed that inability to use the software, poor knowledge about the nature of aphasia and the importance of practice, and cognitive deficits did affect the capability of PwA to adhere to the recommended frequency, intensity and dosage. However, capability affords little without the *opportunity* to engage with therapy. There are both technical and social factors that determine the extent to which one can enjoy the opportunity to engage with therapy. For example, poor usability or frequent software breakdowns would severely limit a person's opportunity to engage with therapy, no matter their ability to use the software, while a well-designed user interface built on top of a robust software would be considered as an enabler of opportunity to practice. Similarly, the presence of social support (e.g. constant and timely input from the therapist, timely assistance from a volunteer or carer whenever a problem related to software/hardware ensued) is the strongest enabler of opportunity to engage with therapy, as results of this study confirmed. Finally, in line with the basic tenets of SDT, motivational factors that promoted high adherence satisfied both the need for competence and for autonomy. Generally, people who adhered to recommended levels of practice shared a strong belief in the critical role of therapy to drive recovery, and in their own capacity to improve via training. Having strong and stable intentions favored the formation of habits around training regularly. Importantly, beliefs could change throughout the course of treatment: indeed, many of the people who failed to adhere with recommended levels of practice shared positive beliefs at the beginning of their experience, but lamented a waning of motivation — due mostly to lack of support and slow progress — that led to attrition. Receiving feedback on the amount of time spent practicing motivated users to keep up with practice, while receiving feedback on task performance could have negative effects on motivation when perceived to be inaccurate, with participants showing a sensitivity specifically for false negatives (i.e., instances where the ASR flagged correct answers as wrong).

Aside from these insights, which are likely to apply to most currently available CBATs, it is worth noting that ASR-based automatic feedback on vocal production has currently been employed in very few applications (e.g., *Constant Therapy*, 2022) other than the one used in the study discussed above (*StepByStep Aphasia Therapy*, n.d.) and, to my knowledge, an assessment of its accuracy, of the feasibility of employing it to assist PwA with their naming, and of users' opinions about its usefulness has been conducted in only one study (Ballard et al., 2019).

It is clear that, while immense progress has been made with advancing the technologies that can enable therapy to be self-administered, there are important steps to be made in order to ensure that they can realize their full potential. As highlighted by the findings of Harrison's work (Harrison et al., 2020), some factors are exogenous to the design of the technologies themselves. For example, a key factor positively affecting adherence is receiving regular support with the use of technology, encouragement and education from volunteers/trained family members. Other aspects are purely associated with the design of the technologies. One pertains to design that promotes engagement and adherence by inducing intrinsic motivation; the other pertains to understanding the role of automatic feedback, which differs in important ways from that provided by therapists in guiding behaviors and promoting motivation.

This dissertation is part of a wider project that attempts to address these two aspects related to design to improve upon the state of the art in CBATs for aphasia. The next two subsections provide a brief overview of an approach that has proven successful in promoting intrinsic motivation in contexts where high levels of task engagement and adherence with therapy are key, and of the role that feedback plays in promoting motivation.

Promoting Engagement via Video-Game Based Activities

Over the last decade lots of progress has been made on expanding the scope (i.e. number and variety of tasks offered) and adaptivity of CBATs (i.e., tracking progress at a fine-grained level of resolution and adapting the level of difficulty accordingly). However, while some key features that are known to promote engagement are present in most applications (i.e. points, leaderboards, progress graphs), current digital technologies afford to take a step further, taking therapeutic activities out of their decontextualized setting (i.e., simple picture-naming) and embedding them into virtual, real-life like scenarios that are meaningful and engaging to users (i.e., playing a character that explores a market and has to make requests for specific items).

In this regard, video game-based programs to test, monitor and treat cognitive disorders have been adopted to serve both children with special needs and older adults with cognitive disorders. When it comes to testing and monitoring cognitive abilities, the key to obtain a reliable measure (i.e. one that is stable across measurements taken at different times) is that subjects do their best to focus on the task at hand. Children and older adults, especially those with impaired cognitive abilities, find this quite difficult, especially with traditional pen and pencil tests (J. Anguera et al., 2016). In this context, video game-based cognitive testing is adopted as a strategy to direct and maintain attention on the desired task. Empirical results corroborate the validity of this approach, enabling the test sensitivity required to discriminate children with and without deficits in cognitive abilities due to genetic disorders (J. Anguera et al., 2016), and to reliably monitor cognitive abilities in children with autism, attention deficit hyperactivity disorder (ADHD), and learning disorders (Flynn et al., 2019).

When it comes to treatment, the key issue is adherence. Here too video game-based programs have proven beneficial. Studies adopting such programs in the context of ameliorating processing speed in patients with multiple sclerosis (MS) (Bove et al., 2019), cognitive control in older adults (J. A. Anguera et al., 2013), showed beneficial effects in terms of adherence and treatment efficacy.

Based on these highly promising results, it is safe to expect that game-based mechanics and video games more in general hold the potential to drastically improve the level of engagement, adherence, and in turn effectiveness of the therapeutic experience of PwA, especially in the context of home-based, self-administered treatment.

The Role of Feedback on Motivation

There is ample consensus among therapists regarding the importance of feedback in driving recovery of language skills (Chapey, 2008; Simmons-Mackie et al., 1999). In this context, feedback is thought to serve mainly two functions: to provide information on response accuracy that reinforces correct retrieval, or to signal the need to adjust responses towards the intended target (Ballard et al., 2019; Chapey, 2008; Nunn et al., 2023; Nunn & Vallila-Rohter, 2023; Palmer et al., 2019, 2020; Swales et al., 2016); and to boost a person’s confidence by providing encouragement (Simmons-Mackie et al., 1999). The first function is all the more important when considering that people with aphasia might fail to realize when they make a mistake (Ballard et al., 2019; Chapey, 2008; Nunn & Vallila-Rohter, 2023; Palmer et al., 2020). The integration of ASRs into CBATs is therefore a necessary step to make it possible for PwA to self-administer naming therapy even in the absence of a therapist (Ballard et al., 2019; Barbera et al., 2021; Palmer et al., 2019, 2020; Swales et al., 2016). Empirical evidence backs up these considerations, suggesting that feedback is an indispensable feature for speech-language therapy to be effective. In a recent study, Sze et al. (2020) presented a meta-analysis of single-case studies investigating the ingredients that were most strongly associated with gains in naming skills. Pooling together 32 studies including more than 100 subjects, the analysis revealed that providing feedback on response accuracy was one of the strongest factors associated with gains during the treatment phase, and the one most strongly associated with the extent to which such gains were maintained at a later follow-up session held more than 3 weeks after the last training session.

However, the empirical evidence on the importance of feedback on patients' motivation and on the extent to which they improve their naming skills does largely come from studies where therapy was delivered face-to-face. There are two important differences between face-to-face and self-administered therapy, wherein the feedback is delivered by the computer.

First, as PwA might find it difficult to detect errors in their speech (Ellis et al., 1983; Maher et al., 1994; Marshall et al., 1985, 1994; Marshall & Tompkins, 1982; Nickels & Howard, 1995; Sampson & Faroqi-Shah, 2011; Schwartz et al., 2016), their self-assessment might sometimes diverge from that of a therapist or of an ASR. These are the instances where feedback would provide more value to the user, provided that it is considered trustworthy. The amount of trust that someone puts in somebody else's judgment, opinion, or recommendation is strongly influenced by the perceived level of expertise of the interlocutor (Chaiken & Maheswaran, 1994), even when this is a technology (Reeves & Nass, 1996; Wienrich et al., 2021). However, while normally a person with aphasia would confidently trust their therapist over their self-assessment of the accuracy of their vocal output, this is not necessarily the case when this judgment is made by a computer program, even one yielding excellent performance in determining the accuracy of single-word utterances. This skepticism could be further exacerbated by the occasional mistakes made by ASR. A failure of an ASR to be perceived as trustworthy by users could have detrimental effects: at best, users might ignore the feedback and rely just on their ability to monitor their speech, which is detrimental in case self-monitoring abilities are poor; at worst, a distrust in the ASR — coupled with a failure to use its feedback — can frustrate users and distract them from the task at hand, potentially even discontinuing use of the program. Currently, among the few studies investigating the effectiveness of CBATs that integrate an ASR, the focus was not placed on understanding what factors affect users' perception of the accuracy and usefulness of ASR feedback, nor how these in turn affect the extent to which they benefit from training.

The second important difference between face-to-face and self-administered therapy is in the way feedback is delivered. Clinical recommendations stress the importance of delivering feedback only when necessary, namely when it has informational and motivational value (Chapey, 2008a). During face-to-face therapy, therapists use a range of explicit (i.e. clients not showing awareness of their errors) or implicit (i.e., body language suggesting clients would like to receive guidance) cues to determine whether feedback is needed or not (Simmons-Mackie et al., 1999). This is in line with the concept of autonomy as proposed by SDT, which states that when a task, or any information received, is perceived to be meaningful, it promotes stronger motivation; conversely, feedback delivered automatically after every response might be perceived as imposed on a person, and processed akin to a controlling form of motivation that induces passivity and reduces the sense of control and ownership of their experience (Biel et al., 2022; Biel & Haley, 2023; Ryan & Deci, 2017). While some authors have proposed to give clients choice over when to receive feedback as a way to foster engagement and motivation (Biel et al., 2022; La Guardia, 2017; Nunn & Vallila-Rohter, 2023; Ryan & Deci, 2017), currently the prevailing approach among studies employing CBAT with (Ballard et al., 2019; Palmer et al., 2019, 2020) and without ASR (see reviews conducted by Lavoie et al., 2017; and Zheng et al., 2016) is to deliver feedback immediately after participants' responses (i.e., be them vocal or written responses, or simply a choice among multiple possible ones in odd-one-out tasks, picture-word tasks etc.). While the theoretical and clinical considerations discussed above suggest that this might not be an optimal strategy, to the best of our knowledge there are no empirical studies investigating the conditions under which automatic feedback is most effective at serving its informational and motivational roles.

The Focus of This Thesis

CBATs play a crucial role in the transition to a service delivery model that is more accessible, effective, and financially sustainable. Not only can such technologies enable anyone with a computer or tablet and an internet connection to engage with the necessary amount of therapy from the comfort of their home, at a fraction of the price per month of what they pay for an hour of face-to-face therapy. CBATs can also empower therapists to cover up additional crucial functions related to patient management, and to serve more clients. Finally, and not least important, the large-scale deployment and adoption of CBATs unlocks the possibility to easily build up datasets encoding enough variability among biographical, neuropsychological, and specific therapy parameters to understand what role they play in recovery from aphasia. This effectively brings the prospect of delivering personalized therapy within reach for the first time after decades of unsuccessful efforts. This thesis forms the foundation of a broader project whose long-term goal is to design, validate and commercialize a CBAT, with an emphasis on improving upon some limitations that were identified in the state of the art.

As a first step on this journey, this dissertation presents a preliminary assessment of the feasibility and efficacy of a CBAT that integrates an ASR by: engaging with PwA, their carers, and therapists to identify the most common unmet needs of individuals with aphasia, the barriers and facilitators in the use of technologies that can inform the design of a prototype, and their opinions and preferences regarding the adoption of video games and voice recognition (chapter 2); using this information to inform the design and development of the first prototype of a CBAT, assess its usability by having a group of PwA test it directly, and use the results of this investigation to improve upon its design (chapter 3); conducting a preliminary assessment of the feasibility of deploying the prototype to enable PwA to self-administer naming and word-finding therapy from their home, and of the benefits that PwA can reap from doing so; quantifying whether delivering feedback on response accuracy via the ASR yields additional benefits relative to delivering no feedback, and providing insights into the factors that might predict when feedback is necessary or not (chapter 4).

The work conducted as part of this dissertation will hopefully provide the basis upon which to integrate additional improvements (i.e., novel game-mechanics, wider variety of tasks and stimuli, adaptive level of difficulty, optimized delivery of feedback etc.) and to make the training more enjoyable and, ultimately, effective.

REFERENCES

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

Allen, L., Mehta, S., McClure, J., Teasell, R. (2012). Therapeutic interventions for aphasia initiated more than six months post stroke: A review of the evidence. *Topics in Stroke Rehabilitation*. 19, 523-535.

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., Kong, E., Larraburo, Y., Rolle, C., Johnston, E., & others. (2013). Video game training enhances cognitive control in older adults. *Nature*, *501*(7465), 97–101.

Anguera, J., Brandes-Aitken, A., Rolle, C., Skinner, S., Desai, S., Bower, J., Martucci, W., Chung, W., Sherr, E., & Marco, E. (2016). Characterizing cognitive control abilities in children with 16p11. 2 deletion using adaptive 'video game' technology: A pilot study. *Translational Psychiatry*, *6*(9), e893–e893.

Arcury, T. A., Preisser, J. S., Gesler, W. M., & Powers, J. M. (2005). Access to transportation and health care utilization in a rural region. *The Journal of Rural Health*, *21*(1), 31–38.

Baddeley, A., & Wilson, B. A. (1994). When implicit learning fails: Amnesia and the problem of error elimination. *Neuropsychologia*, *32*(1), 53–68.

Baker, J. C., LeBlanc, L. A., & Raetz, P. B. (2008). A behavioral conceptualization of aphasia. *The Analysis of Verbal Behavior*, *24*, 147–158.

Ballard, K. J., Etter, N. M., Shen, S., Monroe, P., & Tien Tan, C. (2019). Feasibility of Automatic Speech Recognition for Providing Feedback During Tablet-Based Treatment for Apraxia of Speech Plus Aphasia. *American Journal of Speech-Language Pathology*, *28*, 818–834.

Bersano, A., Burgio, F., Gattinoni, M., & Candelise, L. (2009). Aphasia burden to hospitalized acute stroke patients: Need for an early rehabilitation programme. *International Journal of Stroke*, *4*(6), 443–447.

Berthier, M. L., & Pulvermüller, F. (2011). Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nature Reviews Neurology*, *7*(2), 86–97.

Biel, M., Enclade, H., Richardson, A., Guerrero, A., & Patterson, J. (2022). Motivation theory and practice in aphasia rehabilitation: A scoping review. *American Journal of Speech-Language Pathology*, *31*(5S), 2421–2443.

Biel, M., & Haley, K. L. (2023). Motivation in Aphasia Treatment: Self-Determination Theory Applied to the FOURC Model. *American Journal of Speech-Language Pathology*, *32*(3), 1016–1036.

Bove, R. M., Rush, G., Zhao, C., Rowles, W., Garcha, P., Morrissey, J., Schembri, A., Alailima, T., Langdon, D., Possin, K., & others. (2019). A videogame-based digital therapeutic to improve processing speed in people with multiple sclerosis: A feasibility study. *Neurology and Therapy*, *8*, 135–145.

Boyle, M. (2010). Semantic feature analysis treatment for aphasic word retrieval impairments: What's in a name? *Topics in Stroke Rehabilitation*, *17*(6), 411–422.

Brady, M. C., Ali, M., Berg, K. V., Williams, L. J., Williams, L. R., Abo, M., Becker, F., Bowen, A., Brandenburg, C., Breitenstein, C., & others. (2022). Complex speech-language therapy interventions for stroke-related aphasia: The RELEASE study incorporating a systematic review and individual participant data network meta-analysis. *Health and Social Care Delivery Research*, *10*(28).

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *Cochrane Database of Systematic Reviews*, *6*. https://doi.org/10.1002/14651858.CD000425.pub4.

Braley, M., Sims Pierce, J., Saxena, S., De Oliveira, E., Taraboanta, L., Anantha, V., Lakhan, S. E., & Kiran, S. (2021). A Virtual, Randomized, Control Trial of a Digital Therapeutic for Speech, Language, and Cognitive Intervention in Post-stroke Persons With Aphasia. *Frontiers in Neurology*, *12*(626780). https://doi.org/10.3389/fneur.2021.626780

Canadian Stroke Network. (2011). *The Quality Of Stroke Care in Canada*. https://www.canadianstroke.ca/sites/default/files/resources/QoSC-EN.pdf

Canadian Stroke Best Practices. (2024). https://www.strokebestpractices.ca/recommendations/cerebral-venous-thrombosis/2-acute-treatment-of-symptomatic-cerebral-venous-thrombosis

Caporali, A., & Basso, A. (2003). A survey of long-term outcome of aphasia and of chances of gainful employment. *Aphasiology*, *17*(9), 815–834.

Carmichael, S. T. (2016). The 3 Rs of stroke biology: Radial, relayed, and regenerative. *Neurotherapeutics*, *13*, 348–359.

Carrera, E., & Tononi, G. (2014). Diaschisis: Past, present, future. *Brain*, *137*(9), 2408–2422.

Caughlin, S., Mehta, S., Corriveau, H., Eng, J. J., Eskes, G., Kairy, D., Meltzer, J., Sakakibara, B. M., & Teasell, R. (2020). Implementing telerehabilitation after stroke: Lessons learned from Canadian trials. *Telemedicine and E-Health*, *26*(6), 710–719.

Chang, Y.-N., & Lambon Ralph, M. A. (2020). A unified neurocomputational bilateral model of spoken language production in healthy participants and recovery in poststroke aphasia. *Proceedings of the National Academy of Sciences*, *117*(51), 32779–32790.

Chapey, R. (2008a). *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders* (5th ed.). Lippincott Williams and Wilkins.

Chapey, R., Duchan, J., Elman, R., Garcia, L., Kagan, A., Lyon, J., & Simmons-Mackie, N. (2000). *Life Participation Approaches to Aphasia: A statement of values for the future.* (5; pp. 4–6). ASHA Leader.

Cherney, L. R., & Van Vuuren, S. (2012). Telerehabilitation, virtual therapists, and acquired neurologic speech and language disorders. *Seminars in Speech and Language*, *33*(03), 243–258.

Choi, Y.-H., Park, H. K., & Paik, N.-J. (2016). A telerehabilitation approach for chronic aphasia following stroke. *Telemedicine and E-Health*, *22*(5), 434–440.

Christensen, J. M., & Anderson, J. D. (1989). Spouse adjustment to stroke: Aphasic versus nonaphasic partners. *Journal of Communication Disorders*, *22*(4), 225–231.

Code, Chris, & Petheram, B. (2011). Delivering for aphasia. *International Journal of Speech-Language Pathology*, *13*(1), 3–10. https://doi.org/10.3109/17549507.2010.520090

Constant Therapy. (2022, February). https://constanttherapyhealth.com/science/

Copland, D. (2003). The basal ganglia and semantic engagement: Potential insights from semantic priming in individuals with subcortical vascular lesions, Parkinson’s disease, and cortical lesions. *Journal of the International Neuropsychological Society*, *9*(7), 1041–1052.

Cordella, C., Munsell, M., Godlove, J., Anantha, V., Advani, M., & Kiran, S. (2022). Dosage frequency effects on treatment outcomes following self-managed digital therapy: Retrospective cohort study. *Journal of Medical Internet Research*, *24*(7), e36135.

Cramer, S. C. (2008). Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. *Annals of Neurology*, *63*(3), 272–287.

Crosson, B., Benefield, H., Cato, M. A., Sadek, J. R., Moore, A. B., Wierenga, C. E., Gopinath, K., Soltysik, D., Bauer, R. M., Auerbach, E. J., & others. (2003). Left and right basal ganglia and frontal activity during language generation: Contributions to lexical, semantic, and phonological processes. *Journal of the International Neuropsychological Society*, *9*(7), 1061–1077.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 318.

Damasio, A. (1998). Signs of aphasia. In *Acquired aphasia* (M.T.Sarno, pp. 25–41). Academic Press.

Davis, G. A. (2013). *Aphasia and related cognitive-communicative disorders*. Pearson Higher Ed.

de Silva, D. (2011). *Evidence: Helping people help themselves*.

Debray, T. P., Schuit, E., Efthimiou, O., Reitsma, J. B., Ioannidis, J. P., Salanti, G., Moons, K. G., & Workpackage, G. (2018). An overview of methods for network meta-analysis using individual participant data: When do benefits arise? *Statistical Methods in Medical Research*, *27*(5), 1351–1364.

DeLeon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., Lee, A., & Hillis, A. E. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain*, *130*(5), 1408–1422.

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*, *93*(3), 282–321.

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, *21*(2–4), 125–145.

Dell, G. S., Nozari, N., & Oppenheim, G. M. (2014). Word production: Behavioral and computational considerations. *The Oxford Handbook of Language Production*, 88–104.

Dell, G. S., & O'Seaghdha, P. (1992). Stages of lexical access in language production. *Cognition*, *42*, 287–314.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review*, *104*(4), 801–838.

Dell, G. S., Schwartz, M. F., Nozari, N., Faseyitan, O., & Coslett, H. B. (2013). Voxel-based lesion-parameter mapping: Identifying the neural correlates of a computational model of word production. *Cognition*, *128*(3), 380–396.

DesRoches, C. A., Balachandran, I., Ascenso, E. M., Tripodis, Y., & Kiran, S. (2015). Effectiveness of an impairment-based individualized rehabilitation program using an iPad-based software platform. *Frontiers In Human Neuroscience*, *8*(1015). https://doi.org/doi: 10.3389/fnhum.2014.01015

DesRoches, C. A., & Kiran, S. (2017). Technology-Based Rehabilitation to Improve Communication after Acquired Brain Injury. *Frontiers In Human Neuroscience*, *11*(382). https://doi.org/10.3389/fnins.2017.00382

Dickey, L., Kagan, A., Lindsay, M., Fang, J., Rowland, A., & Black, S. (2010). Incidence and profile of inpatient stroke-induced aphasia in Ontario, Canada. *Archives of Physical Medicine and Rehabilitation*, *91*(2), 196--202.

Doesborgh, S. J., van de Sandt-Koenderman, M. W., Dippel, D. W., van Harskamp, F., Koudstaal, P. J., & Visch-Brink, E. G. (2004). Effects of semantic treatment on verbal communication and linguistic processing in aphasia after stroke: A randomized controlled trial. *Stroke*, *35*(1), 141–146.

Dorze, G. L., & Brassard, C. (1995). A description of the consequences of aphasia on aphasic persons and their relatives and friends, based on the WHO model of chronic diseases. *Aphasiology*, *9*(3), 239–255.

Efstratiadou, E. A., Papathanasiou, I., Holland, R., Archonti, A., & Hilari, K. (2018). A systematic review of semantic feature analysis therapy studies for aphasia. *Journal of Speech, Language, and Hearing Research*, *61*(5), 1261–1278.

Ellis, C., Simpson, A. N., Bonilha, H., Mauldin, P. D., & Simpson, K. N. (2012). The one-year attributable cost of poststroke aphasia. *Stroke*, *43*(5), 1429–1431.

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

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

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

Flowers, H., Skoretz, A., Silver, F., Rochon, E., Fang, J., Flamand-Roze, C., & Marinto, R. (2016). Poststroke aphasia frequency, recovery and outcomes: A systematic review and meta-analysis. *Archives of Physical Medicine and Rehabilitation*, *97*(12), 2188---2201.

Flynn, R. M., Colón-Acosta, N., Zhou, J., & Bower, J. (2019). A game-based repeated assessment for cognitive monitoring: Initial usability and adherence study in a summer camp setting. *Journal of Autism and Developmental Disorders*, *49*(5), 2003–2014.

Foley, N., McClure, J. A., Meyer, M., Salter, K., Bureau, Y., & Teasell, R. (2012). Inpatient rehabilitation following stroke: Amount of therapy received and associations with functional recovery. *Disability & Rehabilitation*, *34*(25), 2132–2138. https://doi.org/10.3109/09638288.2012.676145

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language*, *43*(2), 182–216.

Fridriksson, J., den Ouden, D.-B., Hillis, A. E., Hickok, G., Rorden, C., Basilakos, A., Yourganov, G., & Bonilha, L. (2018). Anatomy of aphasia revisited. *Brain*, *141*(3), 848–862.

Fridriksson, J., & Hillis, A. E. (2021). Current approaches to the treatment of post-stroke aphasia. *Journal of Stroke*, *23*(2), 183–201.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 27–52.

Garrett, M. F. (1975). The analysis of sentence production. In *Psychology of learning and motivation* (Vol. 9, pp. 133–177). Elsevier.

Geschwind, N. (1970). The Organization of Language and the Brain: Language disorders after brain damage help in elucidating the neural basis of verbal behavior. *Science*, *170*(3961), 940–944.

Gleichman, A. J., & Carmichael, S. T. (2014). Astrocytic therapies for neuronal repair in stroke. *Neuroscience Letters*, *565*, 47–52.

Godecke, E., Armstrong, E., Hersh, D., & Bernhardt, J. (2014). Missed opportunities: Communicative interactions in early stroke recovery. *International Journal of Stroke*, *9*(48).

Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*, *21*(6), 649–683.

Goodglass, H., & Wingfield, A. (1997). Anatomical and theoretical considerations in anomia. In *Anomia: Neuroanatomical and cognitive correlates* (H. Goodglass & A. Wingfield, pp. 20–29). Academic Press.

Graham, J. R., Pereira, S., & Teasell, R. (2011). Aphasia and return to work in younger stroke survivors. *Aphasiology*, *25*(8), 952–960.

Grefkes, C., & Fink, G. R. (2020). Recovery from stroke: Current concepts and future perspectives. *Neurological Research and Practice*, *2*(1), 1–10.

Grossman, M. (2014). Anomia. In *Encyclopedia of the Neurological Sciences* (M. J. Aminoff & R. B. Daroff (Eds.), Vol. 1, pp. 194–197). Academic Press.

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

Halai, A. D., Woollams, A. M., & Ralph, M. A. L. (2018). Triangulation of language-cognitive impairments, naming errors and their neural bases post-stroke. *NeuroImage: Clinical*, *17*, 465–473.

Harmon, T. G., Hardy, L., & Haley, K. L. (2018). Proactive social validation of methods and procedures used for training speech production in aphasia. *Aphasiology*, *32*(8), 922–943.

Harrison, M., Palmer, R., & Cooper, C. (2020). Factors Associated With Adherence to Self-Managed Aphasia Therapy Practice on a Computer—A Mixed Methods Study Alongside a Randomized Controlled Trial. *Frontiers in Neurology*, *11*.

Hartwell, J. (2015). *Access Denied: An Investigation of the Impact of Aphasia on Social Inclusion in Long-Term Care Facilities* [University of Louisiana at Lafayette]. https://www.proquest.com/openview/866791505f5dae3b6c9368461ae49702/1?pq-origsite=gscholar&cbl=18750

Heart & Stroke. (2017). *Stroke report 2017*. Heart and Stroke Foundation of Canada. https://www.heartandstroke.ca/-/media/pdf-files/canada/stroke-report/strokereport2017en.ashx?rev=48bbc6fe37d84c618687e259acc69920

Heart & Stroke. (2020). *Where every dollar goes*. Heart and Stroke. https://www.heartandstroke.ca/what-we-do/finances/where-every-dollar-goes

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*(2), 135–145.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402.

Hilari, K. (2011). The impact of stroke: Are people with aphasia different to those without? *Disability & Rehabilitation*, *33*(3), 211–218.

Holland, A., Fromm, D., Forbes, M., & MacWhinney, B. (2017). Long-term recovery in stroke accompanied by aphasia: A reconsideration. *Aphasiology*, *31*(2), 152–165.

Howard, D., Patterson, K. E., Franklin, S. E., Orchard-Lislie, V. M., & Morton, J. (1985). The treatment of word retrieval deficits in aphasia: A comparison of two therapy methods. *Brain*, *108*, 817–829. https://doi.org/10.1093/brain/108.4.817

Howard, G., & Goff, D. C. (2012). Population shifts and the future of stroke: Forecasts of the future burden of stroke. *Annals of the New York Academy of Sciences*, *1268*, 14–20. https://doi.org/10.1111/j.1749-6632.2012.06665.x

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 255.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144.

Jarso, S., Li, M., Faria, A., Davis, C., Leigh, R., Sebastian, R., Tsapkini, K., Mori, S., & Hillis, A. E. (2013). Distinct mechanisms and timing of language recovery after stroke. *Cognitive Neuropsychology*, *30*(7–8), 454–475.

Jones, F., & Riazi, A. (2011). Self-efficacy and self-management after stroke: A systematic review. *Disability and Rehabilitation*, *33*(10), 797–810.

Joy, M. T., & Carmichael, S. T. (2021). Encouraging an excitable brain state: Mechanisms of brain repair in stroke. *Nature Reviews Neuroscience*, *22*(1), 38–53.

Kagan, A., Simmons-Mackie, N., Kapral, N., Hoch, J., Graves, E., Mecredy, G., Whitehurst, D., & Chan, M. (2017, June 3). *Health care utilization and costs in people with stroke with and without aphasia in Ontario, Canada* [Poster presentation at the Clinical Aphasiology Conference].

Kerr, E., Ivanova, B., & Strijkers, K. (2022). Lexical Access in Speech Production. Psycho- and neurolinguistic perspectives on the spatiotemporal dynamics. In *Current issues in the Psychology of Language.* Routledge Press (Taylor and Francis).

Kiran, S., Meier, E. L., & Johnson, J. P. (2019). Neuroplasticity in aphasia: A proposed framework of language recovery. *Journal of Speech, Language, and Hearing Research*, *62*(11), 3973–3985.

Kiran, S., & Thompson, C. K. (2003). *The role of semantic complexity in treatment of naming deficits*.

Kiran, S., & Thompson, C. K. (2019). Neuroplasticity of language networks in aphasia: Advances, updates, and future challenges. *Frontiers in Neurology*, *10*, 295.

Kleim, J. A. (2011). Neural plasticity and neurorehabilitation: Teaching the new brain old tricks. *Journal of Communication Disorders*, *44*(5), 521–528.

Kleim, J. A., & Jones, T. A. (2008). *Principles of experience-dependent neural plasticity: Implications for rehabilitation after brain damage*.

Krakauer, J. W., & Carmichael, S. T. (2022). *Broken movement: The neurobiology of motor recovery after stroke*. MIT Press.

Kristinsson, S. (2021). *Towards Understanding Therapy Response in Chronic Aphasia* [PhD Thesis]. University of South Carolina.

Kristinsson, S., Basilakos, A., Elm, J., Spell, L. A., Bonilha, L., Rorden, C., Den Ouden, D. B., Cassarly, C., Sen, S., Hillis, A., & others. (2021). Individualized response to semantic versus phonological aphasia therapies in stroke. *Brain Communications*, *3*(3), fcab174.

Kümmerer, D., Hartwigsen, G., Kellmeyer, P., Glauche, V., Mader, I., Klöppel, S., Suchan, J., Karnath, H.-O., Weiller, C., & Saur, D. (2013). Damage to ventral and dorsal language pathways in acute aphasia. *Brain*, *136*(2), 619–629.

La Guardia, J. G. (2017). *Self-determination theory in practice: How to create an optimally supportive health care environment*. éditeur inconnu.

Laine, M., & Martin, N. (2006). *Anomia. Theoretical and Clinical Aspects* (1st ed.). Psychology Press.

Lam, J. M., & Wodchis, W. P. (2010). The relationship of 60 disease diagnoses and 15 conditions to preference-based health-related quality of life in Ontario hospital-based long-term care residents. *Medical Care*, 380–387.

Laska, A. C., Hellblom, A., Murray, V., Kahan, T., & Von Arbin, M. (2001). Aphasia in acute stroke and relation to outcome. *Journal of Internal Medicine*, *249*(5), 413–422.

Lavoie, M., Macoir, J., & Bier, N. (2017). Effectiveness of technologies in the treatment of post-stroke anomia: A systematic review. *Journal of Communication Disorders*, *65*, 43–53.

Lazar, R. M., Minzer, B., Antoniello, D., Festa, J. R., Krakauer, J. W., & Marshall, R. S. (2010). Improvement in aphasia scores after stroke is well predicted by initial severity. *Stroke*, *41*(7), 1485–1488.

Le Belle, J. E., Orozco, N. M., Paucar, A. A., Saxe, J. P., Mottahedeh, J., Pyle, A. D., Wu, H., & Kornblum, H. I. (2011). Proliferative neural stem cells have high endogenous ROS levels that regulate self-renewal and neurogenesis in a PI3K/Akt-dependant manner. *Cell Stem Cell*, *8*(1), 59–71.

Leonard, C., Rochon, E., & Laird, L. (2008). Treating naming impairments in aphasia: Findings from a phonological components analysis treatment. *Aphasiology*, *22*(9), 923–947.

Levelt, W. J. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, *42*(1–3), 1–22.

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

Li, S., Nie, E. H., Yin, Y., Benowitz, L. I., Tung, S., Vinters, H. V., Bahjat, F. R., Stenzel-Poore, M. P., Kawaguchi, R., Coppola, G., & others. (2015). GDF10 is a signal for axonal sprouting and functional recovery after stroke. *Nature Neuroscience*, *18*(12), 1737–1745.

Linebarger, M.C., Romania, J.F. (2007). Aphasia Therapy System. (U.S. Patent No. 7203649). U.S. Patent and Trademark Office.

Liu, H., Cordella, C., Ishwar, P., Betke, M., & Kiran, S. (2023). Consistent long-term practice leads to consistent improvement: Benefits of self-managed therapy for language and cognitive deficits using a digital therapeutic. *Frontiers in Digital Health*, *5*, 1095110.

Lorenz, A., & Ziegler, W. (2009). Semantic vs. Word-form specific techniques in anomia treatment: A multiple single-case study. *Journal of Neurolinguistics*, *22*(6), 515–537.

Madden, E. B., Robinson, R. M., & Kendall, D. L. (2017). Phonological treatment approaches for spoken word production in aphasia. *Seminars in Speech and Language*, *38*(01), 062–074.

Martin, N., & Dell, G. S. (2019). Maintenance versus transmission deficits: The effect of delay on naming performance in aphasia. *Frontiers in Human Neuroscience*, *13*, 406.

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, *41*(1), 3–29.

Meltzer, J. A., Allison, J. B., Steele, R. D., & Harvey, S. J. (2018). Computer-based treatment of poststroke language disorders: A non-inferiority study of telerehabilitation compared to in-person service delivery. *Aphasiology*, *32*(3), 290–311. https://doi.org/10.1080/02687038.2017.1355440

Middleton, E. L., Chen, Q., & Verkuilen, J. (2015). Friends and foes in the lexicon: Homophone naming in aphasia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 77.

Middleton, E. L., Rawson, K. A., & Verkuilen, J. (2019). Retrieval practice and spacing effects in multi-session treatment of naming impairment in aphasia. *Cortex*, *119*, 386–400.

Middleton, E. L., Schuchard, J., & Rawson, K. A. (2020). A review of the application of distributed practice principles to naming treatment in aphasia. *Topics in Language Disorders*, *40*(1), 36–53.

Middleton, E. L., Schwartz, M. F., Rawson, K. A., & Garvey, K. (2015). Test-Enhanced Learning Versus Errorless Learning in Aphasia Rehabilitation: Testing Competing Psychological Principles. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, *41*(4), 1253–1261.

Middleton, E. L., Schwartz, M. F., Rawson, K. A., Traut, H., & Verkuilen, J. (2016). Towards a Theory of Learning for Naming Rehabilitation: Retrieval Practice and Spacing Effects. *Journal of Speech, Language, and Hearing Research*, *59*, 1111–1122. https://doi.org/10.1044/2016_JSLHR-L-15-0303

Miozzo, M., Pulvermüller, F., & Hauk, O. (2015). Early parallel activation of semantics and phonology in picture naming: Evidence from a multiple linear regression MEG study. *Cerebral Cortex*, *25*(10), 3343–3355.

Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O. K., Coslett, H. B., & Schwartz, M. F. (2015). Neural organization of spoken language revealed by lesion–symptom mapping. *Nature Communications*, *6*(1), 6762.

Mostany, R., Chowdhury, T. G., Johnston, D. G., Portonovo, S. A., Carmichael, S. T., & Portera-Cailliau, C. (2010). Local hemodynamics dictate long-term dendritic plasticity in peri-infarct cortex. *Journal of Neuroscience*, *30*(42), 14116–14126.

Mozzafarian, D. (2016). Heart Disease and Stroke Statistics—2016 Update A Report From the American Heart Association. *Circulation*, *133*(4), e38–e48.

Nardo, D., Holland, R., Leff, A. P., Price, C. J., & Crinion, J. T. (2017). Less is more: Neural mechanisms underlying anomia treatment in chronic aphasic patients. *Brain*, *140*(11), 3039–3054.

National Aphasia Association. (2020). *2020 Aphasia Awareness Survey*. https://www.aphasia.org/wp-content/uploads/2020/05/Aphasia-2020-Survey-Results.pdf

National Aphasia Association. (2022). *2022 Aphasia Awareness Survey*. https://www.aphasia.org/2022-aphasia-awareness-survey/

Nichol, L., Hill, A. J., Wallace, S. J., Pitt, R., Baker, C., & Rodriguez, A. D. (2019). Self-management of aphasia: A scoping review. *Aphasiology*, *33*(8), 903–942.

Nichol, L., Pitt, R., Wallace, S. J., Rodriguez, A. D., & Hill, A. J. (2023). "There are endless areas that they can use it for": Speech-language pathologist perspectives of technology support for aphasia self-management. *Disability and Rehabilitation: Assistive Technology*, *18*(8), 1473–1488.

Nichol, L., Rodriguez, A. D., Pitt, R., Wallace, S. J., & Hill, A. J. (2023). "Self-management has to be the way of the future": Exploring the perspectives of speech-language pathologists who work with people with aphasia. *International Journal of Speech-Language Pathology*, *25*(2), 327–341.

Nichol, L., Wallace, S. J., Pitt, R., Rodriguez, A. D., Diong, Z. Z., & Hill, A. J. (2022). People with aphasia share their views on self-management and the role of technology to support self-management of aphasia. *Disability and Rehabilitation*, *44*(24), 7399–7412.

Nichol, L., Wallace, S. J., Pitt, R., Rodriguez, A. D., & Hill, A. J. (2022). Communication partner perspectives of aphasia self-management and the role of technology: An in-depth qualitative exploration. *Disability and Rehabilitation*, *44*(23), 7199–7216.

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

Nickels, L., & Best, W. (1996). Therapy for naming disorders (Part II): Specifics, surprises and suggestions. *Aphasiology*, *10*(2), 109–136.

Nozari, N., & Dell, G. S. (2013). How damaged brains repeat words: A computational approach. *Brain and Language*, *126*(3), 327–337.

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

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

Nozari, N., & Pinet, S. (2020). A critical review of the behavioral, neuroimaging, and electrophysiological studies of co-activation of representations during word production. *Journal of Neurolinguistics*, *53*, 100875.

Nunn, K., & Vallila-Rohter, S. (2023). Theory-driven treatment modifications: A discussion on meeting the linguistic, cognitive, and psychosocial needs of individual clients with aphasia. *Journal of Communication Disorders*, *103*. https://doi.org/10.1016/j.jcomdis.2023.106327

Nunn, K., Vallila-Rohter, S., & Middleton, E. L. (2023). Errorless, Errorful, and Retrieval Practice for Naming Treatment in Aphasia: A Scoping Review of Learning Mechanisms and Treatment Ingredients. *Journal of Speech, Language, and Hearing Research*, *66*(2), 668–687. https://doi.org/10.1044/2022_JSLHR-22-00251

Ovbiagele, B., Goldstein, L. B., Higashida, R. T., Howard, V. J., Johnston, S. C., Khavjou, O. A., Lackland, D. T., Lichtman, J. H., Mohl, S., Sacco, R. L., & others. (2013). Forecasting the future of stroke in the United States: A policy statement from the American Heart Association and American Stroke Association. *Stroke*, *44*(8), 2361–2375.

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M. C., Bowen, A., Latimer, N., Julious, S., Cross, E., Alshreef, A., Harrison, M., Bradley, E., Witts, H., & Chater, T. (2019). Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): A multicentre, single-blinded, randomised controlled trial. *Lancet Neurology*, *18*, 821–833.

Palmer, R., Dimairo, M., Latimer, N., Cross, E., Brady, M. C., Enderby, P., Bowen, A., Julious, S., Harrison, M., Alshreef, A., Bradley, E., Bhadhuri, A., Chater, T., Hughes, H., Witts, H., Herbert, E., & Cooper, C. (2020). Computerised speech and language therapy or attention control added to usual care for people with long-term post-stroke aphasia: The Big CACTUS three-arm RCT. *Health Technology Assessment*, *24*(19). https://doi.org/10.3310/hta24190

Parr, S. (2007). Living with Severe Aphasia: Tracking Social Exclusion. *Aphasiology*, *21*, 98–123.

Patterson, R., Robert, A., Berry, R., Cain, M., Iqbal, M., Code, C., Rochon, E., & Leonard, C. (2015). Raising public awareness about aphasia in southern Ontario, Canada: A survey. *International Journal of Speech-Language Pathology*, *17*(2), 121–126.

Pedersen, P. M., Stig Jørgensen, H., Nakayama, H., Raaschou, H. O., & Olsen, T. S. (1995). Aphasia in acute stroke: Incidence, determinants, and recovery. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *38*(4), 659–666.

Pedersen, P. M., Vinter, K., & Olsen, T. S. (2003). Aphasia after stroke: Type, severity and prognosis: The Copenhagen aphasia study. *Cerebrovascular Diseases*, *17*(1), 35–43.

Peeva, M. G., Guenther, F. H., Tourville, J. A., Nieto-Castanon, A., Anton, J.-L., Nazarian, B., & Alario, F.-X. (2010). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *Neuroimage*, *50*(2), 626–638.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 539.

Ralph, M. A. L., & Fillingham, J. K. (2007). The importance of memory and executive function in aphasia: Evidence from the treatment of anomia using errorless and errorful learning. In *Automaticity and control in language processing* (pp. 211–234). Psychology Press.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460.

Raymer, A. M., Beeson, P., Holland, A., Kendall, D., Maher, L. M., Martin, N., Murray, L., Rose, M., Thompson, C. K., Turkstra, L., & others. (2008). *Translational research in aphasia: From neuroscience to neurorehabilitation*.

Raymer, A., Thompson, C. K., Jacobs, B., & Le Grand, H. (1993). Phonological treatment of naming deficits in aphasia: Model-based generalization analysis. *Aphasiology*, *7*(1), 27–53.

Roberts, S., Bruce, R. M., Lim, L., Woodgate, H., Ledingham, K., Anderson, S., Lorca-Puls, D. L., Gajardo-Vidal, A., Leff, A. P., Hope, T. M., & others. (2022). Better long-term speech outcomes in stroke survivors who received early clinical speech and language therapy: What's driving recovery? *Neuropsychological Rehabilitation*, *32*(9), 2319–2341.

Robey, R. R. (1998). A meta-analysis of clinical outcomes in the treatment of aphasia. *Journal of Speech, Language, and Hearing Research*, *41*(1), 172–187.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1–3), 107–142.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*(3), 249–284.

Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: The WEAVER++/ARC model. *Cortex*, *59*.

Roelofs, A., & Ferreira, V. S. (2019). The architecture of speaking. *Human Language: From Genes and Brains to Behavior*, 35–50.

Rose, M., Ferguson, A., Power, E., Togher, L., & Worrall, L. (2013). Aphasia rehabilitation in Australia: Current practices, challenges and future directions. *International Journal of Speech-Language Pathology*, *16*(2), 169–180. https://doi.org/10.3109/17549507.2013.794474

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford publications.

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., & others. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, *105*(46), 18035–18040.

Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., & Weiller, C. (2006). Dynamics of language reorganization after stroke. *Brain*, *129*(6), 1371–1384.

Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., Kümmerer, D., Klöppel, S., Glauche, V., Lange, R., & others. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *Neuroimage*, *49*(4), 3187–3197.

Schuchard, J., & Middleton, E. L. (2018a). The Roles of Retrieval Practice Versus Errorless Learning in Strengthening Lexical Access in Aphasia. *Journal of Speech, Language, and Hearing Research*, *61*, 1700–1717.

Schuchard, J., & Middleton, E. L. (2018b). Word repetition and retrieval practice effects in aphasia: Evidence for use-dependent learning in lexical access. *Cognitive Neuropsychology*, *35*(5–6), 271–287.

Schuchard, J., Rawson, K. A., & Middleton, E. L. (2020). Effects of distributed practice and criterion level on word retrieval in aphasia. *Cognition*, *198*, 104216.

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*(2), 228–264.

Simmons-Mackie, N. (2017). *Aphasia in North America*. https://www.aphasiaaccess.org/white-papers/

Simmons-Mackie, N., Code, C., Armstrong, E., Stiegler, L., & Elman, R. (2002). What is aphasia? An international survey of public awareness of aphasia. *Aphasiology*, *16*, 837–848.

Simmons-Mackie, N., Damico, J. S., & Damico, H. L. (1999). A Qualitative Study of Feedback in Aphasia Treatment. *American Journal of Speech-Language Pathology*, *8*(3), 218–230.

Simmons-Mackie, N., Kagan, A., Christie, O., Huijbregts, M., McEwen, S., & Willems, J. (2007). Communicative access and decision making for people with aphasia: Implementing sustainable health care systems change. *Aphasiology*, *21*(1), 39–66.

Stefaniak, J. D., Halai, A. D., & Lambon Ralph, M. A. (2020). The neural and neurocomputational bases of recovery from post-stroke aphasia. *Nature Reviews Neurology*, *16*(1), 43–55.

StepByStep Aphasia Therapy. (n.d.). [Computer software]. https://aphasia-software.com/

Stockert, A., Kuemmerer, D., & Saur, D. (2016). Insights into early language recovery: From basic principles to practical applications. *Aphasiology*, *30*(5), 517–541.

Strijkers, K., Costa, A., & Pulvermüller, F. (2017). The cortical dynamics of speaking: Lexical and phonological knowledge simultaneously recruit the frontal and temporal cortex within 200 ms. *NeuroImage*, *163*, 206–219.

Sze, W. P., Hameau, S., Warren, J., & Best, W. (2020). Identifying the Components of a Successful Spoken Naming Therapy: A Meta-Analysis of Word-Finding Interventions for Adults with Aphasia. *Aphasiology*, *35*(1), 33–72.

The RELEASE Collaborators, Brady, M. C., Ali, M., VandenBerg, K., Williams, L. J., Williams, L. R., Abo, M., Becker, F., Bowen, A., Brandenburg, C., & others. (2022). Precision rehabilitation for aphasia by patient age, sex, aphasia severity, and time since stroke? A prespecified, systematic review-based, individual participant data, network, subgroup meta-analysis. *International Journal of Stroke*, *17*(10), 1067–1077.

Ueno, T., Saito, S., Rogers, T. T., & Ralph, M. A. L. (2011). Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*, *72*(2), 385–396.

van Hees, S., Angwin, A., McMahon, K., & Copland, D. (2013). A comparison of semantic feature analysis and phonological components analysis for the treatment of naming impairments in aphasia. *Neuropsychological Rehabilitation*, *23*(1), 102–132.

van Hees, S., McMahon, K., Angwin, A., de Zubicaray, G., & Copland, D. A. (2014). Neural activity associated with semantic versus phonological anomia treatments in aphasia. *Brain and Language*, *129*, 47–57.

Vestling, M., Tufvesson, B., & Iwarsson, S. (2003). Indicators for return to work after stroke and the importance of work for subjective well-being and life satisfaction. *Journal of Rehabilitation Medicine*, *35*(3), 127–131.

Wafa, H. A., Wolfe, C. D. A., Emmett, E., Roth, G. A., Johnson, C. O., & Wang, Y. (2020). Burden of Stroke in Europe. Thirty-Year Projections of Incidence, Prevalence, Deaths, and Disability-Adjusted Life Years. *Stroke*, *51*. https://doi.org/10.1161/STROKEAHA.120.029606

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

Warner, G., Packer, T., Villeneuve, M., Audulv, A., & Versnel, J. (2015). A systematic review of the effectiveness of stroke self-management programs for improving function and participation outcomes: Self-management programs for stroke survivors. *Disability and Rehabilitation*, *37*(23), 2141–2163.

Williams, S. E., & Freer, C. A. (1986). Aphasia: Its effect on marital relationships. *Archives of Physical Medicine and Rehabilitation*, *67*(4), 250–252.

Wilson, S. M., & Schneck, S. M. (2020). Neuroplasticity in post-stroke aphasia: A systematic review and meta-analysis of functional imaging studies of reorganization of language processing. *Neurobiology of Language*, *2*(1), 22–82.

Winstein, C., Stein, J., Arena, R., Bates, B., Cherney, L., & Cramer, S. (2016). Guidelines for adult stroke rehabilitation and recovery: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, *47*, e98–e169.

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

Yourganov, G., Smith, K. G., Fridriksson, J., & Rorden, C. (2015). Predicting aphasia type from brain damage measured with structural MRI. *Cortex*, *73*, 203–215.

Zheng, C., Lynch, L., & Taylor, N. (2016). Effect of computer therapy in aphasia: A systematic review. *Aphasiology*, *30*(2–3), 211–244. https://doi.org/10.1080/02687038.2014.996521

Chapter 2 - INFORMING THE DESIGN OF AT-HOME THERAPIES FOR APHASIA: A NEEDS FINDING ANALYSIS

ABSTRACT

Computer-based therapies for aphasia (CBATs) yield an immense potential to make cost-effective treatment available to anyone with a computer and an internet connection. Nonetheless, state-of-the-art CBATs present important limitations which stand in the way to their large-scale adoption. This study represents the first step towards the implementation of a CBAT that transcends these limitations. Specifically, it sets out to: a) provide a solid understanding of the main unmet needs experienced by the end-users of such a product, namely people with aphasia and therapists; b) gather their opinions about the desirable features that such technologies should have; c) gain relevant knowledge on the kind of barriers and facilitators that they experience when interacting with modern devices; d) gather their views on the adequacy of the solutions we are proposing to improve upon state-of-the-art computer-based therapies: gamifying speech and language therapy to make it more engaging, and integrating state-of-the-art automatic speech-recognition (ASR) to provide users with real-time feedback on the accuracy of their naming performance. To achieve these objectives, we combined background research with interviews involving people with aphasia and therapists. The conversations were analyzed using thematic analysis.

Results confirmed the urgent need for more accessible therapy services and revealed that the solutions we proposed to improve upon currently available computer-based therapies were endorsed by both people with aphasia and therapists. Also, this study provided a wealth of information that is relevant to design a prototype of CBAT with high patient and clinician

usability. As such, besides the relevance to this dissertation, this study is of interest to anyone involved in the design of CBATs.

INTRODUCTION

The previous chapter provides an overview of the state-of-the-art of CBATs and identifies two potential limitations that might severely limit their efficacy: the monotonous and repetitive nature of most popular exercises for language therapy and the lack of a system that can provide feedback on how well people speak. Two fundamental innovations were proposed to address these limitations: (a) embracing ideas from the field of serious gaming to make treatment more engaging, promoting motivation and resilience; (b) harnessing recent improvements on state-of-the-art ASR to give users feedback on the accuracy of target words/utterances.

Before moving on with building a prototype that can be used by the target users, we set out to engage with the first two steps of the customer discovery phase recommended for new product development (Blank & Dorf, 2012): first test the problem, namely ensure that its diagnosis is accurate, and that it is worth solving; second, test the solution – once the problem is accurately defined, assess what the target users think about the adequacy of the proposed solution. This study was conceived to implement these steps and to gather information that is relevant to inform the design of a minimum viable product (MVP, typically the earliest version of a prototype integrating the set of features that are *necessary* and *sufficient* for it to address the needs that it was conceived to) that can be tested with end users.

To achieve the above-mentioned objectives, we integrated relevant information gathered from the literature with semi-structured interviews conducted with people with aphasia and speech-language therapists. Specifically, the questions that guided our research and our conversations with study participants were:

1. What are the most pressing unmet needs that people with aphasia experience, and how can technology-based solutions meet them?
2. What, according to people with aphasia and therapists, are the main strengths and weaknesses of currently available CBATs?
3. How does aphasia affect the interaction with modern technology devices?
4. What are the attitudes of people with aphasia towards modern technology devices and video games?
5. Are there common interests and pastimes that can inform how to design an MVP?

The following section presents an overview of the literature addressing these questions.

The Main Unmet Needs of People with Aphasia

The field of aphasia rehabilitation has recently seen an influx of new ideas on the most appropriate ways to deliver treatment. In part, the discussion has been inspired by the mounting evidence suggesting that improvements in language symptoms induced by speech-language therapy do not necessarily translate into meaningful improvements in the quality of life (Chapey, 2008). In fact, even after language symptoms have improved, people with aphasia commonly complain about the pervasive feelings of social isolation, loneliness, loss of autonomy, and of being disengaged from life (Simmons-Mackie, 2008). This important limitation of speech-language therapy is commonly explained by the fact that, traditionally, it has adopted an impairment-based approach wherein communication abilities are assessed and treated in isolation from contexts that might instead be relevant to clients (i.e. holding a conversation with family members or friends, to a level that might enable them to put their talents and ambitions in the service of a community, company, or society writ large; Chapey et al., 2000). To counter this situation, there has been a strong call for expanding the scope of therapy from solely adopting an impairment-based approach to one where real-life contexts are the primary guide for the type of activities and content that clients engage with (Chapey et al., 2000; Simmons-Mackie, 2017; Worrall et al., 2011).

This re-evaluation of the means employed by therapy to improve the quality of life of people with aphasia was accompanied by efforts to identify and address the ways in which the service delivery model fails to meet their needs throughout the continuum of care. These efforts have revealed that, indeed, people with aphasia commonly report a specific set of needs that are not adequately met. The most frequently reported goals are: returning to work and to a lifestyle that is as close as possible to the one that was lived before the onset of aphasia (Berg et al., 2017; Rohde et al., 2012; Worrall et al., 2011); gaining the ability to communicate not only basic needs, but also opinions, ideas and to have normal conversations (Berg et al., 2017; Pettit et al., 2017; Rohde et al., 2012; Wallace et al., 2017; Worrall et al., 2011); learning alternative ways to communicate (i.e. singing along with speaking) (Rohde et al., 2012); engaging in social activities (Pettit et al., 2017; Rohde et al., 2012; Wallace et al., 2017; Worrall et al., 2011) and being helpful to others (Worrall et al., 2011). People also expressed the desire to engage with more speech therapy; to receive more information about aphasia, strokes, and available services (Wallace et al., 2017; Worrall et al., 2011) and ways to support people with communication difficulties (Code & Heron, 2003), and to make this information available to their family members and to society at large; to learn ways to better cope with aphasia and to maximize well-being (Pettit et al., 2017; Wallace et al., 2017) and autonomy (Pettit et al., 2017; Worrall et al., 2011); to be treated with respect and dignity (Pettit et al., 2017; Worrall et al., 2011).

Aphasia also has a significant impact on the psychological well-being of family members and caregivers, who play a crucial role in meeting the physical and psychological needs of a person with aphasia. Therefore, it is important to ensure that their needs are carefully understood and addressed (Draper & Brocklehurst, 2007; Simmons-Mackie, 2017). Howe and colleagues (2012) investigated this topic by interviewing people with aphasia and their family members, showing that caregivers do value being involved in the process of rehabilitation, especially when it comes to setting goals for therapy; they value learning how to communicate effectively with their partner; and they value learning how to promote their well-being and cope with the emotional stress resulting from their new responsibilities. In addition, Howe et al. reported caregivers' desire to be provided with emotional support by professionals or other people that have experienced a similar situation, emphasizing the need to always be given hope and to keep a positive outlook on things. The views of both people with aphasia and family members that were discussed above are in line with those of therapists on the same issues (Wallace et al., 2017).

How Does Aphasia Affect the Interaction with Technology?

Technologies have long been considered as potential tools to assist people with aphasia with communication (for instance Alternative Augmentative Communication [AAC] devices) or to promote recovery of communication skills (i.e. CBATs). Nonetheless, as aphasia affects physical (motor), cognitive (attention, memory, perception), and communicative faculties, PwA might encounter a variety of barriers when using modern devices that are easy to master for healthy users (Kaufman et al., 2019). Therefore, to fully realize the potential of modern technology in helping people with aphasia, it is essential to have a detailed understanding of, on the one hand, the kind of barriers that people with aphasia face when using modern devices, and, on the other hand, the aspects that facilitate interaction with hardware and software.

Barriers to Use of Technologies

The most intuitive guidelines to help minimize barriers to use of technology relate to the physical properties of the user interface. Devices that have small screens and buttons, such as mobile phones, or applications that require users to perform smooth and continuous movements on the touchscreen, such as scrollbars on menus, pose a barrier to use for people with aphasia who frequently have a reduced ability to perform fine movements with their dominant hand or that have difficulties reading (Brandeburg et al., 2013, 2017; Greig et al., 2008; Hill & Breslin, 2016; Palmer et al., 2013).

As people with aphasia might also present with deficits in memory, attention, and executive functioning — all cognitive abilities that play an important role in decision-making (Cahana-Amitay & Albert, 2015; Kim et al., 2020) — features of modern devices that tax these abilities are likely to pose a barrier to use to people with aphasia. Examples are buttons with multiple functions, screens filled with icons, menus offering several options, tasks requiring too many steps (Greig et al., 2008). Similarly, the frequent intrusion of popups and notifications, or the unexpected changes in layout that comes with software updates, could be annoying to many users (Brandeburg et al., 2017).

People with aphasia typically complain about CBATs that cannot adequately keep track of their changing level of competence and adapt to it. This design feature, together with a lack of variety of stimuli and tasks, would make such applications insufficiently challenging and stimulating to the users (Finch & Hill, 2014a; Hill & Breslin, 2016; Palmer et al., 2013).

Aspects that Facilitate the Use of Technologies

Among the aspects that facilitate the use of modern devices and applications, people with aphasia have indicated that touchscreens, when used to control an application via pointing, rather than swiping, are more user-friendly than a mouse or touchpad (Brandeburg et al., 2017; Finch & Hill, 2014a). They also discussed the importance of having access to a wide variety of tasks and stimuli that are relevant to them (Hill & Breslin, 2016; Palmer et al., 2013), that provide the right level of challenge, and that sustain their motivation (Finch & Hill, 2014a; Hill & Breslin, 2016; Palmer et al., 2013). Further, people with aphasia noted the importance of receiving instantaneous feedback and accessing metrics that keep track of their performance (Finch & Hill, 2014a; Galliers et al., 2011; Hill & Breslin, 2016).

Attitudes of People with Aphasia Towards Technology

Another aspect that is important to account for when designing CBATs is people's attitude towards the use of technologies in general, and towards a given CBAT more specifically. As discussed by Chen and Bode (Chen & Bode, 2011), therapists give serious consideration to the attitudes of people with aphasia towards CBATs when deciding whether to adopt them to facilitate therapy. The literature suggests that there are two important factors that influence such decisions: a person's familiarity with the devices that are compatible with their CBAT of choice, and the availability of adequate support to the user. Finch and Hill (2014a) interviewed 34 people with aphasia to investigate how aphasia affects the attitudes towards the use of computers. One of the main findings of this study is that familiarity with a device strongly predicts whether it will be used after the onset of aphasia. Most participants reported using computers regularly before the onset of aphasia - prevalently for work - and they continued to do so after losing their jobs by repurposing computers into tools used mostly for leisure activities, establishing and maintaining relationships with other people via the web, or doing language therapy. This finding is corroborated by other studies as well, wherein a lack of familiarity with technology was found to be a barrier to adoption (Brandeburg et al., 2017; Greig et al., 2008). Also, in Finch and Hill, the familiarity of people with aphasia with computers likely explained the widespread awareness of the ways technologies could potentially benefit them, with the most frequently appreciated benefit being the opportunity to practice when it worked best for them, at their own pace, and independently (Finch & Hill, 2014; Hill & Breslin, 2016).

Another important insight that emerged from these studies is the need to receive adequate and timely assistance whenever problems arise, especially for problems that are technical in nature (Brandeburg et al., 2013, 2017; Finch & Hill, 2014a; Greig et al., 2008; Hill & Breslin, 2016; Palmer et al., 2013). This is particularly important in light of the fact that people with aphasia often report or demonstrate difficulties learning a program or remembering how to use it (Gamberini et al., 2006). For these reasons, people with aphasia deemed the availability of adequate support - including aphasia-friendly manuals (Brandeburg et al., 2017; Greig et al., 2008) or having speech-therapists or volunteers available in person or via the phone (Brandeburg et al., 2017; Hill & Breslin, 2016; Palmer et al., 2013) — essential to adopt CBATs for aphasia therapy.

Therapists' Views on the Desirable Features of CBATs

The other group of end-users to be considered when designing CBATs are therapists. To better understand their needs and opinions, Chen and Bode conducted a survey among occupational therapists, physical therapists and speech-language therapists (Chen & Bode, 2011). This work revealed the decisive factors that influence a therapist's decision about adopting a new technology to assist with work: effectiveness; design (i.e. availability of timely and fine-grained feedback on performance, availability of support, ease of use); a client's interest in using the technology, which was strongly linked to the attainment of tangible outcomes (Chen & Bode, 2011). It was particularly important to therapists that their clients' families were willing to assist their loved one with using the technologies designed to supplement face-to-face therapy, when needed. These findings echo (at least in part) the opinions provided directly by people with aphasia (described in the previous section) and are corroborated and expanded upon by the work of Swales and colleagues (2016). The results of this work supplement those of Chen and Bode in three important ways. First, they highlight the importance of calibrating activities to each person's level of competence, interests, and goals. For example, stimuli should be salient and culturally appropriate, and the program should be able to adjust the level of difficulty of a given task based on a person's progress. Second, language therapists stressed the importance of designing applications that are easily accessible to anyone with aphasia and to clinicians as well, with an emphasis placed on making them available across different types of devices (i.e. computer, tablet, smartphone) and operating systems. Along with this, CBATs should require little training to be used effectively. Third, speech-language therapists discussed the importance of getting access to a number of metrics that are relevant to properly monitor a client remotely. Examples of such metrics are the amount of time spent on each task, on any given session; the type of stimuli that clients engaged with, along with their performance throughout all attempts to perform a task; the timing, frequency and type of cues that were used. Fourth, therapists also discussed the importance of getting access to the raw data to conduct their own analysis. Finally, most therapists look for CBATs that offer a wide variety of tasks, preferably encompassing both impairment-based and function-based therapy.

Attitudes of Healthy Seniors Towards Video Games

Despite the abundance of literature on the use of technologies by people with aphasia, to the best of our knowledge there are no studies investigating their motives to engage with video games, their preferences regarding game genres, or their views about what makes games optimal from a usability perspective. Nonetheless, these issues have been investigated at length in healthy seniors. As this group of people shares the most with people with aphasia in terms of cultural background, familiarity and attitude towards technology, the literature on game preferences of healthy seniors can provide valuable insights. For example, Salmon and colleagues (Salmon et al., 2017) conducted a survey to investigate how preferences about video games change across the lifespan, exploring a range of aspects such as desirable features of video games, devices that were most frequently used, genre preferences, and common social contexts for gaming. Most seniors emphasized the importance for games to be easy to learn and play. The devices most commonly used to play were desktop computers and laptops (Salmon et al., 2017), probably due to the fact that those were the only available technologies when seniors first got exposed to modern devices. Other studies also reported that seniors really valued video games that provide adequate support with learning (Doroudian, 2019; Kaufman et al., 2019).

Another aspect that has been explored is the type of motives driving older adults to engage with video games. In this regard, one of the main findings is that seniors often expect to get some value out of the time spent playing video games other than mere fun (Kaufman et al., 2019), be it the possibility to maintain or improve cognitive or physical function (De Schutter & Malliet, 2014; Diaz-Orueta et al., 2012; Nap et al., 2009), to improve emotional well-being by doing activities that distract the user from the problems of daily life and are perceived as relaxing (De Schutter & Malliet, 2014; Diaz-Orueta et al., 2012; Nap et al., 2009), or to gain new knowledge about topics that are relevant to their interests (De Schutter & Malliet, 2014; Diaz-Orueta et al., 2012; Salmon et al., 2017). Another common motive was the need to nourish social relationships, and the perceived opportunity to do that by playing together with other people (Diaz-Orueta et al., 2012; Kaufman et al., 2019; Seah et al., 2018), especially family members and friends (Kaufman et al., 2019). Importantly, some people specified the desire to engage with multiplayer gaming in person, rather than remotely on the internet. Conversely, in other studies seniors reported that most frequently they played video games alone (Salmon et al., 2017) or that, despite placing a great importance on social relationships, they preferred to cultivate them outside of time spent playing video games (Nap et al., 2009).

A recent review of the state of the art of video games designed to be used by older adults highlights the need to ensure that the game engages users with content they are familiar with and that is relevant to them, that provides real-time feedback on their performance during gameplay, and a summary of their learning outcomes/performance at the end of each playing session (Doroudian, 2019; Kaufman et al., 2019). Games should be designed to flexibly adapt the level of difficulty to users' evolving skills throughout gameplay, providing interesting challenges and fun (Salmon et al., 2017). All the motives mentioned above echo those expressed by people with aphasia regarding their attitudes towards use of technologies, in particular about CBATs (Finch & Hill, 2014).

There was also common ground among seniors about game genre preferences. Studies have investigated seniors' preferences regarding non-digital games, finding that cards, board games and puzzles tend to be the most popular among this group of people (Diaz-Orueta et al., 2012; Kaufman et al., 2019; Mortenson et al., 2017; Seah et al., 2018). Similarly, popular digital games included casual games (Cota & Ishitani, 2015; Kaufman et al., 2019; Schutter, 2011) (i.e. Solitaire, Patience and Bejeweled), along with strategy and educational games (Salmon et al., 2017). In light of these findings, Diaz-Orueta recommends designing games around common interests among seniors such as jigsaw puzzles, mental challenges, physical activities, topics centered around culture and arts, and daily life skills (i.e. cooking) (Diaz-Orueta et al., 2012).

What is Missing?

All in all, the literature provides rich information that informs on how to best design a software-based therapy that can serve the needs of people with aphasia. Previous sections reviewed studies discussing what the most common goals for people with aphasia and their family members are; they also present findings from studies that analyzed what type of barriers and facilitators people with aphasia do typically experience when interacting with technology, and from studies that investigated the opinions of therapists on the desirable features that a video game based CBAT should present. Related to the latter point, as we could not find any studies that addressed these questions with PwA directly, we reviewed findings on the preferences of healthy seniors about the use of video games. Naturally, while there are reasons to think that most insights are valuable to inform the design of CBATs, it is also likely that people with aphasia hold unique views that need to be considered to adequately tailor video games to them. Also, although the literature provides a wealth of information on the barriers and facilitators experienced by people with aphasia when using technology, it is generally vague about what makes a given device easy or hard to use. Given the plethora of devices that can nowadays deliver therapy at home, we wanted to look deeper into whether there are any specific devices that people with aphasia prefer to use, and why.

The Present Study

To supplement the information gathered from the literature, this study was conceived to engage directly with the end-users of CBATs, people with aphasia and therapists. We did that by using semi-structured interviews designed to answer the following questions. For some, we had hypotheses as to the outcomes whereas others were purely exploratory:

1. Confirm or disconfirm our hypotheses about what the most common unmet needs that can be addressed by a software-based solution are.

    H1 - Most people with aphasia experience a lack of easy and affordable access to therapists.

2. Explore whether there are any unmet needs that have not been documented in the literature.

3. Validating our ideas on how to improve upon the currently available CBATs by gathering the opinions of people with aphasia and therapists about the two innovations we plan to introduce: ASR and game-based tasks.

    H2 - People with aphasia and therapists consider the use of an ASR and game mechanics as potentially effective innovations.

4. Gather insights about common interests and pastimes of people with aphasia that can inform the design of engaging tasks, to understand how to make tasks more engaging to people with aphasia.

5. Gain a detailed understanding about what constitutes a barrier or facilitator to the use of modern technologies for people with aphasia and therapists.

H3 - Games will appeal to people with aphasia if they incorporate common themes of interest and have gameplay mechanics that commonly appeal to older adult

METHODS

The procedures of this study were approved by Dalhousie University Research Ethics Board, and all participants provided written informed consent to take part in this study. All research staff who interacted with people with aphasia were trained beforehand by certified speech language pathologists. This was done to ensure that they were familiar with appropriate ways of facilitating communication with people with aphasia, and sensitive to the special needs of this population. The consent form was specifically designed for readability and comprehension by people with mild to moderate aphasia. For example, to account for the fact that people with aphasia may not be able to answer open-ended questions easily, questions were posed in a way that a yes/no answer could be provided to determine whether the participant has understood the information about the study (e.g., "Will participating in this research study benefit you directly?"). To assess participants' capacity to provide informed consent for this study, the person conducting the consent discussion asked participants questions about the scope of the study, its procedures, and the benefits of taking part in it.

The study was explained verbally to potential participants and, if they desired, a communication partner such as a spouse whom the potential participant requested to be involved in the consent process. The potential participant was then asked to read the consent form and ask any questions they may have had about it or the study; if a participant had impaired reading ability, the research assistant read the consent form to them. After reading the consent form, participants were reminded that they could ask the researcher any questions about the procedures.

Participants

To participate in this study, people had to satisfy the following inclusion criteria: having had a single incident of brain damage causing the aphasia; having had aphasia for at least 6 months; being a fluent English speaker before the onset of aphasia; and having normal or corrected-to-normal vision and hearing. People with severe aphasia could not take part in this study, since such a level of impairment would have made the conversation very difficult to complete within a reasonable amount of time. Other exclusion criteria were the presence of primary progressive aphasia, along with suffering from cognitive impairment. Nine people with aphasia (3 females), and 3 therapists (all female) took part in this study. Participants from both groups were all residents of the province of Nova Scotia (Canada). All therapists served in rural communities across the province. Data from one participant with aphasia could not be used as the interview, which was conducted via Skype, failed to record.

Materials

The semi-structured interviews conducted with people with aphasia were designed to target the following topics: consequences of aphasia on everyday life; past and present experience with language therapy; barriers and facilitators with the use of modern technologies;

desirable and undesirable features of video games and of computer-based therapies designed to help with aphasia; interests and pastimes. In contrast, the semi-structured interviews conducted with therapists targeted the following topics: most important unmet needs of people with aphasia; personal approach to the treatment of aphasia; opinions about the potential role of ASR and gamified tasks in the context of at-home treatment of aphasia; common goals that individuals with aphasia aim to achieve through rehabilitation.

Procedure

Whenever possible, the interviews with people with aphasia were conducted in person, otherwise via Skype, and lasted less than 1 hour. All interviews with therapists were conducted via Skype. In-person interviews with people with aphasia were conducted in a soundproof room, with no windows, and were video-recorded using a video camera set to video record the conversations. The interviews were guided by questions selected to ensure that a given topic of interest was adequately covered during the conversation. Each question was used as a starting point to explore a specific aspect of a topic, but the interviewer was encouraged to inquire further whenever participants' responses suggested that more relevant information could be obtained.

Before the beginning of the in-person interviews, the interviewer asked participants to sit whenever it pleased them around a table. Then, after having ensured that participants' arms and face were adequately captured by the camera, the interviewer started the video recording, following the same procedure with participants who took part in the study remotely.

Table 2. 1 - Demographics

Subject	Age	Sex	Time Since Stroke	Type of Aphasia	Education	Language		Experience with Therapy	
						L1	L2	Past	Present
S1	59	F	14y	Receptive; anomia	Diploma (X-Ray technologist)	English	·	8y of 1-to-1 2d/w	5y group therapy 3d/w
S2	67	M	9y	Expressive; reading deficits	BSc. (Civil Eng.)	English	·	8y of 1-to-1 1h/w	Group + 1-to-1 2h/w
S3	61	M	8y	Expressive	Bachelor (Insurance)	English	French	2y of 1-to-1 2d/w	-
S4	61	M	4y	Mild expressive	High School	English	·	ND	ND
S5	51	F	7y	Expressive	Bachelor (Education)	English	·	1-to-1 1d/month	-
S6	-	M	4y	-	Radio College of Canada	English	·	1-to-1 2d/w	Group therapy
S7	69	M	7y	Expressive; reading and writing	High School	English	·	1.5y of 1-to-1, 1h/w	-
S8	59	F	20y	Receptive; expressive	Diploma (Hairdresser)	French	English	3y 1-to-1	1-to-1, 2d/month
S9	-	M	4y	Mild expressive	-	English	·	-	1-to-1

Data Analysis

Thematic Analysis vs. Qualitative Analysis

Two methods that are commonly employed to analyze the content of conversations are *thematic analysis* (Braun & Clarke, 2006) and *qualitative content analysis* (Graneheim & Lundman, 2004). These methods vastly overlap in the scope and procedures used to analyze the data (Vaismoradi et al., 2013), as they both aim to guide researchers in the process of a) identifying the themes that recur throughout textual/conversational corpora and that are relevant to the research questions of the study; b) analyzing and interpreting themes in relation to the topic under discussion; c) reporting results (Braun & Clarke, 2006; Graneheim & Lundman, 2004). Due to these similarities in scope, the terms thematic analysis and qualitative content analysis are often used interchangeably or are otherwise conflated (Vaismoradi et al., 2013). Nonetheless, these methods differ in ways that are important to underline, especially in the context of this study. The most relevant, albeit nuanced, difference regards how the two methods define the concept of "theme". While in qualitative content analysis the importance of a theme is strictly associated with its prevalence throughout a text or conversation (measured by the frequency of codes that refer to it; Graneheim & Lundman, 2004; Morgan, 1993), in thematic analysis it is based on the extent to which a theme contains information that is relevant to address a specific research question (Braun & Clarke, 2006). This distinction is important because some of the research questions of this study were not necessarily best addressed by analyzing and interpreting the themes that occurred most frequently throughout a given topic of discussion. For example, when it comes to gathering information on how to best design software that is aphasia-friendly, important insights might come from just one person who has had the opportunity to test technologies that most of their peers have never used. For this reason, we employed thematic analysis to analyze the conversations.

Phases of Thematic Analysis

All conversations were first transcribed by research volunteers. When analyzing interviews for which a video was also available, volunteers were instructed to pay attention to gestures or facial expressions of participants that could convey additional meaning, and to take note of those.

Data analysis was conducted entirely by the author of this thesis. Using transcriptions as a starting point, the analysis followed the 6 stages described by Braun and Clarke (2006) (see Appendix A for a worked example). The first stage entails familiarizing with the entire dataset to get a first impression of the possible themes that might be present in it. This is done by reading all conversations until the reader has acquired a good grasp on the main ideas that recur throughout the dataset. Next, the dataset is analyzed more thoroughly to identify and label the main topics that were discussed. The parts of the conversation covering a given topic are then scanned to generate initial codes. This phase entails searching for passages (i.e. words, phrases or sentences) that express a coherent idea. Such passages are defined as meaning units (Graneheim & Lundman, 2004). Each meaning unit is then pruned from all redundant and irrelevant words or phrases, such that the core idea can be preserved intact while being highlighted in the condensed meaning unit (Graneheim & Lundman, 2004).

Once this step is completed, each condensed meaning unit is assigned a code according to the central idea that it expresses[1]. The third phase aims to identify potential themes by analyzing the relationship among the codes: those that refer to related ideas are grouped together, and the theme that threads through all codes and ideas is defined. Finally, the last stage entails a quality check to ensure that all themes do indeed represent meaningful abstractions of meaning units that can be used to support data interpretation. This assessment is guided by two criteria: one is that the ideas expressed in the meaning units underlying a given theme are coherent, and that the theme does adequately reflect them; the other is that there should be clear differences in the messages underpinning different themes (Braun & Clarke, 2006; Patton, 1990).

Data from people with aphasia and therapists were analyzed separately, but are presented together whenever a theme emerged in the conversation with members of both groups.

RESULTS

Results of the thematic analysis are organized following the order in which topics were covered during the conversations. Each topic is broken down into themes and sub-themes. Throughout this section, quotes from participants are presented to better elucidate a given point. The person associated with a quote will be referred to as Sx and Tx for people with aphasia and therapists, respectively (the letter 'x' is a number that reflects the order in which the participant from a group enrolled in the study).

---

[1] The terms *meaning unit* and *condensed meaning unit* are borrowed from the work of Graneheim and Lundman 2004, on qualitative analysis. Here, the only difference with the procedure proposed by Braun and Clarke, 2006, to conduct thematic analysis is that, when explaining how to generate the initial codes, not only an instruction on what to do is provided (i.e. assign a label to relevant features of the data in a systematic fashion), but also a description of how to effectively generate a label (i.e. distill the core idea contained in a given passage, starting from a generic meaning unit to a condensed meaning unit etc.). I decided to adopt this idea taken from qualitative analysis because I found it helpful for the sake of data analysis, while not conflicting with any of the tenets of thematic analysis.

Consequences of Aphasia

When discussing the consequences of aphasia, three themes were identified: long-lasting physical, cognitive and linguistic deficits; difficulties holding a job; changes in lifestyle.

Theme 1: Long-lasting Physical, Cognitive and Linguistic Deficits

Two sub themes were identified as part of this theme: difficulties with communication; strategies used to cope with aphasia (see Table A.2 at Appendix A for the relevant meaning units). All participants reported issues associated with the well-known symptoms of aphasia (i.e. struggling with a particular aspect of communication such as speaking, understanding etc.). One person (S1) expressed frustration for the general lack of public awareness on the nature of aphasia. As she presented with only mild anomia and receptive aphasia, it was not obvious to other people that she had a neurological condition that affected her ability to understand ["...you know, even just white noise is a challenge too because I may not understand what people are saying for that reason. Most people do not really, they don't know what aphasia is. So if I make a mistake, I hesitate now to tell them because some people will say, "well, you look fine".]" To S1, a lack of awareness about aphasia leads people to be insensitive to the struggles and needs of a person with aphasia.

Among the strategies employed to help getting the message across, participants mentioned using writing to provide cues about the intended message (i.e. the first letter of the target word (S2), or a list of keywords to guide the conversation (S9)), and the utility of the speech-to-text function as a way to easily send messages/emails (S5).

Theme 2: Difficulty Holding a Job

Four participants reported having lost their job as a result of aphasia. None of them was able to be employed again, even when symptoms progressed from severe/moderate to mild. This was the case of S1, who stated that "The problem that I have still, it's better than it was before, is understanding what people say. So really, an employer wants you to know what people are saying. I have tried several times to go back to work, but it wasn't very successful" (see Table A.2). According to T3, holding a job is one of the biggest challenges experienced by individuals with aphasia, even after having completed therapy and having improved their symptoms to a moderate or mild degree.

Theme 3: Changes in Lifestyle After the Onset of Aphasia

Two sub themes were identified within the theme 'Changes in lifestyle after the onset of aphasia': disruption of lifestyle due to aphasia, and strategies adopted by people with aphasia to cope with it. The first subtheme was exemplified by S1 reporting that she had to quit one of her favorite hobbies (curling) due to aphasia, while S3 reported having had to let go with his habit of reading intensely. S3, who could not retain his work or find a new one due to aphasia, said that he managed to find new ways to contribute to the well-being of his family by taking on house-chores. He also adopted healthier dietary choices and quitted smoking to prevent the occurrences of other cardiovascular issues. On the other hand, T3 and T2 said that one of the things that disrupt a person's lifestyle the most is the strong reduction of their social network, which contributes to the loss of self-esteem and high incidence of depression, an illness which is very frequent among people with aphasia (Hilari, 2011; Hilari et al., 2010; Kauhanen et al., 2000; Morris et al., 2017; Ross et al., 2010). Poor mental health conditions, in turn, further prevent people from returning to activities they used to engage with prior to the onset of their aphasia, or to take the initiative to cultivate novel interests.

Strengths and Weaknesses of Current Therapy Services

When investigating the main unmet needs of people with aphasia, we identified six themes: barrier to access therapy services; inadequate service delivery model; factors whose importance is underestimated; opinions of people with aphasia about positive aspects of therapy; opinions of people with aphasia about aspects to improve in therapy; opinions of therapists about most appropriate way to deliver treatment.

Theme 1: Barriers to Access Therapy

This theme was further divided into six subthemes (see Table A.3 for relevant meaning units): a long-waiting list to enroll in publicly-funded therapy; costs; lack of staff; poor assistance with navigating therapy options; lack of engagement; mobility issues.

Four participants reported having had problems with receiving adequate therapy after being discharged from the hospital, mostly due to excessively long waiting-lists. For example, S2's partner stated that "With the rehab we had to wait at least 6 months and that's not good. We left the hospital and we had to wait 6 months. Fortunately, we have coverage for private specialist; if he had waited 6 months doing nothing, I think doing speech not much help, you know from the government." In addition to it, S5 said that while she was hospitalized she received poor information on the steps she had to take to enroll in publicly-funded therapy: "I was in the hospital for about three months rehab... and I was warned that [she had] to set up services before you go home... I was in the hospital for about a week longer than I needed because the services were not set up for me, if I had gone home a week before I would have gone to the bottom of the list, do you know what I mean?... I stayed the extra week and got the services all lined up, I had no idea. My sister was a nurse and she said "wait for the services to be lined up or you're screwed".

Participants also complained about the costs required to engage with therapy of adequate intensity which, given the relatively short period that is covered by the government (in Canada, 5-6 weeks; Canadian Stroke Network, 2011), can only be offered by private therapists. S5 said that she was open to adopting a CBAT to help with her aphasia, but that prices were too high for her. Five participants reported about being enrolled in a private therapy program at the moment of the interview.

Finally, S6 discussed his difficulties and frustration with finding a therapist that could help him push his limits and improve his communication skills past a certain point. These difficulties were probably exacerbated by the fact that he lived in a rural and underserved area. This point was further stressed by T1 and T3 (see Table A.5 for relevant meaning units). As people living in rural areas rely almost exclusively on their caregivers for transportation, visits have to be planned far in advance to accommodate caregivers' schedules, placing additional constraints on the feasibility of planning meetings. People with aphasia who do not enjoy the support of family members or caregivers usually have no means to access the much-needed services (T1). Life is sometimes not much easier for people living in urban centers where speech-therapists operate. Although here people are less affected by transportation issues, to some, having to pay for a cab, a bus, or a parking lot might represent a barrier to attending therapy (T3). Related to it, T1 noted that, in her experience, individuals whose ability to walk with ease was impaired by stroke did typically experience more difficulties attending therapy. According to her, motor impairments not only represent an obstacle as they make it harder to move, but also because the prospect of dealing with the difficulties of moving inhibits people's drive to leave their house to attend a clinic.

An important sub theme that was identified when analyzing this topic was clients' lack of understanding about the role they should embrace in their recovery. According to T1, most of her clients do not realize the importance of actively carving out some time to exercise on their own, seeking out opportunities to practice communication skills with other people, etc. Instead, most people limit themselves to following instructions during therapy sessions. Relatedly, many of her clients show difficulties with accepting the idea that the therapist can help them improve up until a certain point, after which they need to be proactive in creating the conditions for themselves to flourish. The predominant mindset is one of overreliance on the therapist as a necessary support for them to sustain their communication skills, even if after a certain point little progress is made. This phenomenon was echoed by T3, who said that "...when you let them know that "I think we've kind of reached the end", they start to panic, and they can't sort of see themselves being independent communicators without having a lifeline connected to [them], even though you know that they would be able to manage".

Aside from the difficulties experienced with enrolling in therapy, T3 said that when enrolling in multiple programs targeting different disabilities (i.e. motor, cognitive, in addition to language), some people find it hard to adequately adhere to all of them.

Theme 2: Inadequate Service Delivery Model

One recurrent theme that was discussed by all the therapists was the inadequacy of the service delivery model at meeting patients' needs. One of the limitations that were discussed was the lack of a protocol to establish a connection with the families of individuals with aphasia immediately after a stroke occurred. This would have the intent to provide family members with information about aphasia, rehabilitation, available resources, training on how to enable and facilitate communication with a person with aphasia (T1, Table A.5), and, not least, to gather their input on how to tailor therapy to best suit the needs of their loved one. As T2 said, a failure to engage family members in this process limits the effectiveness of therapy.

Another important limitation of the service delivery model is the lack of organized community-support system to help people with aphasia along their recovery path once they are discharged from inpatient rehabilitation (T2): "... Like therapy — maintenance I guess you could call it, so they are at a point where there is a plateau and a discharge from clinical programs [is] necessary but there are functional goals that can still be achieved in a supportive community environment, and whether that same kind of accessible support group or whether that's community therapy in house support." In addition, a chronic lack of staff specialized in delivering intensive therapy was considered to be among the main reasons for the long waiting-lists to enroll in inpatient treatment.

Theme 3: Factors whose Importance is Currently Underestimated

According to T1, there are two factors that often jeopardize people's access to therapy, or the success thereof: issues with mental health and illiteracy. To her, mental health issues that are very common among people with aphasia are not currently taken as seriously as they should be, and so are treated adequately. T1 hinted to the fact that in most cases the only approach to treating depression is to prescribe drugs, while instead it would need to be at least complemented with other kinds of support ("I think that one remedy is to say "We'll prescribe a pill for you", but it's not about prescribing the pill. It's about providing mental health, actual treatment to coincide with other therapy, and I think it should be a huge part of the recovery process"; see

Table A.5). To T1 this issue is particularly serious: "it is quite common that there are mental health issues that emerge following stroke...we are seeing a lot more patients who are younger...and definitely mental health is a factor for a lot of these individuals, simply because they're still in the prime of their life, and all of a sudden they are aware that they can't do what they used to do or don't have the communication skills to be able to return to work."

Finally, a phenomenon that is not adequately taken into account by current approaches and in the resources used to deliver speech-therapy is that a significant number of seniors, especially among those living in rural areas, have poor literacy skills that predate the onset of their aphasia. The fact that assessing these skills in people with aphasia is not a trivial matter, along with the reluctance of many to admit their limitation, creates a blind-spot that, to T1, increases the chances of making therapists' efforts counterproductive.

Theme 4: Positive Aspects of Therapy

When analyzing participants' opinions about what they found most helpful about therapy we identified three subthemes: interaction; play; monitoring and progress (see Table A.4). People enjoyed interacting with the therapist, or with other people. S3 thought that his therapist's firmness in ensuring that he complied with the prescribed homework was key to ensure he made progress. Therapy helped him learn strategies to better monitor his behavior, to keep track of what he struggled with and to report these issues back to his therapist on a weekly basis, when solutions to such problems were to be discussed and approved together. S7 and S9 said that working in tandem with a therapist and receiving feedback about their skill level promoted engagement and a sense of security that motivated them pushing their limits. Finally, S1 and S6 said that they found it very helpful when therapists or people leading book-club sessions used games (or game mechanics such as enforcing time limits) to perform tasks involving comprehension and word-finding skills.

Theme 5: Aspects to Improve in Therapy

One participant (S5), reflecting on ways therapy could have been more beneficial to her, expressed a desire to engage more with technology, and to use it as a tool to complement traditional methods ("Looking back we never did much technology and I would have liked more technology and less going over words. She gave me a binder with words and I had to go home and say words but...that was pretty boring but it was great too"; see Table A.4).

Theme 6: Therapists' Opinions on the Most Appropriate Way to Deliver Treatment

All therapists firmly endorsed interventions that are guided by a client-centered approach. Here, as T1 put it, "They [their clients] decide what they want to work on...it is [an approach to therapy] directed on what the patient's goals are, not what my goals for them are..." (see Table A.6 for relevant meaning units). In short, this approach entails shaping the therapeutic experience in a way that engages directly with scenarios that are relevant to a given client, or, in other words, in "...finding tasks, communication goals that are meaningful, practical, and functional for the client because there doesn't seem to be as much progress in the patient's outcome in recovery when a lot of their time and energy is devoted to tasks that don't consequently relate to day-to-day activities" (T1). An important corollary of the client-centered approach is that the therapist's role is to help clients reach their goals by providing assistance, and not by imposing specific directions, throughout therapy (T1).

For therapists, the most obvious advantages of this approach are that it promotes engagement, motivation and as a result clients' adhesion to homework (T2 and T3). Examples of strategies used to motivate clients are engaging caregivers in therapeutic activities to "facilitate a discussion in which they [the client] get to be a contributing network member and to show off how competent they are... That generally motivates them to somewhat participate...in the therapy activities, it gives them the "this is why I'm doing this because I can do better"" (T2). Taking part in group sessions also alleviates the sense of loneliness and helplessness. In this context, observing people facing similar issues helps PwA feel better about their struggles and frustrations, and, importantly, feel inspired by seeing improvements of others (T2).

Barriers and Facilitators in Use of Technology

Participants reported being familiar with at least one device among personal computers (PC, 6 people), tablet (6 people), and smartphone (4 people). They commonly used the PC and tablet to send emails and read the news; S1 reported using both devices to play games, while S2 typically used the PC and S3 the tablet for this purpose. Only S1 reported using the smartphone, mostly as a way to communicate through emails, calls and texts (see Table A.9). Below are the views of participants with aphasia about the barriers and facilitators they experienced when using each of the above-mentioned devices (see Table A.11 and Table A.10, respectively, for relevant meaning units).

PC

There was not a consistent theme related to aspects that facilitate the use of PCs. Relative to other devices, people preferred computers due to their versatility (S1), familiarity (S2), processing power (S3) and user-friendliness (S9). S1 favored the PC because she could play games or use programs that were not available for tablet; the needs of S9 were best suited by computers because he frequently had to work with material that is organized in nested folders, and he found this task much easier to perform using the computer rather than the smartphone.

Two distinct themes emerged regarding the main barriers to the use of PCs. One theme reflected a struggle understanding how to navigate programs - in particular those requiring the execution of multiple steps in sequence - and finding adequate support (S1). The other one emerged from S1 and S2 and related to a difficulty using the touchpad and mouse, both of which require users to perform fine movements. To S2, the mouse cord caused problems as it tends to get tangled up when the mouse is used, and getting a cordless mouse solved the problem.

Tablet

Five participants expressed very positive opinions about tablets. Three of them praised tablets' user-friendliness — in particular S3, who suffers from hemiparesis and does highly value devices that allow him to interact with simple touch. One participant (S5) liked the tablet because she had lots of experience using it, and because she found it well suited to play games. Finally, one participant (S8) said tablets had everything she needed.

Three participants discussed barriers in the use of tablets: to S2 the touchscreen is too sensitive; S3 found tablets to be slow relative to PCs; S7 found it hard to type using tablets.

Smartphone

Smartphones were found to be user-friendly (three participants) and portable. S1 found the callback function very useful as it allowed her to issue a call without having to type the number multiple times; S5 praised the speech-to-text function as it allowed her to take notes without having to write, a task that was difficult for her; S9 enjoyed the ease with which he could use the smartphone to schedule events using Google calendar.

The main barriers experienced using a smartphone were related to low familiarity with it (S1 and S8), difficulty typing numbers or reading on a small screen (S6) (a problem addressed by changing the size of the font) and with navigating nested folders (S9).

Favorite device

Four participants (S1, S2, S3 and S9) expressed preference for PC over other devices; 3 participants preferred iPhone (S5, S6 and S9); 1 preferred the iPad (S7); 1 had no preferences (S8).

Desirable Features of CBATs and VGs

Two themes were identified regarding participants' experience with and opinions about desirable features of CBATs: feedback and ability to adapt level of difficulty; usability (see Table A.15 and Table A.8 for relevant quotes from people with aphasia and therapists, respectively).

Theme 1: Challenge

This theme had four subthemes: adaptivity, variety, engagement, and feedback. Four participants discussed the importance of playing video games that adjust the level of difficulty based on the user's performance. Participants praised games offering a wide variety of exercises (S1) and those that do not only provide the opportunity to train language skills, but to also engage the user with enriching experiences (S5). According to both people with aphasia and therapists, CBATs should be able to provide accurate and real-time feedback on performance ("...I like to know if it's kind of where you went wrong, so you can fix it... you can fix it because you can do them over and over, but there's no point if you're going to be wrong all the time. But you don't know where, why you're wrong. You need to know where you're wrong.", S1; "I like that they show how performance is going, especially for those that have issues with awareness", T3).

Theme 2: Usability

The theme of usability did also build upon four subthemes: ease of use, tailoring of material, familiarity, relevance of tasks. Regarding ease-of-use, participants stressed the importance of an application to be easily used with one hand (S3); to use voices with a Canadian accent (S2); and to be cheap (S5). Therapists (T1 and T2) placed an emphasis on aphasia-friendliness: "I think they have to be very, very easy to navigate. Very few kinds of buttons and whistles and what's not involved with it" (T1).

To therapists, very important features of video games were the ability to tailor the content to a person's interests ("If they're meaningful words and phrases to the individual, then possibly [it is beneficial]. If they're not I would question their relevance and use… ", T1) and to integrate real-life scenarios ("What I would like to see is more integration with speech-recognition software, more real-life scenarios in a game setting, for example: if you are practicing ordering...Those people with fluent aphasia do much better with tasks that are more functional, and I think that that's the biggest challenge, so I hope to see improvements", T3).

Finally, therapists expressed positive opinions about the possibility of integrating voice-recognition software to provide feedback on naming accuracy ("In terms of the feedback that clients are getting back on their production, in most apps, it's recording the production, playing it back and giving them a self-evaluation or having someone else evaluate it. The tablet itself is not evaluating the correctness of the production... [Automatic feedback] It will be more reliable and accurate...the more independent we can make the clients the better, but they are reliable.", T2).

Common Pastimes and Goals of PWA

When it came to activities and interests, we did not find a common thread running across participants. Three participants (S1, S2 and S5) reported engaging with substantial physical training and outdoor activities, while S7 reported enjoying walking; gardening was among the favorite pastimes of S3 and S6, woodworking was for S6 and S9, and drawing with charcoal was for S6. S7 reported enjoying playing cards and board games such as monopoly, along with the team games played at the stroke and aphasia clubs he was engaged with. Finally, S1 mentioned reading as one of her favorite activities (see Table A.13). When asked about what kind of goals people with aphasia do most commonly set for themselves before starting therapy, we could find common ground only among very general everyday activities (see Table A.7). T2 and T3 discussed the fact that, in general, their clients expressed the willingness to gain back the ability to take part in conversations again, being in-person via spoken language, or via email/texting. One of therapists' favorite examples was clients expressing their desire to be able to take part in family conversations without feeling overwhelmed.

DISCUSSION

Computer-based therapies for aphasia have long been regarded as a promising tool to help make best-available therapy services easily accessible to people with aphasia. Nowadays, as smartphones, tablets, computers and the internet are ubiquitous and easily accessible, the potential of CBATs can be finally fully realized. Yet, most, if not all, of available products present some limitations that do strongly curb their effectiveness: they do not provide feedback on how well people speak, hence limiting the ability of users to learn from mistakes; and they do not account for important human factors such as user-engagement and motivation, which are essential to ensure that users adhere to the required regimes of therapy (Swales et al., 2016). In light of these limitations, we identified the opportunity to greatly improve upon available products by using state-of-the-art ASR technology, and by adopting best practices in game design to gamify computerized aphasia therapy. Broadly, this study was conceived as the very first step of customer discovery, aiming to test, and eventually adjust, the core assumptions about the nature of the problem that we seek to solve and about the adequateness and real potential of the solution that we are proposing. In addition to it, this study was conceived to gather information on how to design a program that is user-friendly and whose gamified experiences appeal to the tastes of people with aphasia. To achieve these objectives, we conducted semi-structured interviews with target end-users (people with aphasia and therapists) and with caregivers.

Summary of Pain Points Throughout the Continuum of Care

By and large, when it comes to characterizing the nature of the problem that CBATs can help address, this study largely confirmed the main findings presented in the previous chapter, highlighting how the current service delivery system falls short at ensuring adequate access to therapy services and, more in general, to provide support that is appropriate to meet the multifaceted needs of PwA and their carers.

Acute Care

The review of the literature presented in the previous chapter underscores two main aspects that negatively impact the quality of the service provided to PwA during the acute hospitalization phase (1-7 days from the onset of stroke). One is the markedly little time that clients spend engaging with therapy (Code & Heron, 2003), which goes counter recommendations (Hebert et al., 2016).

The other is the lack of adequate information about aphasia, its consequences, about ways to promote recovery, and about strategies that can be employed to facilitate communication (Avent et al., 2005; Foster et al., 2016; Simmons-Mackie et al., 2007). As stressed in the previous chapter, this issue leads to limited engagement of PwA in matters that are very important to them: plans for discharge from the hospital, for starting therapy, and the statement of therapy's goals are often discussed by service-providers only with caregivers, bypassing the person that should be at the center of the decision-making process (Godecke et al., 2014). Unsurprisingly, the lack of opportunity to connect with other people and to exert control over their lives leads clients to experience an increased sense of frustration and helplessness (Simmons-Mackie et al., 2007). This sense of frustration is often shared by caregivers as well, who frequently complain about not receiving enough information on aphasia from clinicians, and express the desire to be better assisted to, and prepared for, supporting their loved one (Simmons-Mackie, 2017).

In this regard, therapists who took part in this study discussed the importance of establishing a connection with family members and caregivers in order to provide information about aphasia and its consequences, along with information about the importance of their role as active partners during therapy and as assistants outside of therapy. These comments reflect the recommendations on best practices to adequately serve people with aphasia during the acute hospitalization phase (Hebert et al., 2016; Simmons-Mackie, 2017).

Inpatient and Outpatient Rehabilitation

When talking about their experience with rehabilitation, people with aphasia confirmed what has been documented at length in the literature, namely that limited subsidies, the high costs of private therapy and the long waiting-lists pose significant barriers to access therapy. According to some participants, long waiting lists prevented them from starting therapy for six months or longer after being discharged from the hospital. Due to the high costs of engaging with private therapy, most people have no choice but to wait, unassisted, until they can start the subsidized treatment. The lack of engagement with therapy in the first few months after the onset of stroke can have significant negative consequences for PwA. For one, recent studies have unambiguously demonstrated that speech-therapy delivered during the acute and subacute phases post-stroke onset yields significant benefits, over and above the well documented positive effects that are attributable to spontaneous recovery (Roberts et al., 2022). Therefore, while the debate over whether speech-language therapy delivered during the acute and subacute phases is more effective than when delivered during the chronic phase continues (Brady et al., 2022; Fridriksson & Hillis, 2021), missing the opportunity to engage with therapy early on causes needless delays in recovery. This might exacerbate the negative impact that aphasia has on the psychological well-being of a person and of her relatives, and on the quality of her social relationships (Hilari, 2011; Hilari et al., 2010; Kauhanen et al., 2000; Morris et al., 2017; Ross et al., 2010).

Other factors that limit access to therapy are the lack of staff available to serve clients, and issues with transportation. According to therapists, the latter factor represents a barrier to accessing therapy not only to people living in rural areas, but also to people living in metropolitan areas where access to rehabilitation centers is relatively easy.

Therapists also highlighted the fact that oftentimes many clients present mental health issues that hinder the efficacy of aphasia therapy, and that issues like depression are not taken seriously enough in the context of rehabilitation. This opinion is widely shared among therapists (Simmons-Mackie, 2017), and is supported by evidence suggesting that, despite an estimate 70% of people with aphasia experience depression in the first 3 months after stroke, depression is often undiagnosed or untreated (Townend et al., 2007).

The limitations described by our participants fit within the wider context of the state of inpatient and outpatient rehabilitation in North America, which documents exceedingly low rates of admission of stroke patients with a moderate-severe diagnosis are referred to an inpatient rehabilitation facility (Simmons-Mackie, 2017), along with levels of intensity and dosage that are far below what is recommended by best practices (i.e., an average of 13 minutes of therapy per day in inpatient (Foley et al., 2012), as opposed to the 1 hour that is recommended (Hebert et al., 2016)).

Long-Term Care

A common concern expressed by therapists is that the service delivery model currently provides very little in the way of support once a person ends their inpatient or outpatient rehabilitation. While aphasia centers have been taken as a promising model of service delivery to address this problem, it has been estimated that, as of 2016, throughout North America there were only about 26 such centers (Simmons-Mackie, 2017). Therefore, despite their promise, they currently serve only a fraction of the people in need.

How Can CBATs Help Address the Identified Problems?

Our interviews with therapists and people with aphasia, along with the information that we gathered from the literature, clearly identify important pain points that people with aphasia and their families experience throughout the entire continuum of care.

- *During the acute hospitalization phase, there is a clear need for people with aphasia and their families to receive adequate education about aphasia, along with training on how to provide communication support to people with aphasia.*

In this regard, CBATs could be designed to act as forums that facilitate the connection between people who have been recently diagnosed with aphasia and those that have been in a similar situation for enough time to feel comfortable in sharing their experience, providing insights into the most common challenges, and how they can be approached. Similarly, a CBAT application can provide comprehensive and aphasia-friendly materials (i.e. videos) to better educate people about the nature of aphasia; the potential of people to recover, and what a successful recovery does entail; and about the strategies that can be adopted to successfully communicate throughout the different stages of recovery. Similar solutions are already in place (Lingraphica, n.d.).

- *Following acute hospitalization, the public healthcare system typically funds rehabilitation for a very limited amount of time, which is by no means sufficient to meet the needs of people with aphasia.*

In this context, CBATs can provide a way to deliver treatment that is cost-effective and that can be carried out for as long as needed throughout the chronic stage of aphasia.

- *There is the lack of a system, at the community-level, that provides support for people with aphasia to help them reintegrate in community.*

This is a problem that is particularly accentuated in small towns, where there may only be few people living with aphasia and therefore not enough demand to justify investing in community services (Kaye, 1997; *Report to Congress on Traumatic Brain Injury in the United States: Epidemiology and Rehabilitation*, 2015). In this case, a software-based application can be designed to provide people with the opportunity to engage in virtual communities, coordinate multiplayer activities etc.. Again, examples of solutions similar to the one proposed are already in place and very popular among people with aphasia and caregivers (Lingraphica, n.d.; *Online Aphasia Communities*, n.d.).

Stakeholders Opinions About the Proposed Solutions

One of the objectives of this study was to assess the opinions of people with aphasia and speech-language therapists about our ideas on how to innovate upon current computer-based therapies for aphasia. In general, participants from both groups endorsed our ideas: all people with aphasia said they would try a novel computer-based therapy to help with their aphasia, and all except one (S7) said they would play a video game designed to do that. All therapists expressed positive opinions about both the idea of using video games, and that of using voice-recognition systems to provide feedback on performance. These opinions echo those reported in other studies as well (Finch & Hill, 2014a), suggesting that there is a consensus on the potential to improve upon the current state of the art in CBATs by integrating novel technologies such as ASR and video games.

Design of a Minimum-Viable Product: Lessons Learned from the Literature and from the Interviews with Study Participants

Overall, all people with aphasia that we interviewed in this study reported using at least one CBAT-capable device regularly. The most favorite ones were the PC and smartphone. Participants' opinions about the desirable features of CBATs and video games lined up with the information that we gathered from the literature: CBATs have to be aphasia-friendly (i.e. allow for easy interaction by people with hemiparesis; keep amount of information down to levels that can be easily processed by people with cognitive, in addition to language, impairment etc.); they have to dynamically adapt the level of difficulty to a person's changing skill levels; they have to offer a variety of exercises, and the opportunity to receive detailed, real-time feedback on performance. To therapists, a key element to ensure that exercises are engaging and effective is that they engage users with content and activities that are relevant to them.

In order to determine how to design a CBAT to maximize usability we wanted to understand what type of challenges people with aphasia experience when using specific devices. We did not find a striking pattern across participants: some expressed frustration when using the mouse and touchpad of a PC, and reported favoring devices, such as tablets, that feature a touchscreen. Others described difficulties with interacting with a touchscreen, or writing on a virtual keyboard. This is not surprising, given the high heterogeneity in cognitive/motor deficits experienced by people with aphasia and in their preferences and familiarity with technology more in general. For this reason, one way to accommodate the range of preferences expressed by the target users is to design an MVP that can be used on any device/platform.

The insights that we gathered from surveying the literature and from the interviews that we conducted in this study have been summarized in Table A.16, which contains a list of user needs that are essential to inform the design of a prototype of a CBAT.

Next Steps

As a next step, we plan to build the first prototype of a computer-based therapy featuring ASR. The main objective will be to test the feasibility of using the ASR online while employing an evidence-based approach (picture naming) to remediate word-finding skills, rather than on a dataset of speech samples, and to assess users' perceptions about its accuracy and usefulness. The prototype will implement the naming task in a simple game, using ASR to provide users with immediate feedback on whether their naming is accurate. We made this choice for two reasons: a) word-finding difficulties (anomia) are a frequent complaint among people with aphasia, and evidence supports a practice-based approach to recover this ability (Chapey, 2008); b) a picture naming task is very easy to implement, and therefore it is an ideal starting point to assess the overall usability of the program.

This preliminary test is key to validate one of the main rationales for this thesis, as the integration of the ASR is the main innovation introduced in this project. Once feasibility (measured by the ease with which users can adopt the program to perform simple tasks, by their opinions about its usefulness, and by the accuracy of the ASR system in providing feedback) will be firmly established, additional features that were identified in this chapter — offering a wide range of tasks targeting multiple language functions; allowing users to select training stimuli that are relevant to them; implementing an adaptive algorithm that regularly calibrates the level of difficulty to a person's progress — will be progressively introduced to improve the quality and effectiveness of the program.

Limitations

This study has some limitations. First, we planned to integrate the perspectives of caregivers of people with aphasia in an effort to have a comprehensive view of the most pressing unmet needs, and how technology could address them. Nonetheless, we were not able to recruit anyone from this group. This was due to the fact that all the caregivers that we got in touch with had a partner with aphasia that took part in the study. As they all assisted their partner during the interviews, they were not willing to commit to another session to be interviewed on their own. Similarly, we could only interview 3 therapists, as opposed to the target sample size of 10. One reason behind our difficulties in meeting these targets is that we limited our recruitment efforts to the province of Nova Scotia: we recruited participants with aphasia via the mailing list of a charitable association that we collaborate with, and that operates exclusively in Nova Scotia; similarly, we recruited therapists via the Nova Scotian registry of licensed professionals. One solution to this problem is to exploit the online format to expand our reach outside of Nova Scotia, as indeed we did to recruit participants for subsequent studies presented in this thesis. Second, the thematic analysis was conducted only by one person. This goes counter to best practices regarding the adoption of this method. Thematic analysis is a procedure that, by its own nature, does strongly rely on a person's judgment when identifying the main themes underlying a conversation. Therefore, to ensure that the effect of bias is minimized, it is recommended that at least 2 people first go independently over all the steps of thematic analysis. Results should then be compared, and eventual inconsistencies should be resolved until a consensus is reached on the type of themes that are identified (Braun & Clarke, 2006; Graneheim & Lundman, 2004). Results of this study might therefore be biased, in the sense that (some of) the themes that were identified might not survive the test of comparison across multiple independent analyses. On the other hand, results of our study converged with previous findings, at least in regard to the topics that we investigated and for which there was an overlap with the existing literature. Although this does not warrant that results are as strong as if the proper procedure would have been followed, it suggests that the issue might also not be too severe.

CONCLUSIONS

In general, this study contributed to a better understanding of the nature of the problem that we intended to address, namely the fact that the current service delivery model does not adequately meet the needs of people with aphasia. It also provided support for the solution we proposed to tackle this problem. Not only all participants of this study manifested positive attitudes towards adopting CBATs, but they also endorsed our ideas on integrating an ASR to provide feedback on naming accuracy and on employing gamification techniques to make the therapeutic exercises more appealing and motivating to people with aphasia, with the goal to maximize adherence to treatment. Both people with aphasia and therapists provided important insights on what barriers and facilitators affect the experience of using modern devices, and on the features that they would like to see in a CBAT. This is of particular importance to inform about the optimal design of a prototype of CBATs, and as such the insights from this study will be helpful to those involved in the design of CBATs.

ACKNOWLEDGEMENTS

I want to thank the following people and institutions for providing help with the design and execution of this study: Lynda Wozniak and Ellina Kostopoulos; Glenn Hougan; Aaron Newman; Morgan Johnson, who (a) contributed to the layout of the questionnaires and (b) together with Lisa Beck, helped with recruiting and interviewing participants; the students who transcribed the interviews (Andrew Son, Hayam Ashraf Mahmoud-Ahmed, Salma Ismail, Archa Rajagopalan, Natalie Smith and Abigail Howard-Gosse); the School of Communication Sciences and Disorders, of Dalhousie University, for gently allowing us to make use of their spaces to conduct the in-person interviews with individuals with aphasia. Finally, I would like to thank all the members of Aphasia Nova Scotia who helped advertise this study, and who partook in it.

REFERENCES

Avent, J., Glista, S., Wallace, S., Jackson, J., Nishioka, J., & Yip, W. (2005). Family information needs about aphasia. *Aphasiology*, *19*(3–5), 365–374. https://doi.org/10.1080/02687030444000813

Berg, K., Askim, T., Balandin, S., Armstrong, E., & By Rise, M. (2017). Experiences of participation in goal setting for people with stroke-induced aphasia in Norway. *Disability & Rehabilitation*, *39*(11), 1122–1130. https://doi.org/10.1080/09638288.2016.1185167

Blank, S., & Dorf, B. (2012). *The Startup Owner's Manual* (First). K&S Ranch Publishing Division.

Brandeburg, C., Worrall, L., Copland, D., & Rodriguez, A. D. (2017). Barriers and facilitators to using the CommFitTM smart phone app to measure talk time for people with aphasia. *Aphasiology*, *31*(8), 901–927. https://doi.org/10.1080/02687038.2016.1219016

Brandeburg, C., Worrall, L., Rodriguez, A. D., & Copland, D. (2013). Mobile computing technology and aphasia: An integrated review of accessibility and potential uses. *Aphasiology*, *27*(4), 444–461. https://doi.org/10.1080/02687038.2013.772293

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/doi.org/10.1191/1478088706qp063oa

Cahana-Amitay, D., & Albert, M. (2015). *Redefining Recovery From Aphasia*. Oxford University Press.

Canadian Stroke Network. (2011). *The Quality Of Stroke Care in Canada*. https://www.canadianstroke.ca/sites/default/files/resources/QoSC-EN.pdf

Chapey, R. (2008). *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders* (5th ed.). Lippincott Williams and Wilkins.

Chapey, R., Duchan, J., Elman, R., Garcia, L., Kagan, A., Lyon, J., & Simmons-Mackie, N. (2000). *Life Participation Approaches to Aphasia: A statement of values for the future.* (5; pp. 4–6). ASHA Leader.

Chen, C. C., & Bode, R. K. (2011). Factors Influencing Therapists' Decision-Making in the Acceptance of New Technology Devices in Stroke Rehabilitation. *American Journal of Physical Medicine & Rehabilitation*, *90*, 415–425. https://doi.org/10.1097/PHM.0b013e318214f5d8

Code, C., & Heron, C. (2003). Services for aphasia, other acquired adult neurogenic communication and swallowing disorders in the United Kingdom, 2000. *Disability & Rehabilitation*, *25*(21), 1231–1237. https://doi.org/10.1080/09638280310001599961

Cota, T. T., & Ishitani, L. (2015). Motivation and benefits of digital games for the elderly: A systematic literature review. *Revista Brasileira de Computacao Aplicado*, *7*(1), 2–16. https://doi.org/10.5335/rbca.2015.4190

De Schutter, B., & Malliet, S. (2014). The older player of digital games: A classification based on perceived need satisfaction. *Communications*, *39*(1), 67–88. https://doi.org/10.1515/commun-2014-0005

Diaz-Orueta, U., Facal, D., Nap, H. H., & Ranga, M.-M. (2012). What Is the Key for Older People to Show Interest in Playing Digital Learning Games? Initial Qualitative Findings from the LEAGE Project on a Multicultural European Sample. *Games for Health Journal: Research, Develoment, and Clinical Applications*, *1*(2). https://doi.org/10.1089/g4h.2011.0024

Doroudian, A. (2019). *Design, Development, and Evaluation of an Online Escape Game for Older Adults*. Simor Fraser.

Draper, P., & Brocklehurst, H. (2007). The impact of stroke on the well-being of the patient's spouse: An exploratory study. *Journal of Clinical Nursing*, 264-- 271. https://doi.org/10.1111/j.1365-2702.2006.01575.x

Elman, R. (2016). Aphasia Centers and the Life Participation Approach to Aphasia. A Paradigm Shift. *Topics in Language Disorders*, *36*(2), 154–167. https://doi.org/10.1097/TLD.0000000000000087

Finch, E., & Hill, A. J. (2014). Computer use by People with Aphasia: A Survey Investigation. *Brain Impairment*, *15*(2), 107–119. https://doi.org/doi.org/10.1017/BrImp.2014.17

Foley, N., McClure, J. A., Meyer, M., Salter, K., Bureau, Y., & Teasell, R. (2012). Inpatient rehabilitation following stroke: Amount of therapy received and associations with functional recovery. *Disability & Rehabilitation*, *34*(25), 2132–2138. https://doi.org/10.3109/09638288.2012.676145

Foster, A., O'Halloran, R., Rose, M., & Worrall, L. (2016). "Communication is taking a back seat": Speech pathologists' perceptions of aphasia management in acute hospital settings. *Aphasiology*, *30*(4–6), 585–608.

Galliers, J. R., Wilson, S., Muscroft, S., Marshall, J., Roper, A., Cocks, N., & Pring, T. (2011). Accessibility of 3D Game Environments for People with Aphasia: An Exploratory Study. *Paper Presented at the 13th International ACM SIGACCESS Conference on Computers and Accessibility, 24-26 Oct, Dundee, Scotland*.

Gamberini, L., Alcaniz, M., Barresi, G., Fabregat, M., Ibanez, F., & Prontu, L. (2006). Cognition, technology and games for the elderly: An introduction to ELDERGAMES Project. *PsychNology Journal*, *4*(3), 285–308.

Ganzfried, E., & Hinckley, J. (2012). *The persistent needs of people living with aphasia: Results of a national survey*. Clinical Aphasiology Conference, Lake Tahoe, CA. http://aphasiology.pitt.edu/id/eprint/2399

Godecke, E., Armstrong, E., Hersh, D., & Bernhardt, J. (2014). Missed opportunities: Communicative interactions in early stroke recovery. *International Journal of Stroke*, *9*(48).

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*, 105–112. https://doi.org/doi:10.1016/j.nedt.2003.10.001

Greig, C. A., Harper, R., Hirst, T., Howe, T., & Davidson, B. (2008). Barriers and Facilitators to Mobile Phone Use for People with Aphasia. *Topics in Stroke Rehabilitation*, *15*(4), 307–324. https://doi.org/10.1310/tsr1504-307

Hebert, D., Lindsay, M., McIntyre, A., Kirton, A., Rumney, P., … Bagg, S. (2016). Canadian Stroke Best Practice Recommendations: Stroke Rehabilitation Practice Guidelines. *International Journal of Stroke*, *11*(4), 459–484.

Hilari, K. (2011). The impact of stroke: Are people with aphasia different to those without? *Disability & Rehabilitation*, *33*(3), 211–218.

Hilari, K., Northcott, S., Roy, P., Marshall, J., Wiggins, R., Chataway, J., & Ames, D. (2010). Psychological distress after stroke and aphasia: The first six months. *Clinical Rehabilitation*, *24*(2), 181–190.

Hill, A. J., & Breslin, H. M. (2016). Refining an Asynchronous Telerehabilitation Platform for Speech-Language Pathology: Engaging End-Users in the Process. *Frontiers In Human Neuroscience*, *10*(640). https://doi.org/10.3389/fnhum.2016.00640

Howe, T., Davidson, B., Worrall, L., Hersh, D., Ferguson, A., Sherratt, S., & Gilbert, J. (2012). *`You needed to rehab. . . Families as well': Family members' own goals for aphasia rehabilitation*. *47*(5), 511–521. https://doi.org/10.1111/j.1460-6984.2012.00159.x

Kaufman, D., Ma, M., Sauvè, L., Renaud, L., & Duplàa, E. (2019). Benefits of Digital Gameplay for Older Adults: Does Game Type Make a Difference? *International Journal of Aging Research*, *2*(43).

Kauhanen, M., Korpelainen, J., Hiltunen, P., Määttä, R., Mononen, H., Brusin, E., & Myllylä, V. (2000). Aphasia, depression, and non-verbal cognitive impairment in ischaemic stroke. *Cerebrovascular Diseases*, *10*(6), 455–461.

Kaye. (1997). *Disability watch: The status of people with disabilities in the United States.* Disability Statistics Center, Institute for Health and Aging, University of California at San Francisco.

Kim, E. S., Suleman, S., & Hopper, T. (2020). Decision Making by People With Aphasia: A Comparison of Linguistic and Nonlinguistic Measures. *Journal of Speech, Language, and Hearing Research*, *63*(6), 1845–1860. https://doi.org/10.1044/2020_JSLHR-19-00182

Laliberte, M., Gaureau, C., & LeDorze, G. (2016). A pilot study on how speech-language pathologists include social participation in aphasia rehabilitation. *Aphasiology*, *30*(10), 1117–1133.

Lingraphica. (n.d.). *Virtual connections: An entire community at your fingertips*. https://www.aphasia.com/virtual-connections-info/

Mayo, N. E., Anderson, S., Barclay, Cameron, Desrosiers, J., Eng, J. J., Huijbregts, M., Kagan, A., MacKay-Lyons, M., Moriello, C., Richards, C. L., Salbach, N. M., Scott, Teasell, R., & Bayley, M. (2015). Getting on with the rest of your life following stroke: A randomized trial of a complex intervention aimed at enhancing life participation post stroke. *Clinical Rehabilitation*, *29*(12), 1198–1211.

Morgan, D. L. (1993). Qualitative Content Analysis: A Guide to Paths Not Taken. *Qualitative Health Research*, *3*(1), 112–121.

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

Mortenson, W. B., Sixsmith, A., & Kaufman, D. (2017). Non-Digital Game Playing by Older Adults. *Canadian Journal of Aging*, *36*(3), 342–350. https://doi.org/10.1017/S0714980817000162

Nap, H. H., de Kort, Y. A. W., & Ijsselsteijn, W. A. (2009). Senior gamers: Preferences, motivations and needs. *Gerontechnology*, *8*(4), 247–262. https://doi.org/10.4017/gt.2009.08.04.003.00

Online aphasia communities. (n.d.). https://www.aphasia.org/aphasia-resources/online-communities/

Palmer, R., Enderby, P., & Paterson, G. (2013). Using computers to enable self-management of aphasia therapy exercises for word finding: The patient and carer perspective. *International Journal of Language & Communication Disorders*, *48*(5), 508–521. https://doi.org/10.1111/1460-6984.12024

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Sage Publications, Inc.

Pettit, L. K., Tonsing, K. M., & Shakila, D. (2017). The perspectives of adults with aphasia and their team members regarding the importance of nine life areas for rehabilitation: A pilot investigation. *Topics in Stroke Rehabilitation*, *24*(2), 99–106. https://doi.org/10.1080/10749357.2016.1207148

Report to Congress on Traumatic Brain Injury in the United States: Epidemiology and Rehabilitation. (2015). National Center for Injury Prevention and Control; Division of Unintentional Injury Prevention.

Rohde, A., Townley-O'Neill, K., Trendall, K., Worrall, L., & Cornwell, P. (2012). A comparison of client and therapist goals for people with aphasia: A qualitative exploratory study. *Aphasiology*, *26*(10), 1298–1315. https://doi.org/10.1080/02687038.2012.706799

Ross, K., LaPointe, L., & Katz, R. (2010). *Risk factors for depression in aphasia: Clinical implications*. Clinical Aphasiology Conference, Isle of Palms, SC.

Salmon, J. P., Dolan, S. M., Drake, R. S., Wilson, G. C., Klein, R. M., & Eskes, G. A. (2017). A survey of video game preferences in adults: Building better games for older adults. *Entertainment Computing*, *21*, 45–64. https://doi.org/10.1016/j.entcom.2017.04.006

Schutter, D. (2011). Never Too Old to Play: The Appeal of Digital Games to an Older Audience. *Games and Culture*, *6*(2), 155–170. https://doi.org/10.1177/1555412010364978

Seah, E. T. W., Kaufman, D., Sauvè, L., & Zhang, F. (2018). Play, Learn, Connect: Older Adults' Experience With a Multiplayer, Educational, Digital Bingo Game. *Journal of Educational Computing Research*, *56*(5), 675–700. https://doi.org/10.1177/0735633117722329

Simmons-Mackie, N. (2008). Social Approaches to Aphasia Intervention. In *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders* (5th ed., p. 1091). Lippincott Williams and Wilkins.

Simmons-Mackie, N. (2017). *Aphasia in North America*. https://www.aphasiaaccess.org/white-papers/

Simmons-Mackie, N., & Holland, A. L. (2011). Aphasia Centers in North America: A Survey. *Seminars in Speech and Language*, *32*(3), 3–105. https://doi.org/10.1055/s-0031-1286175

Simmons-Mackie, N., Kagan, A., Christie, O., Huijbregts, M., McEwen, S., & Willems, J. (2007). Communicative access and decision making for people with aphasia: Implementing sustainable health care systems change. *Aphasiology*, *21*(1), 39–66.

Swales, M. A., Hill, A. J., & Finch, E. (2016). Feature rich, but user-friendly: Speech pathologists' preferences for computer-based aphasia therapy. *International Journal of Speech-Language Pathology*, *18*(4), 315–328. https://doi.org/10.3109/17549507.2015.1081283

Townend, E., Brady, M. C., & McLaughlan, K. (2007). Exclusion and inclusion criteria for people with aphasia in studies of depression after stroke: A systematic review and future recommendations. *Neuroepidemiology*, *29*(1–2), 1–17. https://doi.org/10.1159/000108913.

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing and Health Sciences*, *15*, 398–405. https://doi.org/doi: 10.1111/nhs.12048

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

Winstein, C., Stein, J., Arena, R., Bates, B., Cherney, L., & Cramer, S. (2016). Guidelines for adult stroke rehabilitation and recovery: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, *47*, e98–e169.

Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersh, D., Ferguson, A., & Davidson, B. (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, *25*(3), 309–322. https://doi.org/10.1080/02687038.2010.5083

Chapter 3 – USABILITY STUDY OF THE FIRST PROTOTYPE OF A COMPUTER-BASED THERAPY FOR APHASIA FEATURING AUTOMATIC SPEECH-RECOGNITION

ABSTRACT

This study set out to assess the usability of the first prototype of a program that is intended to allow people with aphasia to self-administer therapy targeting word-finding and naming skills. Importantly, the program integrates automatic speech-recognition (ASR), an important innovation relative to state-of-the-art programs that enables the provision of real-time feedback on the accuracy of vocal responses.

Seven people with aphasia with mild naming deficits took part in this study, using the prototype to simulate a short training session wherein they had to engage with a picture naming task under the guidance of feedback delivered by the ASR. Subsequently, participants provided their opinions on the usability of the program and on the perceived accuracy and usefulness of the feedback received from the ASR.

The results highlight that, in its current form, most participants could use the program to self-administer therapy targeting naming deficits, independently or with minimal support. Furthermore, most participants expressed positive opinions regarding the usefulness of the ASR feedback and were satisfied with their experience, saying that they would use improved versions of the program to self-administer therapy. Overall, the results of this study provide useful insights that chart a clear path ahead on how to develop the next version of the program.

INTRODUCTION

The previous chapter highlighted major gaps that affect the quality of services available to people with aphasia during the acute and chronic phases of the disorder. While during the acute phase there is an unmet need for adequate education on aphasia, its consequences, and on the strategies to optimally deal with it, in the chronic phase there is a need for adequate access to rehabilitation services, due to high costs and lack of professionals. As a further step towards reaching the goal of this dissertation, which is to develop a product that addresses the latter need, this chapter describes the process of designing, developing and testing the usability of the first prototype of a computer-based therapy (CBAT) featuring an ASR that makes it possible for people with aphasia to self-administer therapy targeting naming deficits.

Defining the Minimum Set of Features of the First Prototype

Intended Users

Since this is the very first step of the development process, we decided to start by implementing an exercise that allows users to practice their word finding and naming skills. There are two main reasons behind this choice. First, difficulties with this specific domain of language are the most frequent complaint among people with aphasia (Chapey, 2008; Goodglass & Wingfield, 1997; Nickels, 2002). Therefore, making a program that targets this need will maximize the number of people who may benefit, which will help with recruitment and further product development. Second, evidence supports a practice-based approach (picture naming) to improve word finding and naming abilities (Brady et al., 2016; Chapey, 2008; Goodglass & Wingfield, 1997; Nickels, 2002). A picture naming task is very easy to implement as a computer program, and therefore it is an ideal starting point to assess the overall usability of the prototype.

Intended Use and Use Environment

The program is intended to help people with aphasia improve their word-finding and naming skills in an independent fashion, without the need to be supervised by a caregiver or a therapist. To facilitate self-administration and adherence to a therapy regime that entails regular training sessions, the program was conceived to be accessible from the web such that users can access it whenever it suits them best. Places with a good internet connection and with minimal background noise are the ideal use environment.

User Needs and Software Requirements

In chapter 2 we discussed the main needs that a CBAT can address. In particular, Table A.16 provided a list of 20 user needs identified during the process of literature review. This study builds on these insights to define the user needs that should be addressed by the first prototype of the CBAT. Since the main aim of this study is to assess the feasibility of allowing people with aphasia to self-administer therapy via a CBAT that provides feedback on the accuracy of vocal responses, we reasoned that it was not necessary to address the user needs that are exclusively related to treatment efficacy, such as user needs 17-20 in Table A.16 (i.e. providing a variety of stimuli and tasks (UN17) that are salient (UN19) and challenging (UN20), and adapting the level of difficult to user's skill level (UN18)). Rather, here we decided to focus only on the user needs that are strictly related to the usability of the program and on the accuracy of the ASR. List of tasks of prototype v2.0 (see Appendix D) contains an adapted list of the user needs, along with the method employed to assess whether the need was met and with the study objective the need is associated with. User need 2 was added to the list to ensure that the program is easily accessible from any of the most commonly used web browsers. On the other hand, besides the user needs 17-20 from Table A.16, the list excludes user need 10, as it can be assessed only as new version releases of the program come about.

Table B.1 (see Appendix B) lists the main software requirements (SWR) that resulted from the design of the product. The section below describes in detail how each requirement was implemented.

Software Development

As discussed above, an important user need is to easily gain access to the program. To this end, the program was conceived to be accessible from the web, such that users could launch it by simply typing a URL on the web browser. This will prevent all issues that a person with aphasia can come across when having to download and install a program on their computer. Also, building the program as a web app circumvents possible issues due to incompatibility of the program with specific operating systems (OS). This feature was implemented by building the project to be compatible with the WebGL application programming interface (API)(SWR_1 in Table B.1), which allows interactive 2D/3D graphics to be rendered on a compatible web browser without the use of plug-ins.

Backend

Webserver. The application is hosted on the cloud using a droplet, a server as a service provided by the company DigitalOcean (DigitalOcean, n.d.) that comes ready with all the components needed to make a functioning web application (i.e. a webserver that handles requests from browsers, a filesystem, etc.). The webserver runs on an Ubuntu 16.04 operative system, deployed under the DigitalOcean droplet. Given that highly sensitive information is transferred from the web browser to the webserver and back (i.e. the audio files of users' voices), the communication between the two sides must be handled via the encrypted Hypertext Transfer

Protocol Secure (https)(SWR_16 in Table B.1). This was enabled by storing a Secure Socket Layer (SSL) certificate on the webserver. To further maximize the security of the webserver, remote access to it is allowed via Secure Shell (SSH) keys (SWR_17 in Table B.1). Finally, to simplify the process of launching the program from the web browser, the webserver's IP address was connected to the domain aphasiav1.com.

User Database. A MySQL user database was installed on the droplet to manage all the user-generated data that were of interest for this study. The database was designed to contain user-specific information about:

- Number of sessions completed
- Total number of words named correctly (according to the ASR) in each study session
- For each session, information about item-level accuracy as determined by the ASR.

For the sake of this study, we did not store any demographics about participants (i.e. age, sex, etc.). Each user account was named following this format: NAMEOFUSER Px, where 'x' stands for the order in which the specific participant enrolled in the study. For example, if a person called Francesco would be the fifth to enroll in the study, his user account would be called 'Francesco P5' (SWR_2 in Table B.1).

In line with best practices on how to best protect users' accounts from malicious attacks, user passwords were salted, then hashed (SWR_15, *Cryptographic Hash Function*, n.d.).

Voice Recognition. While numerous advanced ASRs are readily available for use, as commonly implemented they are optimized to interpret speech from young adults, which represent the large majority of the users who have adopted such technologies - and might therefore be suboptimal when used by people whose voice features diverge critically from those of young adults (i.e. children, elderly adults or people with communication disorders). This was confirmed by various studies aiming at integrating an ASR with computer-based tasks for speech therapy (Jamal et al., 2017). For this reason, before integrating an ASR into our web-based application for naming therapy, we set out to test its performance on a single-word recognition task, using a sample of audio recordings collected from people with aphasia. For this project, we employed an ASR built using the Kaldi (Povey et al., 2011) open source framework. Appendix C provides a high-level description of how the ASR converts raw audio files into (sequences of) words, along with the results of a preliminary assessment of its accuracy, conducted offline and using a sample of vocal responses recorded from people with aphasia while doing a naming task.

As discussed in Appendix C, the results of this analysis were strongly encouraging and supported the integration of the ASR into the prototype of CBAT that is presented in this study.

Frontend

The frontend was developed using Unity2D (2020.1.14.f1, *Unity*, n.d.), and built to be a webapp. The program was designed to unfold over 7 stages, described below.

As specified in the list of user needs, each stage of the program comes with a video tutorial that guides the user on the main task at hand, and on how to execute it (SWR_14 in Table B.1). For example, when the user is at Stage 2, the main task is to insert the credentials to login in the personal account. The associated video tutorial explains which field to tap or click on in order to insert username and password, and which button to press in order to execute the login via the inserted credentials (SWR_4 in Table B.1). All stages, except for Stage 4, engage users with only one task. Stage 4 engages the users with five tasks, hence there is a video tutorial for each task.

Below is a breakdown of each stage, introducing the main objective for the users and the sequence of actions needed to achieve it.

Stage 1: Launching the Program. Users launch the program by typing the associated URL (https://aphasiav1.com/LARRY_v1) on their web browser of choice. Upon loading, the program presents an introductory page. The scope of this stage is only to provide a welcoming message to the user. The associated video tutorial provides a brief introduction to the intended use of the program.

Stage 2: Login on Personal Account. Upon moving to Stage 2, users are prompted to insert the credentials to log into their account. The credentials are a user ID and a password that are provided by the researcher prior to the beginning of the study. Upon inserting their credentials, users are instructed to press the button 'Login' to move to Stage 3.

Stage 3: Providing an Overview of the Training Material. Stage 3 provides users with an overview of the pictures that they will be prompted to name in a given session (see Figure 3.1). Here, the only action that users have to execute is to move on to the next stage by pressing the "Start Task" button (SWR_5 in Table B.1).

Stage 4: Naming Task. On Stage 4, users have to engage with the picture naming task[2], which entails naming a series of pictures that appear on the screen. Figure 3.2 presents a schematic overview of the sequence of events that users experience from the moment they start a trial to the moment they have to choose what to do after they receive feedback from the ASR.

---

[2] The rationale, mechanism of action and evidence for the efficacy of the picture naming task when it comes to improving a person's ability to name words are described in the Introductory chapter of this thesis.

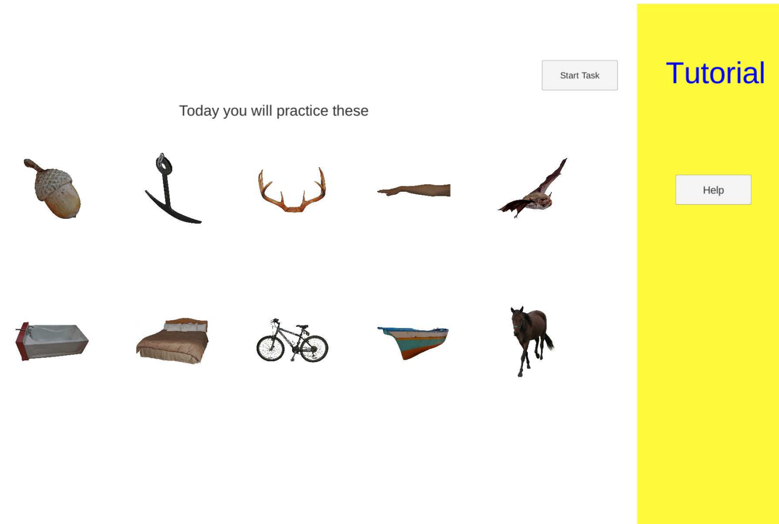

Figure 3. 1 - Screenshot of Stage 3.

Briefly, upon pressing on the button to start a trial, a picture appears at the center of the screen along with a red circle just below it. The center of the circle embeds a number that, starting from three seconds, counts down to zero. During this period, the microphone is turned off, and users have the possibility to prepare a response. When the countdown reaches zero, the microphone turns on and starts recording from the user's microphone. Users are informed about it by the fact that the circle below the picture turns from red to green, and in the middle of it the prompt ‘SPEAK’ appears (see Figure 3.2 for an example). The microphone stays on for 6 seconds. After, the recorded audio file is first saved on the webserver (SWR_3, SWR_7 in Table B.1), then sent to the ASR (which is hosted on a different server, physically located in our research laboratory at Dalhousie University) (SWR_10 in Table B.1). The ASR determines whether the content of the audio file has a satisfactory match with any of the words that make up the space of possible responses (see Appendix C for an explanation of this concept): if so, its output will be the word with the best match, otherwise it will be the word ‘SILENCE’. The software compares this result with the target word to determine whether participants named the word accurately or not. Feedback as to the correctness of the vocal response is delivered in the form of a green check, or red octagon, appearing on the screen. In case users receive positive feedback, they can move on to the next trial. Otherwise they can decide to repeat the current trial until they get it right, or to move on (SWR_9 in Table B.1).

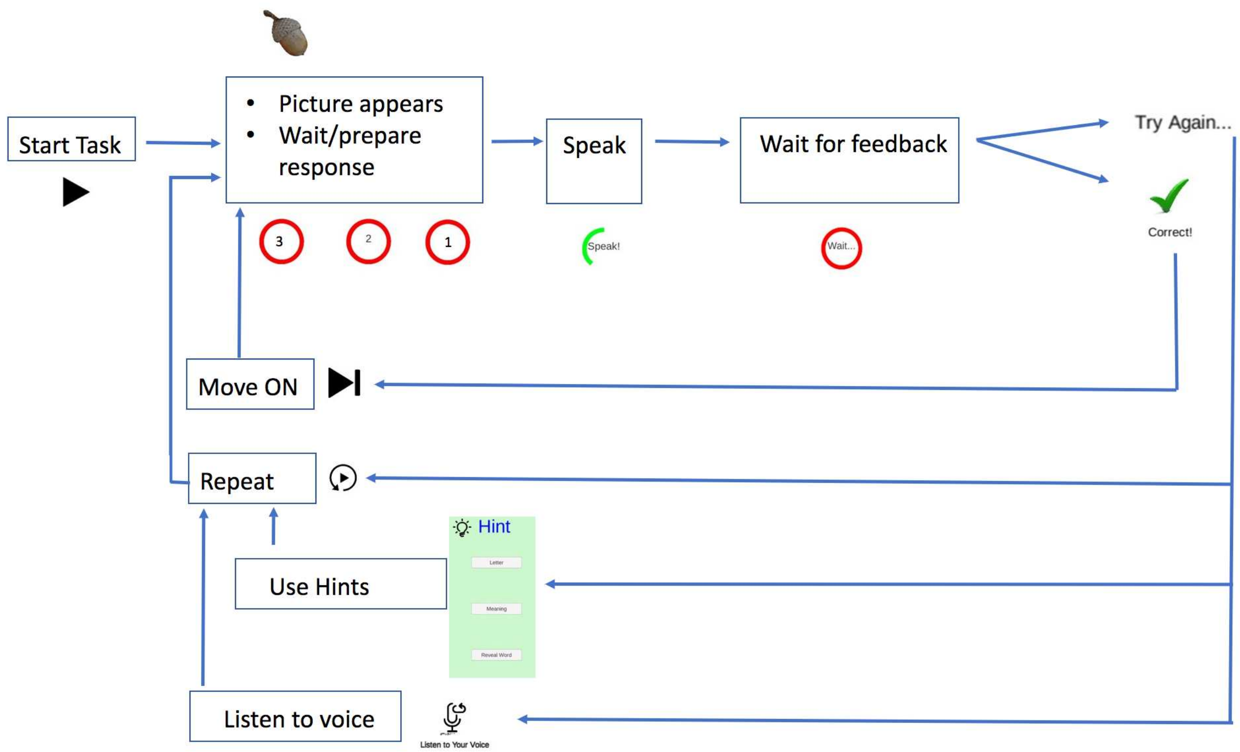

Figure 3. 2 - Schematic overview of the naming task, as implemented in this study.

If they choose to repeat the trial, the program offers the users the option to replay the audio of their speech that was recorded during the last naming attempt (SWR_8 in Table B.1), as this might help them identify the nature of their mistakes and attempt to correct them. Users also have the option to be provided with written cues to help find the right word (SWR_6 in Table B.1). This can be done immediately after the feedback has been displayed, and before the beginning of the next recording, such that there is time for participants to process the cue and prepare an answer before the microphone is activated. There are two kinds of cues: orthographic and semantic. The first provides information about how the word is written, and the second about its meaning. The cues are made available following a hierarchy that goes from less to more specific cues. Participants can make use of the orthographic cues by clicking on the button called 'Letter' that appears on the left side of the screen, under the green banner titled 'Hints'. Clicking on this button will reveal, in order, the following information about the target word: number of syllables (N); first letter; first syllable; first two syllables; first N-1 syllables; full word. The latter cue can be accessed directly by clicking on the button called 'Reveal Word'. Finally, semantic cues can be accessed by clicking on the button called 'Meaning'. When using the semantic cues, the hierarchy will unfold in this sequence: category; skill/feature/function; a high frequency sentence ending with the target word, where only the first letter of the latter is revealed. For example, assuming that the target word is 'acorn', the sequence of cues will be: 'It is a fruit'; 'It is the fruit of the oak'; 'the oak grows from a small *a* _ _ _ _'.

A total of 9 different words are presented during each session.

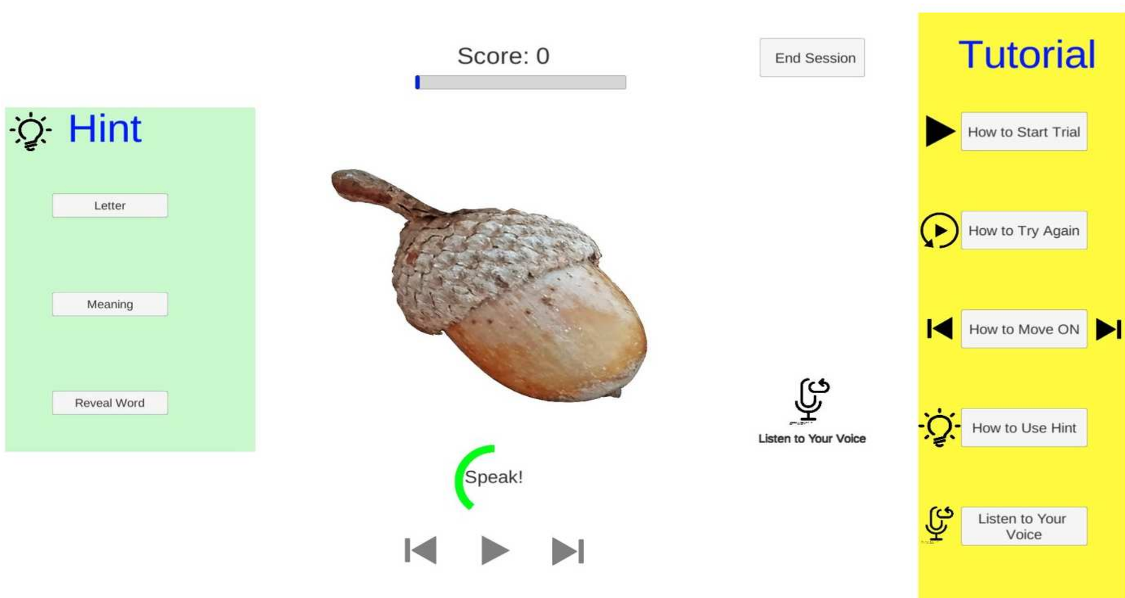

Figure 3. 3 - Overview of the User Interface while participants are engaging with the naming task.

Stage 5: Review of Item-Level Performance. Once users have gone through all the words, they are given the opportunity to review the material that they had just practiced (SWR_12 in Table B.1). Here, all trained items are presented together, with information about whether they were named correctly or not shown below their picture: a green bar if the item was named correctly, red otherwise (see Figure 3.4). Users can then decide to give another try to the words that they could not say correctly. The item-level accuracy score from Stage 4 is updated with that of Stage 5. For example, if at the end of Stage 4 a user had correctly named four words, and during Stage 5 they named one additional word correctly relative to Stage 4, the score for that session would be 5.

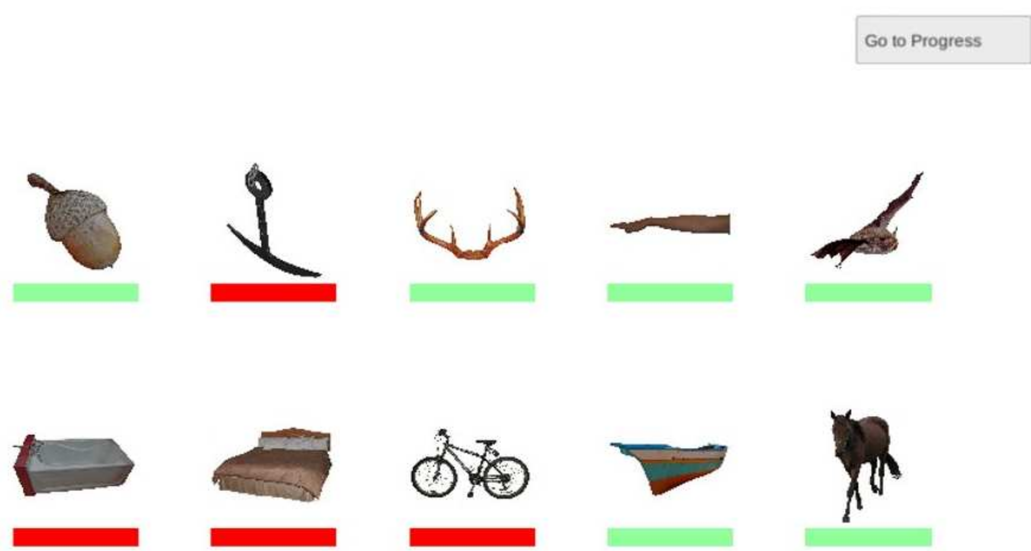

Figure 3. 4 - Appearance of the User Interface during Stage 5

Stage 6: Summary of Progress. Once users are satisfied with their performance on Stage 5 and they are ready to quit the session, they move to Stage 6 where they are provided information about the total scores that they obtained throughout all sessions (SWR_13 in Table B.1, see figure below).

Figure 3. 5 - Appearance of the User Interface during Stage 6, at the end of Day 2

Stage 7: Quit Program. Finally, users can quit the program by closing the browser page.

The Current Study

This study set out to assess the usability of the first prototype of a program that is intended to allow people with aphasia to self-administer therapy targeting word-finding and naming skills. Usability is typically understood as a characteristic of a product that allows users to achieve their goals while easily learning how to use the product autonomously, in the least possible amount of time, and satisfactorily (International Standard Organization., 2016). Therefore, usability studies such as the one presented in this chapter aim to inform the design and development of the product by providing information about: a) whether the product meets the needs that it was conceived to; and b) whether there are any design aspects that should be adjusted to make the product easier to learn and use. Doing so requires a detailed analysis of how users navigate the program via the user interface, and whether the materials that are intended to provide users with support are effective.

As discussed in the introductory chapter, the usability of the program is also strongly affected by users' perception of the accuracy and usefulness of the ASR. People with aphasia might find it difficult to detect errors in their speech (Ellis et al., 1983; Maher et al., 1994; Marshall et al., 1985, 1994; Marshall & Tompkins, 1982; Nickels & Howard, 1995; Sampson & Faroqi-Shah, 2011; Schwartz et al., 2016), and these are the instances where the ASR would provide more value to the user, provided that its output is considered trustworthy. However, while normally a person with aphasia would confidently trust their SLP over their self-assessment of their vocal output, this is not necessarily the case when this judgment is made by a computer program, even one yielding excellent performance in determining the accuracy of single-word utterances, would be trusted by a person with aphasia. For one, the amount of trust that someone puts in somebody else's judgment, opinion, or recommendation is strongly influenced by the perceived level of expertise of the interlocutor (Chaiken & Maheswaran, 1994), even when this is a technology (Reeves & Nass, 1996; Wienrich et al., 2021). This skepticism could be further exacerbated by the occasional mistakes made by the ASR. A failure of an ASR perceived as trustworthy by users could have detrimental effects: at best, users might ignore the feedback and rely just on their ability to monitor their speech, which is detrimental in case self-monitoring abilities are poor; at worst, a distrust in the ASR, coupled with a failure to ignore its feedback, can frustrate users and distract them from the task at hand, potentially even discontinuing use of the program.

As far as we know, the only study that investigated the performance of ASR while providing real-time feedback to people with aphasia, and gauging their opinions afterwards, is the study conducted by Ballard and colleagues (2019). In this study, the authors integrated an off-the-shelf ASR into a program designed to allow people with aphasia to practice word-finding and naming skills. Although the accuracy of the ASR varied substantially across participants (between 65% and 82%), all of them benefited from using the application (as demonstrated by the improvement in naming performance observed when compared pre- to post-training naming accuracy on treated items) and expressed no concerns about the reliability of the ASR. If anything, all participants were extremely positive about the overall usefulness of the program and about continuing to use it, with one participant saying that he really liked the fact that he could receive feedback from the ASR.

In this study, we aimed to assess participants' opinions about the accuracy and usefulness of the ASR more directly, and to assess the usability of the first prototype of CBAT by analyzing how they interacted with it, with the following objectives in mind:

1. To identify any design feature that might negatively affect the usability of the prototype, along with the events that trigger the need for support
2. To assess the clarity and usefulness of the user interface, including the material that is made available to users to learn how to use the program and to support them when needed, and to identify what parts need to be improved
3. To assess users' perceptions about the accuracy and usefulness of the ASR
4. To provide a quantitative measure of the accuracy of ASR responses relative to that of human raters

METHODS

The procedures of this study were approved by Dalhousie University Research Ethics Board, and all participants provided written informed consent to take part in this study.

Participants & Inclusion Criteria

To participate in this study, participants had to satisfy the following inclusion criteria:

- Being at least 18 years or older.

- Present at time of enrollment with "mild" to "moderate" expressive aphasia (cut-off score of 32), as defined by their score on the Western Aphasia Battery – Revised (WAB–R) (Kertesz, 2007).
- Aphasia onset at least 6 months prior to enrolment in the study.
- Have a partner who could assist them during the study.
- Being a fluent English speaker prior to onset of aphasia.
- Have the ability to understand and follow instructions necessary to complete the experiment.
- Have normal or corrected-to-normal vision and hearing.
- Have a computer or tablet with a minimum 9.7 inches.
- Have an internet connection.

Exclusion criteria were:

- Presence of severe aphasia. This criterion was defined because such level of impairment would have made the interviews conducted to gather feedback on usability very difficult to complete within a reasonable amount of time.
- Primary progressive aphasia.
- Other neurological illnesses in addition to aphasia.

Eight participants (3 females) took part in the study (see more details on the demographics in table 3.1). One (S8) could not complete the study because the audio recordings made while she interacted with the program were corrupted.

Measures

WAB-R – Part I

Part I of the WAB-R is designed to estimate the severity of a person's deficit along four domains of communication: spontaneous speech; auditory verbal comprehension; repetition;

naming and word-finding. A composite score (Aphasia Quotient, AQ) is derived by summing the scores obtained in these four domains, with spontaneous speech being weighed twice as any other score. The composite score is used to determine the level of severity along four broad categories: severe (AQ below 31.3); moderate (AQ between 31.3 – 62.6); mild (AQ between 62.6 – 93.7); normal (AQ above 93.7) (Kertesz, 2007). Given that this study was conducted remotely, the examiner administered the digital version of the WAB-R, which allows researchers to access all the materials from their computer and to present them to participants by sharing their screen.

Background Questionnaire - Questionnaire on the Experience with Technologies and Computer-Based Therapies

The Background Questionnaire (see Appendix B) was designed to gather information about (a) participants' background (age, onset and cause of aphasia) (b) and participants' day-to-day experience with technology and computer-based therapies for aphasia (i.e. pros and cons of the devices/program they use the most, barriers encountered while performing most important tasks, etc.). We decided to collect this information to gather insights that, by complementing the information obtained by analyzing participants' interaction with the program, could help us better determine how to optimize the usability of the prototype.

System Usability Scale (SUS). The SUS (see System Usability Scale, Appendix B) is a 5-point Likert scale including 10 statements designed to probe users' opinions about the usability of a given product (Brooke, 1986). Using this scale, a score of 1 corresponds to a strong disagreement with a specific statement, while 5 corresponds to a strong agreement. Six statements with a positive connotation about usability (i.e. *I would use the program frequently*), were alternated with 4 statements with a negative connotation (i.e. *The program is too complex to use*), using the same scale to determine how much a person agreed with them.

Table 3. 1 - Demographics

Subject	**Age**	**Sex**	**Time Since Stroke (years)**	**WAB-R scores**					
				SS	**AVC**	**R**	**NWF**	**AQ**	**Profile**
S1	66	M	9	19	9.75	9.1	9.5	94.7	Anomic
S2	57	F	12	19	9.95	9.6	9.3	95.7	Anomic
S3	72	M	14	13	8.85	6.2	7.9	71.9	Conduction
S4	64	M	9	11	8.15	6	5.9	62.1	Conduction
S5	65	M	3	20	10	9	9.5	97	Anomic
S6	70	F	3	18	9.5	7.4	9.2	88.2	Anomic
S7	47	M	5	11	9	6.2	5.7	63.8	Conduction
S8	57	F	2	19	9.9	9.8	8.6	94.6	Anomic

Note. SS: Spontaneous Speech; AVC: Auditory Verbal Comprehension; R: Repetition; NWF: Naming & Word-Finding; AQ: Aphasia Quotient.

Scores obtained from the statements with a positive connotation were converted from a scale 1-5 to a scale 0-4 by subtracting 1 unit to them. Scores obtained from the statements with a negative connotation were converted into a scale 0-4 by subtracting them from the number 5. This is done to ensure that, ultimately, when summing them up, the minimum and maximum values correspond, respectively, to the lowest and highest possible judgments on usability. Since the sum of individual scores in the range 0-4 would range between 0-40, a 2.5 multiplying coefficient was applied to scale this range to 0-100, which is easier to interpret.

Usability Questionnaire – Questionnaire on the Usability of the Program. To gather participants' views on the usability of the program we designed a questionnaire addressing the quality and usability of specific aspects of the prototype, including the video tutorials, images and audio, and the ASR (see usability questionnaire, Appendix B).

Procedure

Preliminary Phase

Before obtaining consent to participate in the study, the examiner and the prospective participant met for a preliminary video call to ensure that the conditions were optimal to conduct the study via videoconference. The examiner verified that the internet connection was stable, and that the audio and video were of acceptable quality. Whenever an issue arose, the examiner instructed participants on how to fix it. For example, if participants' camera faced a source of intense light that prevented seeing their face clearly, they were instructed to turn the camera enough to alleviate the issue. The partner of the participant was tasked to facilitate the sessions, with two main responsibilities: a) to help in case the quality of communication degraded (i.e. help with adjusting the audio/visual settings; resetting the internet connection if it did not function properly; facilitating communication between the examiner and the person with aphasia); and b) to assist with the scoring of two of the tasks of the WAB-R (administered during session 1, after participants have provided consent to take part in the study). One of such tasks is the "Auditory Word Recognition" task, where participants have to point, upon the request of the examiner, at one of 6 objects (items, colored boxes, numbers, letters etc.) ordered along a 3 x 2 grid. Since the examiner could not see which of the objects the participant was pointing at, the facilitator was tasked to communicate this information. To prevent participants from getting influenced by knowing about whether they were performing the trials accurately or not, information on task accuracy was masked by assigning to each position on the 3 x 2 grid a number from 1 to 6, starting from top-left, moving to the right, then going back to bottom left and moving to the right again. This way, the facilitator could quickly communicate to the examiner which item was pointed at, while hiding information about accuracy from the participant. In one instance the items were not placed in a grid-like fashion. In this case, the facilitator was asked to record which item was pointed at on a piece of paper, and to report the outcomes to the examiner at the end of the exercise. Facilitators were also needed to assist with the “Sequential Commands” task. In this task participants had to use three different objects (a pen, a comb and a book) to perform specific actions in a given sequence (i.e. Point to the pen with the book). Again, the facilitator was tasked to verify and report, for each trial, whether the correct actions were performed in the right sequence. At the end of this session, an electronic copy of the consent form was sent to prospective participants, via email. This was done to allow them to take the time to carefully read the form (if needed, with the support of a person they trusted) and to prepare a list of questions to ask at the beginning of the next session, before making a decision on whether or not they wanted to provide consent to take part in the study.

Session 1 – Consent Form, Standardized Testing and Background Questionnaire

At the beginning of this session, participants were asked whether there was anything regarding the content of the consent form that needed to be clarified. After clarification was provided, participants were given two options to provide formal consent. One was to insert an electronic signature on the copy they received via email, and to share it back with the person conducting the experiment. If this was not possible, the consent was gathered verbally.

Participants were first asked whether they provided consent to take part in the study. If they did, they were informed that from that moment on the video-conversation would be recorded, and that they had to confirm again that they provided consent to take part in the study. This way a proof that consent was given could be stored.

Afterwards, the examiner administered the digital version of the WAB-R. If participants met all inclusion criteria, the examiner also administered the Background Questionnaire (see Appendix B).

Session 2 - Usability Study

At the beginning of this session the examiner provided a short demonstration on how to use the prototype of the program. Then, participants started using the program on their computer, while sharing their screen. This allowed the examiner to video record the entire session for the purpose of data analysis. Once participants completed the tasks, the examiner administered the SUS and the usability questionnaire (see Appendix B).

Data Analysis

Usability

The first objective of this study was to diagnose the main obstacles that limit the usability of the program. It was therefore crucial to identify *which* aspects of the program had to be adjusted, and *how*. To do that, users' experience with the program was broken down into a series of independent tasks (see list of tasks for prototype v1.0, Appendix B for details), thus making it possible to quantify usability separately for each task.

The definition of usability that is provided by the International Organization for Standardization (ISO; 2016) puts a premium on attributes of a product/service that allow users to achieve their goals (a) independently and with high accuracy (effectiveness); (b) using the least amount of resources - minimizing mistakes (efficiency); (c) while being satisfied with the way the product meets their needs (satisfaction). These three attributes can be measured quantitatively, and are typically targeted as primary outcome measures in usability studies (Hill & Breslin, 2016; Simic et al., 2016). In this study, **Effectiveness** was quantified using a 4-point scale wherein scores were assigned according to the extent users could successfully complete a given task independently: 0 (the user was not able to start and complete the task, and the researcher had to assist the user or their partner on how to move on); 1 (the user started the task but was not able to complete it); 2 (the user completed the task with assistance from the researcher, or from watching a video tutorial); 3 (the user completed the task independently). For each task, the researcher took notes on all requests of support made by the user in order to better identify features of the prototype that had to be redesigned. **Efficiency** was quantified by counting the number of times participants pressed the wrong button as they interacted with the prototype. For example, if a participant expressed the intention to repeat a given trial, but instead pressed the button to move on to the next trial, that would be considered as an error. As mentioned before, this information was gathered by analyzing the videos recorded while users shared their screens. **Satisfaction** was quantified by administering the SUS after the user had quit the program. In addition to the metrics described above, we used the usability questionnaire (see Appendix B) to obtain qualitative information that may help with improving the design of the program. The satisfaction scores obtained from the SUS range between 0-100. Scores below 60 are considered to reflect poor satisfaction with the product, and major concerns regarding its usability, while scores above 80 are considered to reflect good satisfaction. Scores between 60-80 indicate moderate concerns with the usability of the product.

The usability questionnaire was also used to assess participants' opinions about clarity, usefulness and areas to improve upon in the instruction manual (second objective), and about their perceptions of the accuracy and usefulness of the feedback received from the ASR (third objective).

Analysis of the Performance of the ASR

Following the same rationale described in Appendix C, we decided to employ signal detection theory (Donaldson, 1992) to assess the performance of the ASR. The judgments of two human raters were used as ground truths to judge the performance of the ASR against.

RESULTS

Participant Language Profile

Four of the seven participants that completed the study had mild aphasia (S1, S5 and S6, all with AQ>88), while the others obtained AQs within the range where people that do not have aphasia typically fall (see Table 3.1). These three participants would likely not benefit from using the program that we developed. Nonetheless, for the sake of getting feedback on the usability of the prototype, we believe that they can provide valuable information as they have a unique perspective on the specific challenges that people with aphasia face when using technology, and so their data were included in analysis.

Usability

Tables 3.2-3.8 show results for each of the outcome measures. Since the analysis of results consistently revealed an association between the outcome measures and participants' AQs, to facilitate the interpretation of results participants are ranked by the severity of their aphasia, from less to more severe.

Effectiveness

Table 3.2 shows effectiveness scores for each stage of the program. To facilitate the interpretation of results, effectiveness was color-coded: green for a score of 3; yellow for a score of 2; orange for a score of 1; red for a score of 0.

Table 3. 2 - Effectiveness

Task	**Participant**						
	S5	**S2**	**S1**	**S6**	**S3**	**S7**	**S4**
Access web and type web address	3	3	3				
Welcome scene	3	3	3	3	2*	3	2*
Login into personal account	3	3	3	3	2*	2*	2*
Overview of material	3	3	3	3	2*	2*	2 (1)
Naming task							
Star trial	3	2 (1)	3	3	2 (1)	2 (1)	2 (2)
Use cues to help with naming					1	2 (1)	2*
Use 'Try Again' function		3			3	2 (1)	2*
Use 'Replay Audio' function						2 (1)	2*
Move to next item	3	2 (1)	3	3	3	2 (1)	2*
Exit task when completed	3		3	3	2*	2*	2*
Review item-level performance	3	2 (1)	3	3	2*	2*	2*
Quit program							2*

Note. Numbers within parentheses indicate the number of times that a participant used the video tutorial to tackle a specific task. The presence of an asterisk indicates that participants requested and received a suggestion from the examiner on how to move on with that task.

It is important to note that some tasks were never executed by some participants. For example, four participants (S3, S4, S6, S7) launched the program by clicking on its URL (which was shared with them via the chat function on zoom) instead of inserting it on the address bar, and therefore we could not measure effectiveness for this task. Similarly, some participants did not use some or any of the functionalities designed to support users to get better at naming words, like the *hints* or the *'replay your voice'* function. In such cases, we did not insert any score in the corresponding cell in Table 3.2. Participants with AQs in the range that is considered to be mild-normal could complete all the tasks with maximum levels of effectiveness.

The only exception was S2, who used the video tutorial to remember how to start a trial and how to move to the next one. In contrast, participants with mild deficits (S3, S4 and S7, with AQ<72) frequently relied on the support of the video tutorials, or on requesting a suggestion from the examiner, to complete the task at hand. Except for S3, who could not successfully use the 'hints' function to help with naming, the provision of support via the video-tutorials or via a nudge by the researcher was sufficient for the other participants to complete a task by themselves, whenever they encountered difficulties. In contrast, participants with lower AQ scores needed help to complete most of the tasks that they engaged with (S3 relied on help to complete 6 out of 9 tasks; S4 relied on help for all tasks (11/11); S7 relied on help in 9 out of 10 tasks).

Efficiency

As anticipated, a participant's AQ was related to efficiency scores as well (see Table 3.3): participants with higher AQ scores made no mistakes (except for S2 who made 1 mistake), while those with lower scores made some mistakes (S3: 2; S4: 3; S7: 1). All mistakes were made when engaging with the naming task.

Satisfaction

The composite scores obtained from the SUS reveal that a person's satisfaction was also associated with AQ. Participants with AQ > 88 had SUS scores > 90; participants with AQ < 72 had SUS scores < 78 (see Table 3.4). Table 3.5 breaks down participants' responses at the level of single statements.

Table 3. 3 - Efficiency

Task	Participant						
	S5	**S2**	**S1**	**S6**	**S3**	**S7**	**S4**
Access web and type web address	0	0	0				
Welcome scene	0	0	0	0	0	0	0
Login into personal account	0	0	0	0	0	0	0
Overview of material	0	0	0	0	0	0	0
Naming task							
Star trial	0	0	0	0	1	0	1
Use cues to help with naming					1	0	0
Use ‘Try Again’ function		0			0	1	1
Use ‘Replay Audio’ function						0	0
Move to next item	0	1	0	0	0	0	1
Exit task when completed	0		0	0	0	0	0
Review item-level performance	0	0	0	0	0	0	0
Quit program							0

Note. The numbers within the cells indicate the number of mistakes made during the associated task.

Table 3. 4 - System usability Scores

	Participant						
	S5	**S2**	**S1**	**S6**	**S3**	**S7**	**S4**
Score	92.5	92.5	90	100	67.5	77.5	45

Note. The green background is associated with scores ranging between 80-100, reflecting good satisfaction with the product; yellow is associated with scores ranging between 60-80; orange is associated with scores below 60.

Table 3. 5 - System Usability Scores - breakdown of results by statement

Statement	Participant						
	5	**2**	**1**	**6**	**3**	**7**	**4**
I would use the program frequently	4	5	3	5	4	4	3
The program is easy to use	5	5	4	5	5	5	3
The various functions in the program were well integrated	4	5	5	5	4	4	4
There is consistency in the program	5	5	5	5	4	5	5
Most people would learn to use the program very quickly	4	5	5	5	2	4	3
I felt very confident using the program	5	5	5	5	4	4	1
The program is too complex to use	1	1	1	1	1	1	3
I would need the support of a technical person to be able to use the program	1	4	2	1	3	3	5
The program is cumbersome to use	1	1	1	1	2	2	3
I needed to learn a lot of things before I could get going with the program	1	1	1	1	4	3	4

Note. In line with the logic of the SUS, higher numbers reflect higher agreement with the statement, on a scale that goes from 1 (highly disagree) to 5 (highly agree), with 3 being equivalent to 'neither agree not disagree'. Statements 7-10 have a negative connotation with respect to the usability of the program. Therefore, in these cases low scores (1 and 2), which reflect disagreement with the statement, are highlighted in green. Vice versa for high scores such as 4 and 5. To facilitate the interpretation of results, scores reflecting agreement/disagreement with a statement with a positive/negative connotation with a statement were highlighted in green; vice versa, in orange. Yellow is associated with response that neither agree nor disagree with a statement.

Clarity and Usefulness of the Video Tutorials

Of the seven participants that completed the study, three did not use any of the video tutorials, hence they could not respond to the three questions targeting their opinion on the quality and usefulness of that material. The other participants provided positive opinions about the usefulness of the video tutorials: three said that they were useful, while one (S4) expressed uncertainty on this matter (see Table 3.6).

Regarding the clarity of the instructions, two participants thought that they were clear, while two expressed uncertainty. S4 discussed more in detail what he thought could be improved in the video tutorials. He pointed out that: a) in the stage of the program dedicated to the naming task, there were too many buttons associated with the video tutorials, and this contributed to creating confusion in a stage that already afforded several functionalities; b) there were some problems with the videos, such as for example the voice stopping in mid-sentence, or the video tutorial stopping before completing the demo of a specific function; c) it was not always clear from the video tutorials which button in the user interface was associated with the function of interest to the user.

Participants' Perception of the Accuracy and Usefulness of the ASR

All participants thought that the feedback that they received from the ASR was useful (see Table 3.7). Five participants thought that the ASR provided accurate feedback on their naming responses, while two participants thought that this was only sometimes the case: S1 and S4 expressed insecurity on this matter because they thought that they named all words right. Regarding S1, the ASR provided negative feedback to his response on two instances, results that were in line with the scores provided by the human raters (see Table 3.8).

Table 3. 6 - Opinions on the usefulness and clarity of the video instructions

	Participant						
	S5	**S2**	**S1**	**S6**	**S3**	**S7**	**S4**
Useful							
Clear							

Note. The green color reflects participants' affirmative response to the perceived clarity and usefulness of the video-instructions. Yellow reflects participants' opinion that the video-instructions were only sometimes clear/useful. Whenever a cell has no color, it means that that participant did not use the video-instructions, and therefore could not provide an opinion on their clarity/usefulness.

Table 3. 7 - Participants' opinion on the accuracy and usefulness of the ASR

	Participant						
	S5	**S2**	**S1**	**S6**	**S3**	**S7**	**S4**
Accurate							
Useful							

Note. The green color reflects participants' affirmative response on the perceived accuracy and usefulness of the ASR. Yellow reflects participants' opinion that the ASR was only sometimes accurate/useful.

The discrepancy between the perception of S4 on the accuracy of his responses and the output of the ASR could be explained by the fact that in some cases S4 named a picture before the microphone was activated, or only after it turned off, hence the ASR did process an incomplete or absent response as an incorrect answer (as did the human raters, see Table 3.8).

Comparison of the Judgement of the ASR vs. Human Rater

Table 3.8 shows that 64 out of 69 responses of the ASR matched with those of the human raters. Results of analysis conducted by applying signal detection theory (see Table 3.9) revealed that the ASR showed good capability to discriminate (a'= .95), and a low bias score (β''=-.23).

Users' Opinions on how to Improve the Prototype

Two participants pointed out specific ways to improve upon the usability of the prototype. S7 said that it would be desirable to have the option to listen to the correct pronunciation of the target word, in addition to the option to listen to one's own last response. S4 recommended two changes. First, to adjust the position of the login button to be consistent with design principles adopted by other popular apps, namely locating the login button below the naming and password input fields, and not above.

Table 3. 8 - Confusion matrices for the performance of the ASR relative to human raters

		ASR output	
		Correct	**Incorrect**
S1	**True Correct**	8	0
	True Incorrect	0	2
S2	**True Correct**	10	0
	True Incorrect	0	0
S3	**True Correct**	6	4
	True Incorrect	0	0
S4	**True Correct**	5	0
	True Incorrect	0	4
S5	**True Correct**	10	0
	True Incorrect	0	0
S6	**True Correct**	10	0
	True Incorrect	0	0
S7	**True Correct**	7	0
	True Incorrect	1	2

Note. Green and red cells indicate, respectively, responses that were correctly and incorrectly labeled by the ASR.

Table 3. 9 Accuracy and bias of the ASR

TP	**FP**	**FN**	**TN**	**a'**	**β''**
56	1	4	8	.95	-.23

Note. A value of a' near 1 indicates good discriminability, while values around .5 indicate chance performance. Values for β'' near 0 indicate no bias; positive values indicate a bias towards scoring a response as incorrect, with a value of 1 indicating that all responses were scored as such. Vice versa for negative values (Makowski, 2018). True/false positives/negatives relate to the total among all participants.

Second, to extend the amount of time where the microphone is active, as the current time window (6 seconds) was judged to be too short to allow users to fully articulate a word.

DISCUSSION

This chapter provides an overview of the first prototype of a web-based program designed to allow people with aphasia to self-administer therapy targeting word-finding and naming skills. Importantly, this program features an ASR. This enables the provision of real-time feedback on naming accuracy, which is a key ingredient to drive recovery of function. A group of people with aphasia took part in a usability study whose goal was to identify the main design factors limiting the usability of the program, to assess the effectiveness of the material that was made available to support users, and to assess the accuracy of the ASR both objectively, and according to users' perceptions. Below is a discussion of the results pertaining the objectives of the study.

The Major Obstacles Affecting the Usability of the Prototype

To identify features of the program that limited its usability, we integrated an analysis of the metrics of effectiveness (the level of support needed to complete the tasks) and efficiency (the number of mistakes made while completing the tasks) with data pertaining users' satisfaction. The analysis of the effectiveness and efficiency scores clearly revealed that everybody was able to successfully go through all the stages of the program either independently or with the support of a video tutorial, while making few errors. Both effectiveness and efficiency were associated with participants' AQs: the lower the AQ, the poorer the effectiveness and efficiency scores. In fact, while participants with higher AQ scores (>90) successfully completed all their tasks without the need for support and making no mistakes, all those with lower AQ scores (<72) needed some kind of support to complete the task, made some mistakes, and admitted that they would need the help from a partner to use the program. This clear gap in performance between the two groups of participants could be due to deficits in cognitive functions such as memory, attention, and learning that are typically more pronounced in people with severe aphasia than in those with milder forms (Fonseca et al., 2018; Huang et al., 2022; Lee et al., 2020; Varkanitsa et al., 2023). In this regard, it is important to reiterate that participants were provided only a brief (about five minutes long) step-by-step demonstration on how the program works. While this demonstration was long and detailed enough for participants with high AQ scores to learn to use the program effectively, this was not the case for those with low AQ scores, wherein the possible presence of cognitive deficits could have interfered with understanding and retaining the instructions. This could also explain why these participants made mistakes only in the stage where they had to engage with the naming task - where they had to remember and understand several different choices that one could make after receiving feedback from the ASR. Nonetheless, in this study we did not test participants' attention, memory, or learning skills, nor their ability to understand and comprehend long conversations (the WAB-R tests auditory comprehension via short questions requiring a yes/no response, via a single-word comprehension task, and via a task that assesses a person's ability to follow short instructions). Therefore, whether the lower effectiveness and efficiency observed in participants with lower AQs can be attributed to deficits in cognitive skills remains, currently, only a hypothesis. The gap in performance between participants with mild vs moderate aphasia could also be explained by a difference in familiarity using CBAT, or with technology more in general. However, all participants taking part in this study reported regular use of their device of choice for this study, and participants with moderate aphasia reported having tried other similar programs in the past.

The breakdown of satisfaction scores presented in Table 3.5 provided three insights. *First*, participants with milder symptoms expressed high levels of satisfaction with the usability of the prototype. *Second*, participants with more severe symptoms of aphasia indicated that they did not find it easy to learn how to use the program and that they would benefit from support when using the program. *Third*, everyone (except for S4) regardless of severity of aphasia said that the program was well designed, and that they would use it frequently.

On the Usefulness and Clarity of the video tutorials

In general, among the four participants who relied on the support of the video tutorials to successfully complete a task, only one did not find them particularly useful, while two of them reported that the main bottleneck was the clarity of the instructions, along with some glitches due to the program interrupting the tutorials before their end.

Participants' Perceived Accuracy and Usefulness of the ASR

A key objective of this study was to assess participants' opinions on the accuracy and usefulness of the ASR. Results were promising, with all participants declaring that the feedback from the ASR was useful and, except in two cases, accurate. This result is encouraging, as it suggests that participants developed trust in the judgment that was made by the ASR, a necessary condition to ensure that the feedback is beneficial to users. Importantly, in line with what Ballard and colleagues (2019) reported, even the two participants that found the ASR sometimes inaccurate said that it was useful, suggesting that the few discrepancies between the output of the ASR and users' judgment on their accuracy did not affect their trust in it.

Nonetheless, there are two reasons to be cautious when interpreting this finding. First, although participants were instructed to do their best to name words correctly, and to repeat a given item multiple times in case they got negative feedback from the ASR or if they felt that they named it incorrectly, the goal of this study was only to test the usability of the program and not its effectiveness at improving word-finding and naming skills. Therefore, participants' judgment on the usefulness of the ASR cannot be attributed to evidence in the usefulness of the feedback, but only to a positive impression that they formed during their brief experience using the program. Second, most participants named all words correctly on their first attempt, and did not need to rely on the ASR to determine whether to try the same item again or not. Therefore, related with the point made above, there were limited opportunities for participants to learn from the feedback received from the ASR.

Accuracy of the ASR

Results of our analysis comparing the output of the ASR with that of human raters revealed that the ASR was excellent at scoring words, with high levels of accuracy and low bias. Analysis of the responses that the ASR scored incorrectly revealed that, in all four cases where a correct response was labelled as incorrect, the audio files that were recorded were of poor quality, although intelligible to human raters. Therefore, it is possible that such mistakes would not have occurred had the audio recording been cleaner. On the other hand, the specificity score is based on nine samples only (due to the low number of naming errors made), and therefore more data is needed before we can draw any conclusions on the ability of the ASR to accurately detect mistakes in naming.

Although these results are very promising and are in line with the best outcomes of ASRs used for scoring utterances of single words reported in related studies (Abad et al., 2013; Ballard et al., 2019; Barbera et al., 2021), a more extensive testing (i.e. more participants, and more samples per participant) is needed to confirm them.

Lessons Learned and Next Steps

The study has a few take-home messages. First, given the intended use of this first prototype, there are only a few minor aspects of the design that require some revision, such as a) editing the login stage to move the login button below the username and password input fields; b) editing the video tutorials.

Second, the usability of the first prototype proved to be satisfactory, according to most participants. As a next step, the usability of the improved prototype will be tested not just following a single-session study, but both before and after participants will have had the opportunity to use the program to self-administer naming therapy over a number of training sessions. This will allow us to investigate users' opinions after having had ample time to familiarize with it, and to properly understand its strengths and weaknesses from a usability perspective.

Third, to have a more reliable estimate about how the ASR performs when integrated in a CBAT, and to what extent people with aphasia trust it as an assistant, more extensive online testing needs to be conducted. The next study will make a step further towards this goal, as participants will be asked to use the program over more, and longer, sessions.

Fourth, and finally, the instructions that will be provided by the examiner before a participant starts using the program should be further simplified to make sure that more people with aphasia can successfully retain them. In this case, simplification could entail both reducing the amount of content, and delivering it over a longer span of time.

Limitations

This study presented some limitations which we plan to overcome in the subsequent one. For one, among the pool of participants that took part in the study, only about half of them (three out of eight) had an AQ score that was below the range where a person would be considered to have no aphasia (that being said, participants with AQ scores above 90 points, where one would be considered to have no aphasia, did show obvious signs of residual language deficits). Therefore, a more reliable test of the usability of the program would need to include far more individuals with pronounced language deficits, those that would ultimately better match the profile of the intended user of this program. However, it is important to stress that we encountered difficulties with participant recruitment in this study, and in order to execute it within a reasonable amount of time we had to settle with recruiting some people that do not match the ideal participant's profile.

Finally, another limitation of this study is that during the session there were too few items to adequately test the accuracy and perceived usefulness of the ASR during use.

CONCLUSIONS

The results of this study highlight that, in its current form, most participants a) could use the program to achieve a simple goal independently or with minimal support; b) expressed positive opinions regarding about the usefulness of the ASR, the main innovation relative to state-of-the-art CBAT that we wanted to test out; c) expressed satisfaction with their experience, and said that they would use improved version of the program to self-administer therapy. Nonetheless, there are important aspects that require improvement, such as for example the clarity of the video tutorials that are available to users for support.

Overall, these results are very promising and chart a clear path ahead on how to develop the next version of the program.

REFERENCES

Abad, A., Pompili, A., Costa, A., Trancoso, I., Fonseca, J., Leal, G., Farrajota, L., & Martins, I. P. (2013). Automatic word naming recognition for an on-line aphasia treatment system. *Computer Speech and Language*, *27*, 1235–1248.

Ballard, K., J, Etter, N., M., Shen, S., Monroe, P., & Tien Tan, C. (2019). Feasibility of Automatic Speech Recognition for Providing Feedback During Tablet-Based Treatment for Apraxia of Speech Plus Aphasia. *American Journal of Speech-Language Pathology*, *28*, 818–834.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459.

Barbera, D. S., Huckvale, M., Fleming, V., Upton, E., Coley-Fisher, H., Doogan, C., Shaw, I., Latham, W., Leff, A. P., & Crinion, J. (2021). NUVA: A naming utterance verifier for aphasia treatment. *Computer Speech and Language*, *69*, 1–25. https://doi.org/10.1016/j.csl.2021.101221

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *Cochrane Database of Systematic Reviews*, *6*. https://doi.org/10.1002/14651858.CD000425.pub4.

Brodeur, M. (n.d.). *The BOSS dataset*. https://sites.google.com/site/bosstimuli/

Brodeur, M., Guerard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. *PlosOne*, *9*(9). https://doi.org/10.1371/journal.pone.0106953

Brooke, J. (1986). *System Usability Scale (SUS)*.

Chaiken, S., & Maheswaran, D. (1994). Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgment. *Journal of Personality and Social Psychology*, *66*(3), 460–473.

Chapey, R. (2008). *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders* (5th ed.). Lippincott Williams and Wilkins.

Criptographic hash function. (n.d.). https://en.wikipedia.org/wiki/Cryptographic_hash_function

DigitalOcean. (n.d.). [Computer software]. DigitalOcean. https://www.digitalocean.com/

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General, 121*(3), 275–277. https://doi.org/10.1037/0096-3445.121.3.275

Ellis, A. W., Miller, D., & Sin, G. (1983). Wernlcke's aphasia and normal languagp processing: A case study in cognitive neuropsychology. *Cognition*, *15*, 111–144.

Fonseca, J., Raposo, A., & Pavao Martins, I. (2018). Cognitive functioning in chronic poststroke aphasia. *Applied Neuropsychology: Adult*. https://doi.org/10.1080/23279095.2018.1429442

Goodglass, H., & Wingfield, A. (1997). *Anomia: Neuroanatomical and Cognitive Correlates*. Academic Press.

Graves, A. (2012). Long Short-Term Memory. *Supervised Sequence Labelling with Recurrent Neural Network*, *385*, 37–45.

Hill, A. J., & Breslin, H. M. (2016). Refining an Asynchronous Telerehabilitation Platform for Speech-Language Pathology: Engaging End-Users in the Process. *Frontiers In Human Neuroscience*, *10*(640). https://doi.org/10.3389/fnhum.2016.00640

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780.

Huang, Q., Shu, T., Xu, W., & Chen, Z. (2022). Alerting, Orienting, Executive Control and Language Function in Post-stroke Aphasia Patients. *Rehabilitation Science*, *7*(1), 5–15. https://doi.org/10.11648/j.rs.20220701.12

International Standard Organization. (2016). *Application of Usability Engineering to Medical Devices*.

Jamal, N., Shanta, S., Mahmud, F., & Sha'abani, M. (2017). Automatic Speech Recognition (ASR) based Approach for Speech Therapy of Aphasic Patients: A Review. *AIP Conference Proceedings*, *1883*. https://doi.org/10.1063/1.5002046

Kertesz, A. (2007). *Western Aphasia Battery: Revised*. https://books.google.com/books/about/Western_Aphasia_Battery.html?hl=&id=bD3DxQEACAAJ

Le, D., Licata, K., & Provost, E. M. (2018). Automatic quantitative analysis of spontaneous aphasic speech. *Speech Communication*, *100*, 1–12. https://doi.org/10.1016/j.specom.2018.04.001

Lee, J. B., Kocherginsky, M., & Cherney, L. R. (2020). Attention in individuals with aphasia: Performance on the Conners' Continuous Performance Test – 2nd edition. *Neuropsychological Rehabilitation*, *30*(2), 249–265. https://doi.org/10.1080/09602011.2018.1460852

Maher, L. M., Gonzales Rothi, L. J., & Heilman, K. M. (1994). Lack of error awareness in an aphasia patient with relatively preserved auditory comprehension. *Brain and Language*, *46*, 402–418.

Makowski, D. (2018, March 29). Compute Signal Detection Theory Indices with R. *The PsychoBlog. Learn and Improve your R skills for Psychology.* https://neuropsychology.github.io/psycho.R/2018/03/29/SDT.html

Marshall, R. C., Neuburger, S. I., & Phillips, D. S. (1994). Verbal self-correction and improvement in treated aphasic clients. *Aphasiology*, *8*(6), 535–547.

Marshall, R. C., Rappaport, B. Z., & Garcia-Bunuel, L. (1985). Self-monitoring behavior in a case of severe auditory agnosia with aphasia. *Brain and Language*, *24*, 297–313.

Marshall, R. C., & Tompkins, C. A. (1982). Verbal Self-Correction Behaviors of Fluent and Nonfluent Aphasic Subjects. *Brain and Language*, *15*, 292–306.

Mel-frequency cepstrum. (n.d.). In *Wikipedia*. Retrieved March 26, 2023, from https://en.wikipedia.org/wiki/Mel-frequency_cepstrum

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

Nickels, L., & Howard, D. (1995). Phonological errors in aphasic naming: Comprehension, monitoring and lexicality. *Cortex*, *31*, 209–237.

NVIDIA Accelerates Real Time Speech to Text Transcription 3500x with Kaldi. (n.d.). Retrieved March 23, 2023, from https://developer.nvidia.com/blog/nvidia-accelerates-speech-text-transcription-3500x-kaldi/

Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hanneman, M., Motlicek, P., Qian, Y., Schwarz, P., Silovsk, J., Stemmer, G., & Vesel, K. (2011). The Kaldi Speech Recognition Toolkit. *Workshop on Automatic Speech Recognition and Understanding*. IEEE, Hilton Waikoloa Village, Big Island, Hawaii, US.

Reeves, B., & Nass, C. (1996). *The media Equation: How People Treat Computers, Television, and New media like Real People and Places.* Cambridge University Press.

Sampson, M., & Faroqi-Shah, Y. (2011). Investigation of self-monitoring in fluent aphasia with jargon. *Aphasiology*, *25*(4), 505–528.

Schwartz, M. F., Middleton, E. L., Brecher, A., Gagliardi, M., & Garvey, K. (2016). Does naming accuracy improve through self-monitoring ofe rrors? *Neuropsychologia*, *84*, 272–281.

Simic, T., Leonard, C., Laird, L., Cupit, J., Höbler, F., & Rochon, E. (2016). A Usability Study of Internet-Based Therapy for Naming Deficits in Aphasia. *American Journal of Speech-Language Pathology.*, *25*, 642–653.

Unity. (n.d.). [Computer software]. https://unity.com/

Varkanitsa, M., Godecke, E., & Kiran, S. (2023). How Much Attention Do We Pay to Attention Deficits in Poststroke Aphasia? *Stroke*, *54*, 55–66. https://doi.org/10.1161/STROKEAHA.122.037936

Wienrich, C., Reitelbach, C., & Carolus, A. (2021). Assistants in the Context of Healthcare Investigating the Effect of Perceived Expertise on the Trustworthiness of Voice Assistants, Providers, Data Receivers, and Automatic Speech Recognition. *Frontiers in Computer Science*, *3*, 1–12. https://doi.org/10.3389/fcomp.2021.68520

Chapter 4 – IMPROVED WORD-FINDING AND NAMING ABILITIES FOLLOWING SELF-ADMINISTERED COMPUTER-BASED THERAPY

ABSTRACT

This study followed up on the one reported in the previous chapter by: a) assessing the usability of an updated version of the prototype, and; b) conducting a preliminary assessment of the potential benefits that participants might gain from using the computerized therapy; c) characterizing in which cases, and to what extent, ASR-based feedback can be beneficial.

Twelve people with aphasia with mild-to-moderate naming deficits took part in this study. In line with results from the previous study, most participants could use the program independently or with minimal support, expressing positive opinions on its usefulness and usability. Only two participants were eligible to complete the study, using the program to self-administer therapy. Both of them benefited from training, showing improvements of 65 and 30% relative to performance in the baseline and control condition, 1 month after the end of therapy. However, no clear advantage from receiving the ASR-based feedback was observed. This was likely due to the fact that participants' ability to detect their mistakes was comparable to that of the ASR, and so there was no additional benefit to the feedback provided by ASR.

These results suggest that provision of external feedback might not always be necessary. However, to determine when ASR-based feedback is most needed, future studies must include larger pools of participants with a wider spectrum of self-monitoring and naming difficulties.

INTRODUCTION

The previous study set out to assess the usability of the first version of the prototype of computer-based aphasia therapies (CBAT). Specifically, the focus was on: a) identifying the main factors limiting usability and b) gathering information on how to further improve it; c) gathering participants' opinions on the accuracy and usefulness of the automatic speech-recognition (ASR) along with d) conducting a quantitative assessment of how accurately it could can detect errors in naming responses, relative to human raters. To this end, a short session was sufficient to get participants' first impressions on what needed to be improved. The study presented in this chapter was more ambitious: it aimed to be a more sensitive test of the strengths and weaknesses of the current version of the prototype by assessing its usability in the intended use case scenario – where people with aphasia use the program to self-administer therapy targeting word-finding and naming skills over multiple training sessions, from their home, whenever it suits them best. A description of how the user interface (and user database) were updated to integrate the feedback that we gathered in the previous study is provided in Appendix D.

Additionally, this study also presents a preliminary assessment of the potential benefits that participants might gain from using the prototype in the intended use case scenario. While it cannot be considered a proof-of-concept, as other studies have demonstrated the feasibility of adopting functionally similar products for the same purpose (Ballard et al., 2019; Palmer et al., 2019, 2020), currently there is scant empirical evidence on the usability and clinical benefits of CBAT featuring an ASR to deliver automatic feedback on naming accuracy. Therefore, this study will also contribute to the nascent knowledge-base on the usability and potential benefits of employing ASR to support people with aphasia in the recovery of their expressive skills.

To the best of our knowledge, only two other studies have contributed to this topic in some capacity. In a small-scale study involving five people with apraxia of speech and aphasia, Ballard et al. (2019) demonstrated that participants could successfully self-administer computerized training consisting of confrontational naming wherein feedback was provided by an ASR, and reap significant benefits from it. Following a training regime consisting of 16 sessions delivered over 4 weeks, participants could successfully name between 60% and 80% of the training material (20 words), relative to an average 23% during the baseline sessions. These gains were retained one month after the end of the training phase.

Similar results were reported from a randomized controlled trial conducted by Palmer et al. (2019, 2020) to assess whether computerized therapy[3] can serve as a cost-effective tool to complement the speech-language therapy that people with aphasia received directly from therapists (defined in this study as *usual care*). This study, in which 240 participants were randomly assigned to three groups (usual care; usual care plus daily computerized therapy targeting word-finding and naming skills; usual care plus attention control) showed that computerized therapy can serve as an effective adjunct to usual care, yielding an average 16.4% and 14.4% gain over the average performance attained by those in other groups, which did not differ significantly from baseline 6 months after the end of therapy.

In addition to these two studies, it is worth mentioning the work of Braley et al. (2021), who tested the efficacy of a digital intervention targeting language and cognitive deficits using the commercially available software Constant Therapy (*Constant Therapy*, 2022). While this

[3] This study employed the latest version of the StepByStep computer program. StepByStep has been available on the market since 2004 (Mortley et al., 2004), but only v5 (as of September 2023 the latest version available) features an ASR to deliver feedback on naming accuracy

product allows users to engage in confrontational naming under the guidance of ASR-based feedback, it also features several other exercises targeting naming and word-finding processes (i.e. picture-word matching; semantic categorization, etc.). While this study showed that participants benefited from training — as demonstrated by improvements on several language and cognitive tasks — the fact that they practiced the same items by engaging with multiple exercises makes it difficult to parse out the benefit, if any, of ASR-based feedback from the other features of the intervention.

Measuring the Benefits of Delivering ASR-Based Feedback: Is Feedback Always Necessary?

There is ample consensus among therapists regarding the importance of feedback for driving recovery of language skills (Chapey, 2008; Simmons-Mackie et al., 1999). In this context, feedback is thought to serve, mainly, to provide information on response accuracy that reinforces correct retrieval, or to signal the need to adjust responses towards the intended target (Ballard et al., 2019; Chapey, 2008; Nunn et al., 2023; Nunn & Vallila-Rohter, 2023; Palmer et al., 2019, 2020; Swales et al., 2016). This is all the more important when considering that people with aphasia might fail to realize when they make a mistake (Ballard et al., 2019; Chapey, 2008; Nunn & Vallila-Rohter, 2023; Palmer et al., 2020). Some studies (de Grosbois et al., 2023) have worked around this problem by presenting participants with the written or spoken version of the target word immediately after their naming attempt. The rationale behind this choice was that, this way, participants had a stronger reference to judge their response against, which would presumably attenuate the effect of their poor self-monitoring skills by helping them to identify when they made a mistake. While results from this study were promising, in that participants could show a stable 25% increase in naming performance (in a set of 20 words) at the end of training, which was maintained up to 5 weeks after the end of training, this task largely amounts to repetition training, at least for the words that participants identified as named incorrectly. This approach contrasts with others, such as errorful learning or retrieval practice (described in detailed in the introductory chapter of this thesis, see section titled The Role of Effort and Success in Word Learning), which encourage participants to retrieve words using only the required level of support via the provision of cues which are delivered in a hierarchical fashion, moving progressively from minimal (e.g., first letter) to maximal (e.g., providing the written form of the word). This way, participants make an effort to name the target word correctly until they find the right level of support required to succeed, as opposed to relying on reading the solution from the very beginning. As discussed in the introductory chapter (see section titled The Role of Effort and Success in Word Learning), while repetition training has been shown to be as effective as effortful approaches at strengthening the connections between lexical and phonological levels of representations, it is less effective at strengthening the connections between semantic and lexical levels of representation (Schuchard & Middleton, 2018a, 2018b).

In contrast with feedback provided constantly, in the form of a full written or spoken word form, binary feedback on naming accuracy provided by an ASR would still allow participants to overcome their self-monitoring deficits without preventing them from engaging with effortful retrieval naming. It is therefore important to explore the potential of ASR-based feedback to positively affect the extent to which for people with aphasia can benefit from self-administer naming therapy, in the absence of a therapist (Ballard et al., 2019; Barbera et al., 2021; Palmer et al., 2019, 2020; Swales et al., 2016). As discussed in the introductory chapter of this thesis, empirical evidence stresses the importance of providing feedback on naming performance for therapy to be successful, suggesting that feedback is an indispensable feature for speech-language therapy to be effective. In a recent study, Sze et al. (2020) presented a meta-analysis of single-case studies investigating the ingredients that were most strongly associated with gains in naming skills. Pooling together 32 studies including more than 100 subjects, the analysis revealed that providing feedback on response accuracy[4] was one of the strongest factors associated with gains during the treatment phase, and the one most strongly associated with the extent to which such gains were maintained at a later follow-up session held more than 3 weeks after the last training session.

Currently the prevailing approach among studies employing CBAT with (Ballard et al., 2019; Palmer et al., 2019, 2020) and without (see reviews conducted by Lavoie et al., 2017; and Zheng et al., 2016; de Grosbois et al., 2023) ASR is to deliver feedback immediately after participants' responses. However, as the ability to self-monitor varies substantially among people with aphasia (Marshall et al., 1994; Marshall & Tompkins, 1982; Middleton et al., 2022; Nozari et al., 2011; Schwartz et al., 2016), we expect feedback to yield higher benefits, and to be perceived as more useful, by people with lower monitoring abilities, or with lower confidence in their ability to detect errors; conversely, it might not always be necessary when self-monitoring abilities suffice to detect naming errors.

As a matter of fact, early studies employing CBATs to treat anomia using confrontational naming tasks did not provide feedback on naming accuracy, forcing participants to rely on their ability to detect mistakes to determine whether additional practice was required (Doesborgh et al., 2004; Fink et al., 2002; Herbert et al., 2012; Laganaro et al., 2003; Mortley et al., 2004; Palmer et al., 2013; Ramsberger & Marie, 2007). To support correct naming, most computerized

[4] While the systematic review conducted as part of this study included studies wherein therapy targeting spoken naming was delivered via computers, no such studies delivered feedback on naming accuracy via ASR.

treatments also afforded the possibility to replay the audio file recorded during the last naming attempt (Fink et al., 2002; Herbert et al., 2012; Laganaro et al., 2003; Mortley et al., 2004; Palmer et al., 2013). This provides users an additional channel to evaluate the correctness of their response, the potential locus of the mistake, and in turn an insight into how to change their response to attain correct naming. In addition to this feature, all studies used programs that provided access to orthographic, phonological, and semantic cues to facilitate word retrieval. All of them reported promising results, showing training-induced improvements in word-finding and naming skills.

This raises two questions. First, how could people with aphasia improve their naming skills, absent guidance from external feedback? Second, would they have experienced bigger benefits had external feedback been delivered? Regarding the first question, we can only speculate that participants' improvements were limited to items whose naming accuracy they could correctly judge. Therefore, we expect participants to benefit from external feedback in a manner that is inversely proportional to their ability to self-monitor: the lower their monitoring skills, the higher the benefit from training under the guidance of external feedback. However, as participants' ability to self-monitor was not measured, this still remains an untested hypothesis. The studies investigating the use of CBAT integrating an ASR (Ballard et al., 2019; Palmer et al., 2019, 2020) did not make a direct comparison with treatment lacking this ingredient, and therefore no conclusions could be made as to what extent the observed improvements in naming where due uniquely to the availability of external feedback.

The Current Study

This study set out to expand the scope of the one described in the previous chapter on two fronts. First, usability was assessed by participants based on much more extensive experience using the program. In the previous study, the objective was to identify the most obvious aspects of the user interface limiting the usability of the prototype. To this end, a short usability session was sufficient to get participants' very first impressions on what needed to be improved. Now that the first round of user feedback had been integrated, we aimed to test the usability of the improved prototype in the intended use scenario.

Second, this study was designed to conduct a preliminary assessment of the benefits that people with aphasia could gain from using the product as intended. Additionally, we wanted to assess whether delivering ASR-based feedback immediately after users' responses is always necessary. We did so by contrasting participants' gains in naming ability in a scenario where they were allowed only to rely on their self-monitoring ability to determine when to practice a given item more or to request support, with a scenario where they could also rely on feedback provided by an ASR.

Objectives and Research Questions

In line with the objectives of the study presented in the previous chapter, here the objectives regarding the usability test are:

1. Diagnose the major obstacles that might affect the usability of the current version of the prototype
2. Obtain qualitative information that may improve the features and design of the application to make it more suitable for people with aphasia
3. Assess users' opinions about the accuracy and usefulness of the feedback provided by the ASR
4. Provide a quantitative measure of the accuracy of ASR responses relative to that of human raters

Additionally, this study addresses the following research questions regarding the preliminary assessment of the training-induced potential benefits:

1. Do participants benefit from self-administered word-finding therapy when having to rely only on their ability to self-monitor to detect errors?
2. Does the provision of feedback benefit participants? If yes, in which cases?
    a. We hypothesize that there is a positive correlation between self-monitoring skills and the extent to which a person improves their naming skills in the set of words trained without feedback
    b. We hypothesize that there is an inverse correlation between people's self-monitoring skills and the extent to which they benefit from receiving feedback relative to no feedback (i.e. the lower the self-monitoring abilities, the higher the gains in the feedback condition relative to the no-feedback condition)
3. Is there a relationship between a person's perceptions on the usefulness of the external feedback and their ability to self-monitor their speech?
    a. We hypothesize that there is an inverse correlation between self-monitoring skills and people's opinions about the usefulness of receiving feedback from the ASR. (i.e. people with lower the self-monitoring abilities will find external feedback more useful than those with better monitoring skills)

METHODS

The procedures of this study were approved by the Research Ethics Boards of Dalhousie University and of the Aphasia Institute of Ontario, a community-based center based in Toronto. All participants provided written informed consent to take part in this study.

This study was conducted entirely remotely, recruiting participants from Canada between June 2022 and March 2023. Recruitment was facilitated by a) Aphasia Nova Scotia, a charitable organization that advertised this study to their members through their mailing list and social media channels; b) the Aphasia Institute of Ontario, whose qualified staff provided support by introducing eligible members to the scope, length and methods of this study, and by connecting us directly with those expressing interest in taking part in it; and c) Speech-language therapists (SLT) across Nova Scotia, who advertised the study to clients that they deemed eligible to take part in the study.

Participants, Inclusion & Exclusion Criteria

To participate in all sessions of this study, in addition to the inclusion criteria described in the previous chapter, participants had to:

- Achieve a minimum score of 70% on the auditory comprehension subtests of the WAB-R.
- Score less than 10 points on the Naming and Word-Finding subtest of the WAB-R. This criterion was added to identify participants with excellent naming abilities, and to prevent them from moving past session 1 of the study.
- Have a partner who could assist with the delivery of the standardized test during session 1, and that could help facilitate communication between the researcher and the person with aphasia during the interviews conducted on sessions 7 and 16.

Additionally, to participate in sessions 8-17 of the study (which involved the naming intervention), participants had to:

- Demonstrate difficulties naming a minimum of 80 words (i.e. not naming, or providing incorrect name) on the two sessions where the words were presented (between session 2 and 6 of the study) (see the subsection titled ‘Procedures’ for details).

We also employed the same exclusion criteria as in the previous study, along with one additional one: to isolate the effects of training naming and word-finding via the computerized therapy developed as part of this thesis, people who engaged in any type of speech-language therapy (SLT) targeting this ability were excluded. Nonetheless, engaging in SLT targeting other language domains was not considered an exclusion criterion in this study, because even if an improvement in other language processes could benefit word-finding as well, such benefits would have manifested equally in the experimental and control conditions that participants were to experience in this study.

Broadly, this study can be divided in three phases (see section 'Procedures', and figure 4.1 for details): after session 1, which was dedicated to gathering participants' consent and administering standardized language testing, sessions 2-6 were dedicated to gather naming and self-monitoring baseline performance; sessions 8-15 were dedicated to delivering training; and sessions 16-17 were dedicated to assess naming performance 1-week and 1-month after the end of training.

Fourteen participants (8 females) took part in the study (see Table 1 for demographics). Two participants completed only the first six sessions (wherein language assessment), but did not take part in any of the later sessions testing the usability or efficacy of the prototype. Twelve took part in the first seven sessions, but only two of them (S2 and S5) were eligible to complete the entire study. Two participants (referred to as S4 and S6 in this study), had taken part in the previous study as well (referred to in that study as S3 and S7 respectively). In the current study, they repeated the procedures that they underwent in the previous one, such as executing all the tasks associated with the WAB-R.

Table 4. 1 - Demographics

Subject	Age	Sex	Time Since Stroke (years)	WAB-R scores					
				SS	AVC	R	NWF	AQ	Profile
S1	66	M	9	19	9.0	9.2	9.1	93	Anomic
S2	44	M	11	17	7.5	3.4	6.1	68	Conduction
S3	59	F	3	20	9.6	9.0	9.3	96	Anomic
S4	72	M	14	13	8.8	6.2	7.9	72	Conduction
S5	47	M	5	11	9.0	6.2	5.7	64	Conduction
S6	70	F	3	18	9.5	7.4	9.2	88	Anomic
S7	67	F	2	14	8.5	8.4	4.9	72	Anomic
S8	83	F	3	18	8.2	5.6	8.2	80	Conduction
S9	62	F	6	17	8.7	5.6	8.8	80	Conduction
S10	40	F	1	18	10	8.8	9.5	93	Anomic
S11	70	M	8 months	18	9.4	9.0	8.0	89	Anomic
S12	74	M	22	13	9.8	8.8	8.6	80	Anomic

Note. SS: Spontaneous Speech; AVC: Auditory Verbal Comprehension; R: Repetition; NWF: Naming & Word-Finding; AQ: Aphasia Quotient

A Note on Sample Size

This study set out to address whether there is a relationship between participants' self-monitoring abilities and the extent to which they benefit from external feedback. Given that this is one of the first studies addressing this question, we aimed to conduct an exploratory analysis to establish whether these two variables are correlated. To this end, to ensure that this analysis was adequately powered to detect at least a large effect size (Pearson's correlation coefficient set at 0.5) while tolerating a 5% chance of Type I errors (threshold to consider the correlation as statistically significant set at .05) and a 20% chance of Type II errors (statistical power set at 80%), we conducted a power analysis using the wp.correlation function from the R library called WebPower. The output indicated a target sample size to 28 participants. This sample size was, unfortunately, not achievable. While initially 24 individuals showed an interest, four did not make themselves available to start the study; six could not start due to lack of adequate technology, or due to the unavailability of a partner that could assist them during the first session; two started, but could not continue past session 6 for personal reasons. At some point, due to time constraints, we had to discontinue data collection. This will be further considered in the Discussion.

Measures

I employed the same measures, and related materials, that I employed for the study described in the previous chapter: WAB-R (Part I); background and usability questionnaires; and the System Usability Scale (SUS). For a description of these materials I refer to Chapter 3 of this dissertation.

Study Design

A single case series, employing an Adapted Alternating Treatment Design (AATD) (Byiers et al., 2012) was adopted for this study. Within the AATD, each experimental condition is assigned a specific set of items, enabling the comparison of training effects.

The study design is depicted in Figure 4.1. This design entails a single baseline phase (sessions 2-6), an intervention phase in which the experimental conditions are alternated within participants (sessions 8-15), and a follow-up phase to assess the extent to which training-induced gains were retained 1-week and 1-month after the end of the training phase (sessions 16 and 17).

Two study sessions were dedicated to assessing the usability of the program: one prior to the beginning of the intervention phase (session 7), and one after its end (session 16). Participants' naming performance was probed three times throughout the intervention phase, at the beginning of sessions 10, 12 and 14.

Experimental and Control Conditions

To isolate the potential role of external feedback on naming performance, the two experimental conditions were designed to be identical except for the fact that, in one of them, the ASR provided feedback on the accuracy of participants' vocal response (feedback condition), while in the other one the feedback was not provided (no-feedback condition).

During the training phase of the study (see the subsection titled 'Procedures' for details), participants experienced the feedback and no-feedback conditions on different sessions (and days), in an alternated fashion. Half the participants experienced the feedback condition first. Participants were assigned a starting condition based on when they enrolled in the study relative to previous participants (i.e. the first participants started with the feedback condition, the second with the no-feedback condition etc.).

The two experimental conditions alternated over four cycles (ABABABAB), and each condition was assigned a unique set of stimuli (in line with the AATD; Byiers et al., 2012; Sindelar et al., 1985).

The study design also included two control conditions, both of which involved distinct sets of words that were not subject to training. One set was probed only during the baseline and follow-up phases, while the other set was also regularly probed (along with the sets associated with each training condition) throughout the training phase (before the beginning of the second, third, and fourth training cycles). These control conditions were set up to make it possible to distinguish between training-induced generalization effects to untreated items versus mere probing effects (Howard et al., 1985, 2015). For example, if participants' performance on both sets of untrained words did not change during the training phase and/or the follow-up sessions relative to the baseline sessions, this would suggest that neither probing effects nor generalization have occurred. Instead, if performance improved in both sets to a similar extent, these changes could be attributed to training-induced generalization, as probing alone cannot explain this pattern of results. Finally, if only the performance on the set of untrained words that was probed during training improved relative to baseline, this would suggest that probing effects, rather than generalization, were at play.

Baseline and Follow-up Sessions

During the baseline phase, participants' ability to name a set of words was tested twice over the course of five sessions (sessions 2-6). The testing served to measure their baseline performance, and to determine which stimuli could be selected as part of the experimental and control conditions.

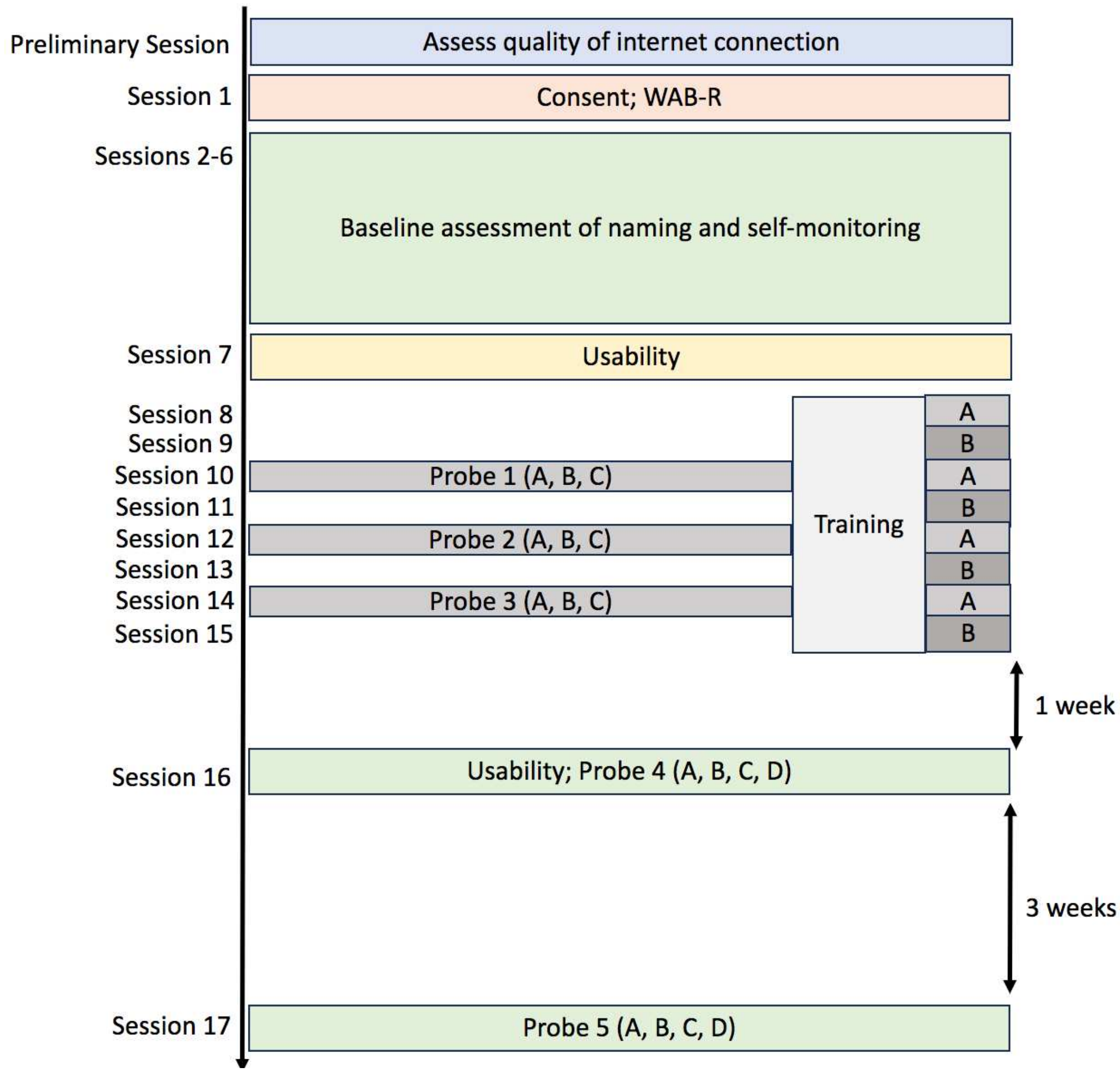

Figure 4. 1- Outline of the study design. At the beginning of sessions 10, 12 and 14 participants' naming performance on the set of words belonging to the two experimental conditions (A and B) and to one control condition (C) was probed. During the probing task, participants had to perform a picture naming task without the support of any cue, and without receiving any feedback from the ASR. Once the probing task was over, participants started the training session. The same probes were repeated 1-week and 1-month after the end of the training phase (sessions 16 and 17), but in these cases also words belonging to the second control condition (D) were probed.

In addition, the stimuli belonging to the experimental conditions were tested a third time: the very first time that participants engaged with them during the training phase, on the first trial of the first cycle of training. In fact, during this trial participants' performance was not affected by previous training, nor by the various forms of support offered by the program: the feedback from the ASR was presented after naming occurred, therefore performance (only on the first trial) was not affected by it; also, as verified after data collection was completed, on the first trial participants did not resort to the support of the orthographic or semantic cues, hence their performance could not be affected by this factor either. The choice to have only two 'pure' baseline scores before the beginning of the intervention phase, as opposed to three (which is recommended to ensure that a clear and stable pattern, if present, can be detected; Kratochwill et al., 2013), was justified by the need to minimize the number of study sessions. This need emerged during pilot testing, where prospective participants and professionals who helped with recruitment declared that 17 sessions spread over a period of about 2 months were already a significant commitment for most participants, and that adding more might have discouraged some from enrolling in the study, or caused high rates of attrition. The alternative option - having participants naming more words per session in order to fill-in more baseline tests in the same number of sessions – was discarded as well, as it would have increased the length of each session to over 1 hour, a threshold that most people with aphasia might not feel comfortable with. These reasons motivated the decision to exploit the first attempts at naming items during the first training session, although this did not apply to the control conditions, for which a third baseline score could not be acquired.

Participants' naming performance in the sets of words associated with the experimental and control conditions were also probed over two follow-up sessions, 1 week and 1 month after the last training session, respectively.

Criteria to Evaluate the Quality of the Study Design

The recent increase in popularity of single-case design studies has called forth an effort to establish guidelines for ensuring that studies meet specific quality criteria. These guidelines support researchers with designing experiments that maximize internal and external validity, and the scientific community in the pursuit of synthesizing the body of knowledge generated by single-case designs to inform evidence-based treatments. To this end, Kratochwill et al. (2013; Kratochwill & Levin, 2010) proposed the following three criteria to gauge the quality of a single-case study adopting an AATD (the one adopted in this study). These criteria were set by a panel of experts working for the What Works Clearinghouse (WWC) initiative, a project established by the Institute of Education Sciences with the mandate to evaluate the internal validity of studies investigating the efficacy of educational and psychological interventions (Kratochwill et al., 2013; Kratochwill & Levin, 2010; *What Works Clearinghouse*, 2023). The three criteria proposed by the WWC are: *first*, that the researcher must be in full control of when and how the independent variable (in this study, the type of intervention) is manipulated; *second*, that the outcome variable (in this study, naming accuracy) is systematically evaluated and documented by more than one observer, using accepted psychometric measures of agreement and at least 20% of the data points within each condition; *third*, that studies employing an AATD should present the two experimental conditions adjacently at least five times for the study to meet the standard, or four times to meet the standard with reservations — in both cases with at most 2 points per phase.

This study meets the standards according to the first and second criteria, and meets the standards with reservation regarding the third criteria. Regarding the second criterion, two observers scored all of the data associated with the outcome variable. An accepted psychometric measure (the inter-rater reliability, or IRR) was employed to assess the extent to which the observers agreed, setting a threshold at .80 as recommended (Kratochwill et al., 2013). Regarding the third criterion, this study presented four repetitions of the AB sequence during the training phase.

Stimulus Selection

The picture stimuli used in this study were taken from the International Picture Naming Project (IPNP) dataset (*The International Picture-Naming Project*, n.d.). From this dataset 300 pictures were selected, at random, from the pool of associated words that had no more than four syllables.

Procedures

Preliminary Phase

The preliminary phase of the study had exactly the same objectives as in the one described in the previous chapter, namely: ensuring that the study could be conducted remotely under the best possible conditions (i.e. stable internet connection; participants being aware of how to optimize the quality of audio and video; participant's partners being aware of how to best assist the person conducting the study during the administration of the WAB-R, etc.) For this reason, this session was conducted following the same procedures as in the preliminary session of the previous study.

Session 1 – Consent Form, Standardized Testing and Background Questionnaire

At the beginning of this session, participants provided informed consent to take part in the study. Then, the examiner administered the digital version of the WAB-R. If participants met all inclusion criteria, the examiner also administered the questionnaire in Background Questionnaire (see Appendix B).

Sessions 2-6: Stimuli Selection and Assessment of Pre-Treatment Baseline Naming Performance (~ 1 hour each)

Sessions 2-6 were designed to select the corpus of words to be used as training material (and control conditions). To this end, participants engaged with a naming task designed to identify which words they were unable to name. As mentioned before, 300 stimuli were selected for this purpose, and each had to be named twice across sessions 2-6, for a total of 600 trials (120/session). This was done using the same program that was later used for the training phase (sessions 8-15). Importantly, sessions 2 to 6 were identical, except for the stimuli that participants had to name. Since the program is able to keep track of the session each participant is in at any given moment, it could adapt the list of stimuli that they engaged with accordingly.

During session 2, which was done via video call, participants were first asked to confirm their consent to take part in the study, then they were instructed on how to access the appropriate website. Upon starting the program participants had to log into their account by inserting their personal username and password (provided to them earlier by the examiner). Once logged in, they watched a video tutorial with detailed instructions on how to perform the task. Briefly, when engaging with the confrontational naming task, at the beginning of each trial a picture appeared at the center of the screen. As discussed in chapter 3, from this moment on participants were given three seconds to prepare a response. During this time, the microphone was turned off.

Then, a written prompt (‘SPEAK’) appeared on the screen, signaling participants that the microphone was on and recording their vocal responses. At this stage of the experiment no feedback was provided on naming accuracy. The microphone recorded vocal responses for 10 seconds, after which the picture disappeared. To gauge participants’ ability to monitor their performance, once the recording was over they were prompted to answer the following question: “Do you think your response was accurate?” by selecting one of the following options: “Yes”, “No”, and “I don’t know”. Once responded, they were in control of determining when to move to the next trial by pressing the associated button.

During sessions 3-6 participants were not required to meet with anyone from the research team.

The training material and matched set of untrained stimuli were selected among the group of words that participants named incorrectly during both sessions they were presented in (see section ‘Analysis of Vocal Responses’ for a description of how the vocal responses recorded throughout sessions 2-6 were scored to determine their accuracy). From this pool, 80 words were selected and divided into four sets (one per condition: two for trained, two for untrained material) of 20 items each. The decision to select 20 words per condition was based on a trade-off between the size of the confidence interval that applies to the result, and the length of the training sessions. In the context of this study, the confidence interval can be interpreted as the range of variability in a person’s naming accuracy that should be expected when this person repeats the same task, with the same stimuli, under the same conditions (i.e. no feedback or support received; no training in-between repetitions of the task; same stimulus order, etc.). It is, in other words, an index of measurement error (Howard et al., 2015).

The confidence interval can be obtained from the binomial theorem (Howard et al., 2015), which, taken a parameter *p* indicating the expected value for the outcome variable (for this study, the percentage change in performance from baseline to follow-up), defines the lower and upper confidence intervals as:

$$Low_{CI} = p - z_{\frac{a}{2}} * \sqrt{(p * \frac{(1-p)}{n})}$$

$$High_{CI} = p + z_{\frac{a}{2}} * \sqrt{(p * \frac{(1-p)}{n})}$$

In both equations, the term $z_{\frac{a}{2}}$ indicates the z-score corresponding to the desired level of confidence (in this case, 95%), and *n* is the size of the item set (Glen, n.d.). Using these formulas, it is possible to plot how the size of the confidence interval varies with the sample size. Figure 2 renders this result, with the sample size of a hypothetical training session varying along the horizontal axis (spanning the range from 5 to 101), while the size of the confidence interval (only half of it) is depicted on the vertical axis (spanning the range from 43.8% to 9.8%). Intuitively, as the size of the training set increases, the confidence interval shrinks. Critically, the rate at which the confidence interval decreases is exponential: about 66% of the maximum decrease in the size of the confidence interval occurs when moving from a sample size of 5 to 20 (from 43.8% to 21.8%). Given the diminishing returns gained in confidence as the size of the training set increases, I considered the value of 20 items to be an optimal trade-off between the need to minimize the size of the confidence interval while keeping the training set small enough to prevent overly lengthy training sessions.

To provide a concrete example (along the lines of the one provided by Howard, 2015, p. 533), if we assume that the set size is 10, and we assume that training results in a change in performance of 50% of the total maximum score (from 0 to 5 words named correctly), the margin of error around the observed score is expected to be +/- 31%, meaning that assuming that repeated naming under the same conditions could yield performance between 2 and 8 correct words. Conversely, assuming the same baseline performance (0 words named correctly pre-training) and amount of change due to training (10/20 words named correctly post-training), a set size of 20 items would yield performance ranging between 8 and 12 words.

It is well-known that, when sampling words at random from a lexicon, there is considerable variation along different factors that are known to influence naming accuracy. These factors include: word frequency and word length (Kittredge et al., 2008); name agreement (Cheng et al., 2010) — an estimate of how commonly a given name is associated to a specific object; visual complexity (Cuetos et al., 2002); object familiarity (Cuetos et al., 2002); and finally, age of acquisition (Cuetos et al., 2002; Valente et al., 2014).

In line with best practices (Howard et al., 2015), to prevent introducing confounding factors in the analysis, once selected the 80 stimuli were randomly assigned to the four conditions to minimize differences in any of the above-mentioned properties among sets.

Session 7 - Usability Study

At the beginning of this session the examiner provided a short demonstration on how to use the prototype of the program. Then, participants started using the program on their computer, while sharing their screen. This allowed video-recording the entire session for the purpose of data analysis. Once participants completed the tasks, the System Usability Scale (SUS) (Brooke, 1986) and the usability questionnaire (see Appendix B) were administered.

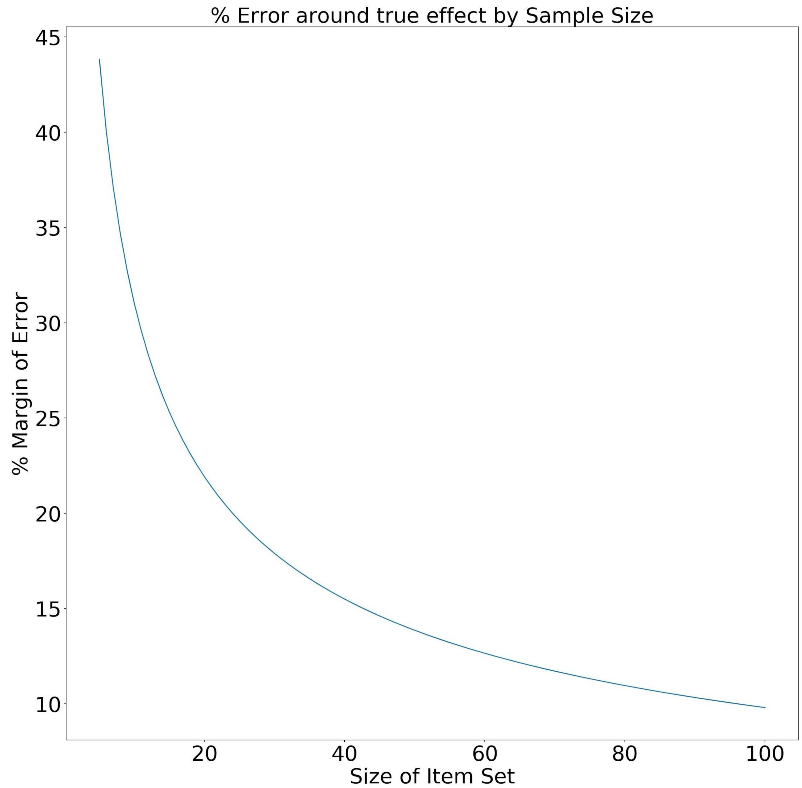

Figure 4. 2 - Relationship between the size of the training set and width of the confidence interval.

Sessions 8-15: At-home Training

To perform these sessions, participants were not required to be in touch via video call with a member of the research team. They were encouraged to do a training session every other day, whenever it worked best for them. Their usage was monitored remotely to ensure that they complied with the recommended training regime. Upon noticing that a participant had not logged in for at least 3 days, we sent an email with a gentle reminder to perform the next session as soon as possible. If we did not receive a response within 3 days, another reminder was sent. Participants were considered to have discontinued the study if they did not respond to the second reminder within 1 week.

In a typical session participants had to name 20 words, three times, to ensure they got enough exposure to the training material. During the training sessions, support with the naming task was provided by the program in the form of written semantic/orthographic cues, verbal cues (listening to the audio file of another person—a native English speaker without aphasia—naming the target word correctly, or listening to the recording of one's own last attempt), or, depending on the session number, in the form of feedback on vocal accuracy delivered by the ASR.

Before the beginning of training sessions 10, 12, and 14, participants' performance was probed to assess whether, and to what extent, any improvement observed in untrained items could be attributed to genuine generalization effects as opposed to mere probing effects (as discussed earlier in Experimental and Control Conditions). Probing entailed naming a total of 60 words — 40 associated with both sets of trained words, and 20 associated with one set of untrained words — while a microphone recorded their vocal responses. No form of support (i.e. provision of written or verbal cues) or feedback on naming accuracy was provided during the probing procedure.

Session 16-17: One Week and One Month Follow-up

This session occurred one week after session 15. The first part of this session repeated the procedures of the usability session described above (session 7): participants used the program while sharing their screen, and with the interaction being video-recorded such that it could be later analyzed to extract insights into the usability of the program; afterwards, SUS and usability questionnaires were administered. In the second part, participants engaged with the same naming task they engaged with during the baseline sessions (sessions 2-6), except for the fact that instead of naming 300 words, they named only the 80 that were selected for the trained and untrained sets that were to be tested during and after the training phase (40 trained words, 40 untrained words). Importantly, no form of support, nor any form of feedback was provided during this phase. The objective was to assess whether, and to what extent, eventual training-induced gains in naming and word-finding were maintained one week after the end of practice.

Session 17 repeated the second part of session 16 and was held one month after session 15.

Data Analysis

Usability

To address the research questions concerning the usability of the program, data analysis was conducted following exactly the same procedures as in the study described in the previous chapter.

Analysis of Data Collected during Training Sessions

Research question 1 was addressed by combining visual analysis with metrics designed to quantify the effect of different experimental conditions on the outcome variable. We planned to address research questions 2 and 3 by means of correlational analysis.

Visual Analysis. Visual analysis was the primary method of choice to analyze performance data, where performance was expressed as the proportion of items named correctly at each time point. In this regard, we followed well-established guidelines to assess the quality of data, the evidence in favor or against the presence of an experimental effect, and its magnitude (Byiers et al., 2012; Kratochwill et al., 2013; Manolov & Moeyaert, 2017). First, baseline data were assessed to determine whether a stable level could be identified, or whether there was any trend in the data. As one of the objectives of this study was to assess the impact of a specific intervention on learning outcomes, the key to establish the presence of an experimental effect was to first ensure that baseline data showed minimal variability, ideally with scores at floor level. This criterion was enforced by selecting, for training, only items that were named incorrectly twice over two of sessions 2–6 (the baseline/stimulus selection phase). Then, data points between adjacent phases (i.e. baseline vs. training; training vs. follow-up) were compared to assess changes in the level or trend from one phase to the other, and the amount of overlap. As mentioned above, visual analysis was complemented with quantitative measures, which are described below. Finally, the naming performance on the probes administered during the training and follow-up phases were analyzed to assess whether there were at least three demonstrations of an experimental effect at different points in time.

Overlap-Based Measures. When comparing data collected across different study phases and conditions by means of visual analysis, the amount of overlap can depend on whether an experimental effect is present or not: the lower the overlap, the higher the level of confidence in concluding that the study conditions had a different effect on the outcome measure. To supplement the information gathered by visual analysis, previous studies have proposed specific metrics to *quantify* the extent to which data overlap.

A popular overlap-based metric is Percentage of Non-Overlapping Data (PND) (Scruggs & Mastropieri, 1998). In the context of this study, PND would be calculated by considering the highest score in the baseline phase, and determining the percentage of scores from the training/follow-up phases that are higher than this point. Therefore, the PND - which ranges between 0, indicating full overlap, to 1, indicating no overlap - tracks the extent to which a data point, taken at random from one phase, is likely to fall outside of the range of data from the other phase. However, this method ceases to be informative in the presence of outliers. To illustrate this point, let's consider two examples. In the first case (see figure 4.3, left), one of the scores recorded in the baseline phase is higher than 60% of the scores from the training phase, but that the rest of the baseline scores are lower than the lowest score from the training phase. In the second case (see figure 4.3, right), 60% of the scores from the training phase fall within the range of the highest 50% of points from the baseline phase.

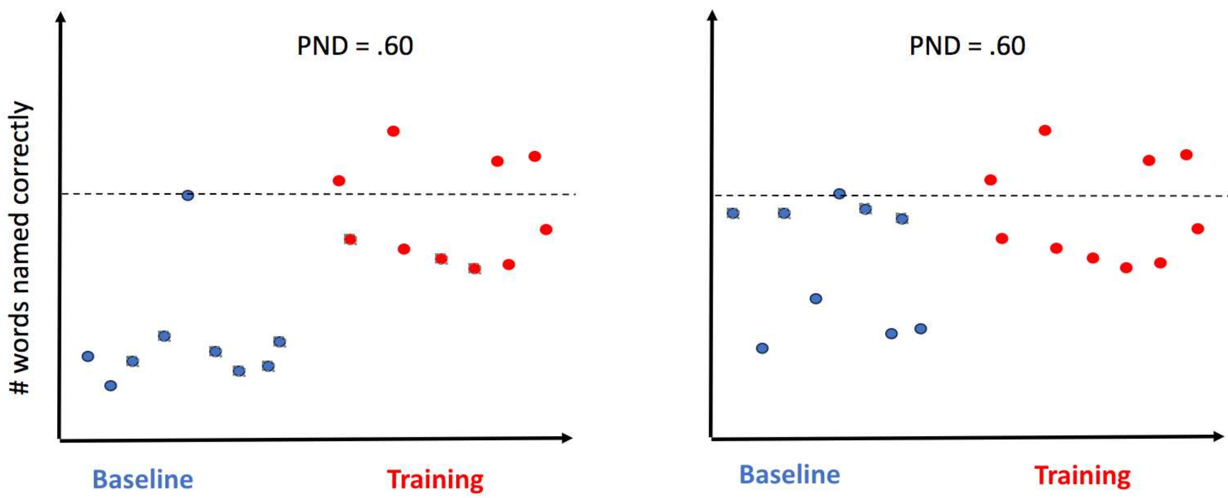

Figure 4. 3 Illustrative example of the limitations of the Percentage of Non-Overlapping Data metric.

Clearly, in the first case, visual analysis would lead one to conclude that, despite a single outlier, the training condition is associated with a sharp increase in the scores, relative to baseline, while in the second case the effect would be much more attenuated. However, in both cases the supplementary quantitative measure PND would be equal to .4, illustrating why in the presence of outliers this metric cease to be informative.

The Percentage of All Non-Overlapping Data (PAND) (Parker et al., 2007) was proposed to address this problem. PAND is calculated by considering all the points in the two contrasting conditions, and not just the highest/lowest in the baseline condition to set up a threshold: this is done by first determining the minimum number of scores that have to be removed from each phase in order to have no overlap between the data. PAND is then the ratio between the number of remaining points versus the total number of points. PAND can be interpreted as the chance that, when selecting at random a point from the baseline or the intervention phase, the point does not belong to the range of overlap.

Percentage of Mean Change Between Conditions. Overlap-based metrics quantify the extent to which data from one phase are likely to overlap with those from another, but do not inform on how strong the effect of one condition might be. For example, there could be two cases where the scores from baseline and experimental conditions do not overlap, yielding PAND of 100% in both instances. Yet, in one case the medium/maximum score of the experimental data might be just 5% higher than the corresponding score in the baseline phase, while in the other case it might be 50%. Clearly, PAND or other overlap-based metrics cannot provide any insight in this regard. To this end, a complementary metric to PAND is the percentage change in performance between two phases of the study relative to the maximum possible change (Dallery et al., 2013; Graham et al., 2012; Manolov & Moeyaert, 2017), which can be expressed as such:

$$\Delta Naming = \frac{Naming_{followup} - Naming_{baseline}}{20}$$

Analysis of Vocal Responses

The audio files of vocal responses recorded across all sessions (from session 2 to session 17) were analyzed using the Audacity software. Two native English speakers (not involved in testing subjects) transcribed the audio files and judged the accuracy of each item in a binary fashion (correct/incorrect). Whenever the transcription included multiple utterances (i.e. an initial incorrect response followed by a successful correction), the rater was instructed to classify it as correct if the correct answer was present among the attempts, or incorrect otherwise.

Agreement between the raters was assessed by a member of the research team, who calculated the IRR as percent agreement across the two raters. In line with the current convention (Kratochwill et al., 2013), the threshold to determine whether the IRR was acceptable or not was set to 80% of agreement.

Calculation of Self-Monitoring Scores

The basis to calculate self-monitoring scores were participants' responses to the question "*Do you think your response was accurate?"* that was posed immediately after each naming attempt. Instances wherein participants responded: "I don't know", were pooled together with responses that were incorrectly evaluated. Instances wherein no response was provided were removed from the analysis. For each study session, the self-monitoring score was then calculated as the percentage of participants' responses that matched the judgement of both human raters.

Analysis of Performance of the ASR, and of Participants' Self-Monitoring Abilities

Following the same rationale described in Appendix C, and employed in the study described in the previous chapter, we decided to employ signal detection theory (Donaldson, 1992) to assess the performance of the ASR. The judgments of two human raters were used as ground truths to judge the performance of the ASR against. We used the same method to analyze participants' self-monitoring abilities, and to determine the extent to which they could discriminate correct vs incorrect responses, and whether or not there was any bias in their judgment.

RESULTS

Participants' Language Profile

Participants' language profiles, as measured via the WAB-R, are documented in Table 1. Five individuals had AQ scores between 88 and 96, which correspond to a minimal level of severity, while seven participants met the inclusion criterion of a mild to moderate level of severity, with AQs between 64 and 80. Only three (S2, S5 and S7) were eligible to complete the study, based on their performance in sessions 2-6 (see Table 4.15 for details). Unfortunately, while S7 completed sessions 8 and 9, she had to discontinue the study due to personal reasons that were not related to a loss of interest in completing the study, or with usability issues.

Usability

Effectiveness

Tables 4.2 and 4.3 show scores reflecting how effectively participants executed all the tasks. The analysis of effectiveness scores revealed that 8 out of 12 participants could complete all tasks independently: two without any support, and six by relying only on minimal support provided by the video tutorial or from a hint from the research. The other four participants could execute all but one of the tasks independently. Two participants could not complete the first two tasks independently (figuring out how to move from the welcome scene to the login scene, and inserting their credentials to login into their personal account), and needed their partner to carry out these steps for them. Inserting the login credentials was difficult because these two participants (S8 and S12) had very limited mobility in their dominant hand and could not easily type in the letters in the username and password input fields. Both of them suffered from hemiparesis affecting the side of their dominant limbs, and this limited their ability to type.

For one participant (S8), this limitation was compounded by the format that was used to set-up the credentials: *name_number*, where ‘name’ was the participant’s name, and ‘number’ reflected the order in which the participant enrolled in the study. Importantly, to save users from having to set the first letter uppercase, all characters had to be typed in lowercase.

Table 4. 2 - Effectiveness

Task	Participant									
	S3	**S1**	**S10**	**S11**	**S6**	**S12**	**S9**	**S8**	**S7**	**S4**
Access web and type web address			3	3			3			
Welcome scene	2	3	3	3	1 (1VT)	1	3	1		3
Login into personal account	3	3	3	3	3	1	3	1	3	3
Naming task	3	3	3	3	3	3	3	3		1
Star trial	3	3	3	3	3	3	3	3	3	1
Use cues to help with naming	3					3	3		3	2
Use 'Try Again' function	2 (1 VT)			2	2 (1 VT)		2		3	2
Use 'Replay Audio' function	3						3			
Use 'Listen to Target Word' function	3				3	3	3		2	
Move to next item	3	3	3	3	3	3	3	3	3	3
Exit task when completed	3	3	3	2	3	3	2			
Review summary of performance	3	3	3	3	3	3	3			
Quit program	3	3	3	3	3	3	3			

Table 4. 3 - Effectiveness (for participants who completed two usability sessions)

Task	**Participant**			
	S2		**S5**	
	pt1	**pt2**	**pt1**	**pt2**
Access web and type web address	3			
Welcome scene	3	3	3	3
Login into personal account	3	3	3	3
Naming task	3	3	3	3
Star trial	3	3	3	3
Use cues to help with naming	3	3	3	3
Use 'Try Again' function	3	3	2	3
Use 'Replay Audio' function	1 (2 VT)			
Use 'Listen to Target Word' function	3	3		
Move to next item	3	3	3	3
Exit task when completed	3	3	3	3
Review summary of performance	3	3	3	3
Quit program	3	3	3	3

Note. Numbers within parentheses indicate the number of times that a participant used the video tutorial to tackle a specific task. The presence of an asterisk indicates that participants requested and received a suggestion from the examiner on how to move on with that task

We made this choice to make the credentials as easy to remember and to type as possible. While most participants did not find it difficult to perform this task, S8, which used an iPad, was frustrated when the device attempted to force the first letter of her name to be in uppercase, which invalidated the credential. As she also found it difficult to memorize the strategy required to force the iPad to accept a proper name in lowercase letters, she had to resort to the help of her husband to log in.

Most participants (6 out of 12) struggled when they had to repeat a given trial. Of the eight users who engaged with this task, six had to rely on the support of a video-tutorial or of their partner to do it correctly, and these same six users also made mistakes while doing so.

Efficiency

In line with the effectiveness scores, the participants who struggled remembering how to repeat a given trial also made mistakes when attempting to do so. Otherwise, all other tasks, except for the one entailing listening to the last attempt at naming a word, were executed by all participants without mistakes (see Table 4.4 and 4.5).

Satisfaction

Table 4.6 and 4.7 show participants' global SUS scores, while Table 4.8 and 4.9 show the SUS scores broken down by statement. These tables highlight a strong association between SUS and AQ scores. As shown in Table 4.6, a participant's satisfaction with the program – as measured by SUS scores – could be easily predicted by their AQ: an AQ >88, which indicates very minimal severity of aphasia (except for S7), was associated with high SUS scores (above 85), while AQs ranging from 64 to 80, which indicate the presence of mild to moderate aphasia, were associated with SUS scores ranging from 47.5 to 77.5.

S8 and S9 were the least satisfied, due to their neutral opinion on the user-friendliness of the program (statements 3, 4 and 8) and to the fact that they did not find it easy to learn enough about the program to be proficient at it. Both commented that they would have preferred to have more time to get familiar with the program before doing the study. S8 also said that she would probably need to rely on somebody else's assistance while using the program.

Table 4. 4 - Efficiency

Task	**Participant**									
	S3	**S1**	**S10**	**S11**	**S6**	**S12**	**S9**	**S8**	**S7**	**S4**
Access web and type web address			0	0			0			
Welcome scene	0	0	0	0	0		0			0
Login into personal account	0	0	0	0	0		0			0
Naming task	0	0	0	.3	0	1	0	0	0	
Star trial	0	0	0	0	0	0	0	0	0	
Use cues to help with naming	0					0	0		0	
Use ‘Try Again’ function	1			1	0	2	0	1	0	1
Use ‘Replay Audio’ function	0				1		0			2
Use ‘Listen to Target Word’ function	0				0	0	0		0	
Move to next item	0	0	0	0	0	0	0	0	0	0
Exit task when completed	0	0	0	0	0	0	0			
Review summary of performance	0	0	0	0	0	0	0			
Quit program	0	0	0	0	0	0	0			

Table 4. 5 - Efficiency (for participants who completed two usability sessions)

Task	Participant			
	S2		S5	
	pt1	pt2	pt1	pt2
Access web and type web address	0			
Welcome scene	0	0	0	0
Login into personal account	0	0	0	0
Naming task	0	0	1	0
Star trial	0	0	0	0
Use cues to help with naming	0	0	0	0
Use 'Try Again' function	0	0	3	0
Use 'Replay Audio' function	0			
Use 'Listen to Target Word' function	0	0		
Move to next item	0	0	0	0
Exit task when completed	0	0	0	0
Review summary of performance	0	0	0	0
Quit program	0	0	0	0

Table 4. 6 - System Usability Scale scores

	Participant									
	S3	S1	S10	S11	S6	S12	S9	S8	S7	S4
Score	97.5	87.5	85	85	100	77.5	47.5	50	82.5	67.5

Note. The green background is associated with scores ranging between 80-100, reflecting good satisfaction with the product; yellow is associated with scores ranging between 60-80; orange is associated with scores below 60.

Table 4. 7 - System Usability Scale scores (for participants who completed two usability session

	Participant			
	S2		**S5**	
	pt1	**pt2**	**pt1**	**pt2**
Score	67.5	82.5	85	85

Note. pt1 indicates the score recorded during the session 7 (first usability session), while pt2 indicates the score recorded during session 16 (second usability session).

Participant S2, who took part in sessions 7 and 16, was more satisfied with the program (as demonstrated by the drastic improvement in his SUS score from pre- (67.5) to post-training (82.5)) due to the fact that he got more familiar and proficient with it (statements 7 and 10), and that he found it more user-friendly (statement 8).

Clarity and Usefulness of the Video-Tutorials

Only five participants could provide their opinions on the clarity and usefulness of the video tutorials, as the others did not make use of them (see Table 4.10 and 4.11). Among these participants, all found the tutorials to be useful to remember the function of specific buttons (except for S4, who expressed a neutral opinion). Nonetheless, S2 and S4 expressed a neutral opinion about the clarity of the instructions, with S2 saying that in some cases it was not easy for him to understand their content.

Participants' Perception of the Accuracy and Usefulness of the ASR

All participants, except for S1 and S10, found the feedback from the ASR to be useful (see Table 4.12 and 4.13). Four participants (S1, S3, S8 and S7) did not find the feedback to be always accurate.

Table 4. 8 - System Usability Scores - breakdown of results by statement

Statement	Participant									
	3	**1**	**10**	**11**	**6**	**12**	**9**	**8**	**7**	**4**
I would use the program frequently	5	3	5	5	5	5	4	4	3	4
The program is easy to use	5	5	5	4	5	5	3	3	5	5
The various functions in the program were well integrated	4	3	4	5	5	3	5	4	3	4
There is consistency in the program	5	4	4	5	5	4	3	4	3	4
Most people would learn to use the program very quickly	5	5	4	5	5	3	4	3	5	4
I felt very confident using the program	5	5	5	5	5	4	1	3	4	4
The program is too complex to use	1	1	2	1	1	2	3	3	1	2
I would need the support of a technical person to be able to use the program	1	1	1	4	1	1	5	5	1	3
The program is cumbersome to use	1	1	2	2	1	2	3	3	1	4
I needed to learn a lot of things before I could get going with the program	1	1	2	2	1	2	4	4	1	3

Note. Statements 7-10 have a negative connotation with respect to the usability of the program. Therefore, in these cases low scores (1 and 2), which reflect disagreement with the statement, are highlighted in green. Vice versa for high scores such as 4 and 5.

Table 4. 9 - System Usability Scores - breakdown of results by statement (for participants who completed two usability

Statement	Participant			
	S2		S5	
	pt1	pt2	pt1	pt2
I would use the program frequently	5	4	4	4
The program is easy to use	4	4	4	5
The various functions in the program were well integrated	5	5	5	4
There is consistency in the program	4	4	4	4
Most people would learn to use the program very quickly	3	4	4	3
I felt very confident using the program	5	5	4	4
The program is too complex to use	2		3	1
I would need the support of a technical person to be able to use the program	3	2	1	1
The program is cumbersome to use	3	2	3	2
I needed to learn a lot of things before I could get going with the program	5	1	3	3

Table 4. 10 - Opinions on the usefulness and clarity of the video instructions

	Participant									
	S3	**S1**	**S10**	**S11**	**S6**	**S12**	**S9**	**S8**	**S7**	**S4**
Accurate										
Useful										

Note. The green color reflects participants' affirmative response to the perceived clarity and usefulness of the video-instructions. Yellow reflects participants opinion that the video-instructions was only sometimes clear/useful

Table 4. 11 - Opinions on the usefulness and clarity of the video instructions (for participants who completed two usability sessions)

	Participant			
	S2		**S5**	
	pt1	**pt2**	**pt1**	**pt2**
Useful				
Clear				

Table 4. 12 - Participants' opinion on the accuracy and usefulness of the ASR

	Participant									
	S3	**S1**	**S10**	**S11**	**S6**	**S12**	**S9**	**S8**	**S7**	**S4**
Useful										
Clear										

Note. The green color reflects participants' affirmative response on the perceived accuracy and usefulness of the ASR. Yellow reflects participants' opinion that the ASR was only sometimes accurate/useful.

Table 4. 13 - Participants' opinion on the accuracy and usefulness of the ASR (for participants who completed two usability sessions)

	Participant			
	S2		**S5**	
	pt1	**pt2**	**pt1**	**pt2**
Accurate				
Useful				

The first three referred directly to incongruences between their self-monitoring and the feedback received, but, as shown in Table 4.14, the judgment of the ASR mismatched only once with that of the human raters. In the case of S7, however, her judgment on the accuracy of the ASR was justified by the frequent mismatch (11 times) between ASR and the human raters, suggesting that the ASR was indeed unreliable for this individual.

Comparison of the Judgement of the ASR vs. Human Raters

Table 4.14 presents data on the performance of the ASR during the usability session. Tables 4.16 and 4.17 are the equivalent for the training session completed by S2 and S5, respectively. During the usability session, the performance of the ASR matched that of human raters 80% of the times (135 out of 169 responses). Application of signal detection theory revealed a good level of discrimination (a'=.87) and low bias (β''= -.19). During the training sessions, 85% of the ASR responses matched with those of human raters (522 out of 611), with a good ability to discriminate between correct and incorrect responses (a'=.89) and a moderate bias towards judging naming responses as correct (β''= -.56).

Table 4. 14 - Performance of the ASR during the usability sessions, relative to human raters

	TP	TN	FP	FN
S1	7	1	1	0
S2	9	13	3	2
S3	7	5	1	0
S4	5	6	3	0
S5	9	11	6	0
S6	8	1	0	0
S7	1	9	2	9
S8	6	2	1	0
S9	8	4	2	0
S10	9	0	0	0
S11	6	0	0	3
S12	6	2	0	1

Note. For S2 and S5, the scores were computed using data collected during the session 7 and 16. TP/TN stand for true positives and negatives, respectively. FP/FN stand for false positives and negative.

Users' Opinions on how to Improve the Prototype

All participants, except for S1, said that they would use the program again. S1 justified his opinion with the fact that his priority was to recover his ability to use numbers (which were not trained in this prototype), while he considered his word-finding and naming deficit very mild and not as important to focus on.

The data collected during the semi-structured interviews revealed positive opinions on specific features of the program, along with several recommendations on how to improve its usability. On the latter point, five participants commented that there were too many buttons in the Naming Scene, and that this made it difficult to remember their function. In particular, S11 found it difficult to remember which button to press to listen to the sound of the target word, while S6 struggled to remember how to listen to the audio file of his last naming attempt. S2 pointed out that some video tutorials did not work adequately, as they frequently interrupted before the end, and that, when listening to voice recordings, the quality of the audio was frequently too poor. S3 recommended to ensure that all pictures could be easily identified and named, as he found that some were not realistic.

Among the positive comments, S3 and S5 expressed appreciation for the usefulness of the buttons providing hints, with S3 adding that he also found the tutorials particularly useful to successfully use the prototype of the program.

Training-Induced Gains in Naming Performance

As noted, only two participants completed the training portion of the study, and each individual's data is reported separately.

S2

Table 4.15 presents participants' naming performance on the 600 trials that they were exposed to during the baseline phase (sessions 2-6). For S2, the error rate during this phase was 53%. Of the mistakes, 81% were for words that were named incorrectly on both sessions they were presented in (i.e., 130 of the 300 unique words presented). These mistakes included 45 non-responses. S2 was able to correctly evaluate his performance in 66% of the trials. As shown in Table 4.15, of the 173 errors committed over 600 total trials, he could correctly detect 67 (39%). Throughout the eight training sessions, S2 indicated that he did not know whether his answer was correct or not only twice, once during session 8 and once during session 9.

Table 4. 15 - Participants' performance in the naming task executed during the baseline phase.

	# Errors (%)	% SM	# Errors over two sessions (%)	# Errors over one session (%)	# NR over two sessions	IRR
S1	32 (.05)	.96	10 (.04)	22 (.08)	0	
S2	320 (.53)	.66	130 (.43)	60 (.20)	45	.90
S3	70 (.12)	.89	24 (.08)	22 (.08)	1	.95
S4	73 (.12)	.86	23 (.06)	27 (.08)	0	.96
S5	317 (.53)	.82	119 (.40)	76 (.25)	48	.95
S6	42 (.07)	.89	13 (.04)	16 (.04)	0	.77
S7	267 (.44)	.73	90 (.30)	87 (.30)	38	.88
S8	106 (.18)	.85	20 (.07)	66 (.22)	0	.75
S9	43 (.07)	.93	6 (.01)	33 (.05)	0	.96
S10	23 (.04)	.92	7 (.01)	9 (.01)	0	.97
S11	68 (.11)	.89	18 (.03)	32 (.05)	6	.93
S12	33 (.06)	.95	5 (.01)	23 (.04)	2	.97

Note. NR = No response; SM = Self-Monitoring score; IRR = Inter-Rater Reliability

Figure 4.4 depicts the naming performance observed during the three phases of the study (baseline, training and follow-up). Each dot represents the number of correct responses for the items associated with a specific condition (out of 20, the maximum score), at a given point during the study. There was very little variability from the first two baseline sessions (where performance was at floor, by design) to the third one, which corresponded to the first naming attempt of the first session of the feedback and no-feedback conditions, where S2 named 0 and 3 words correctly, respectively.

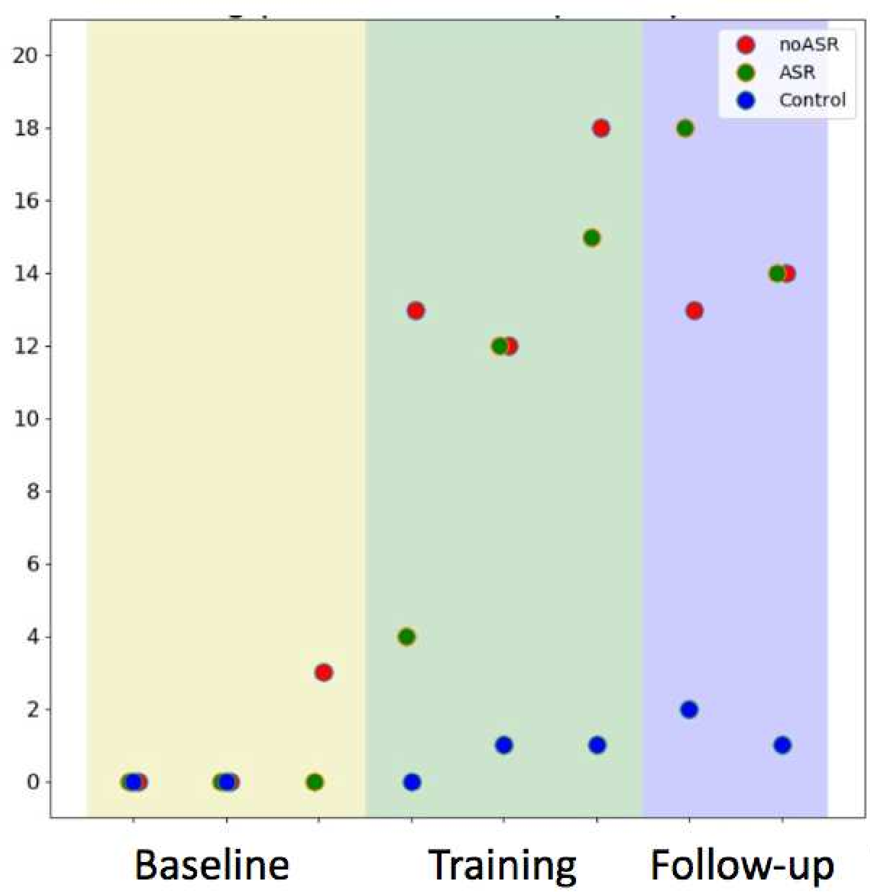

Figure 4. 4 - Naming performance of S2, throughout the three baseline sessions (yellow background), the three probing sessions administered during the training phase (green background), and the two follow-up sessions (blue background).

Visual analysis of the naming performance observed on the probes administered during the training sessions revealed that both experimental conditions are associated with a clear upward trend, peaking at 15 and 18 correct words for the feedback and no-feedback conditions during the last probing session. Naming performance in the set of words that were not trained peaked at 1 correct word. Overall, the gains observed during training appeared to be maintained 1 week and 1 month after training.

One week after the last training session, the gap between the two conditions was reversed (18 and 13 correct words named for the feedback and no-feedback conditions, respectively; 2 words for the control condition), but performance levelled 1 month after the last training session (14 for both experimental conditions, and 1 word for the control condition). The training-induced net positive effect on naming performance evident in Figure 3 was corroborated by the PAND metric, which equaled 1 (the maximum possible value), reflecting the lack of overlap between training and control data. In addition, analysis of the mean change in performance between control and training data showed that, one week after the end of training, there was an average improvement in performance of 12.5 words (62.5% of the maximum possible improvement), while at one month the improvement was of 11 words (55% of the maximum possible improvement). A direct comparison of the two experimental conditions (with and without ASR) showed no evidence in favor of either of them. While during the training phase the naming performance on the probing trials was in favor of no-feedback condition, one week after the end of training the trend was reversed, and there was no difference at one month follow-up.

Visual inspection of the trend in the use of cues by S2 throughout training sessions (illustrated in Figure D.1, Appendix D) revealed that a) S2 favored the use of orthographic cues, written and spoken word form, while semantic cues and the option to listen to their last naming attempt were seldom employed; and b), that the use of cues decreasing significantly throughout the training sessions.

Analysis of Self-Monitoring throughout the Training Sessions. Table A.16 presents data comparing the ability of the ASR and of S2 himself to judge naming accuracy. The table provides confusion matrices contrasting true responses versus the output of the ASR and of S2's self-monitoring, along with the output of the analysis conducted according to signal detection theory. The ability of S2 to discriminate between correct and incorrect responses was similar during both condition, and was similar to that of the ASR. Both S2 and the ASR showed a bias towards judging naming responses as correct. However, the bias was much more pronounced during the condition in which S2 received feedback from the ASR relative to when feedback was not provided. Also, in the latter case, bias was similar to that shown by the ASR.

Table 4. 16 Comparison between the performance of the ASR and the judgement of S2, relative to human raters

	TP	**TN**	**FP**	**FN**	**a'**	**β''**	**No Responses**
Feedback	178	29	44	3	.84	-.87	25
No Feedback	126	41	20	12	.88	-.47	20
ASR	165	51	22	16	.89	-.45	

Note. The rows associated with Feedback and No Feedback contain scores associated S2 self-monitoring performance throughout all the four sessions associated with each condition. Measures are expressed in terms of true/false positives/negatives, and associated a' and β'' scores. For each of these two experimental conditions, the last column indicates the number of instances in which participants did not provide a response. The last row contains the scores associated with the performance of the ASR.

S5

Similar to S2, over the five baseline sessions the error rate of S5 was at 53%, naming 40% of words incorrectly or providing no response on both sessions they were presented in, while adequately evaluating his performance in 82% of the cases. As shown in Table 4.15, of the 139 errors committed over 600 total trials, he could correctly detect 93 (67%).

Throughout the 8 training sessions, S5 never indicated that he did not know whether his answer was correct or not. During the first naming attempt of session 8 and 9, S5 scored 7 and 6 correct words in the feedback and no-feedback conditions, respectively (see Figure 4.5).

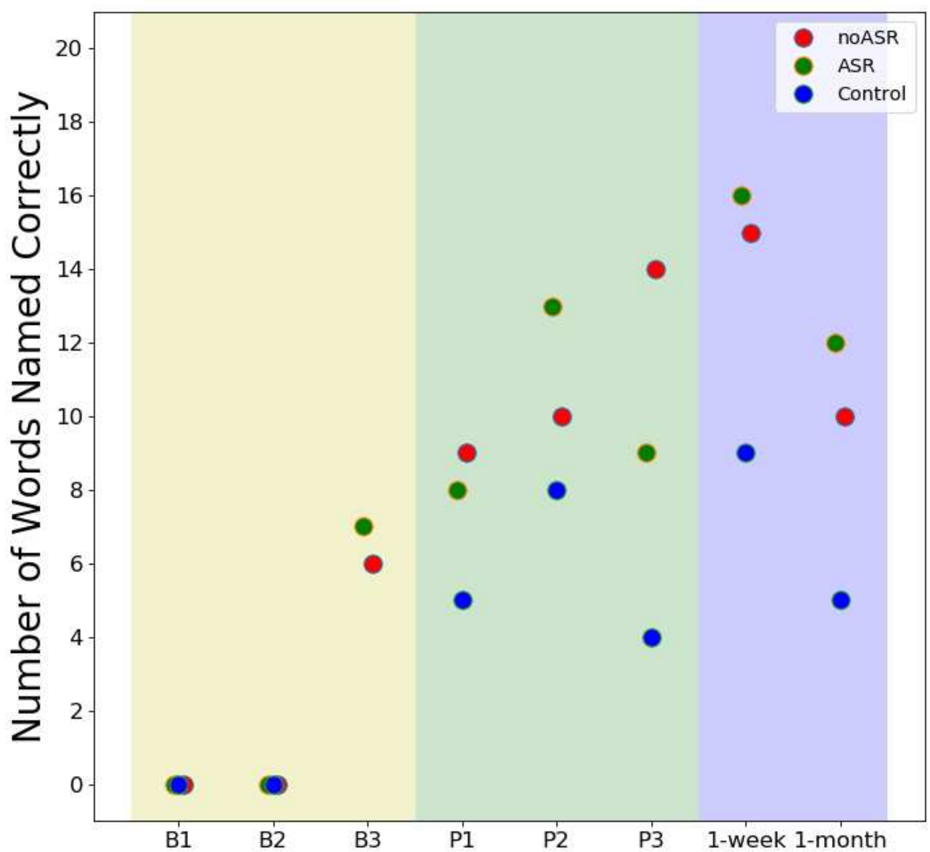

Figure 4. 5 - Naming performance of S5, throughout the three baseline sessions (yellow background), the three probing sessions administered during the training phase (green background), and the two follow-up sessions (blue background).

Naming performance for the words associated with the control condition also rose from floor level during the first two baseline sessions, to a value that oscillated between 5 and 9 throughout the next 5 sessions. However, a visual analysis of the data associated with the control condition does not suggest the presence of an upward or downward trend in the time series.

Conversely, there was an upward trend for the performance on both sets of words associated with the experimental conditions. Except for a drop of performance observed for the feedback condition on the last probe, the performance associated with both experimental conditions gradually increased until they peaked at one week after the end of training, scoring 16 and 15 correct words for the feedback and no-feedback conditions, respectively.

To calculate PAND, one point per phase had to be dropped in order to eliminate all the overlap between the phases. Therefore, as the total number of data points was 23, PAND equaled .91.

One week after the end of training, the mean performance of the two experimental conditions was 32.5% higher than the highest performance obtained during the baseline and control condition. This dropped to 10% at one month follow-up.

When comparing the performance observed in the two experimental conditions directly, there was no trend in favor of one or the other, as evident by the fact that the difference in words named correctly oscillated between 5 in favor of the no-feedback condition to 3 in favor of the feedback condition. During the follow-up session, the difference was minimal (1, and 2 in favor of the feedback condition).

Visual inspection of the trend in the use of cues by S5 throughout training sessions (illustrated in Figure D.2, Appendix D) revealed that a) S5 largely favored the use of orthographic cues, and occasionally spoken word form, while the other types of cues were seldom employed; and b), that the use of cues decreasing significantly throughout the training sessions.

Analysis of Self-Monitoring throughout the Training Sessions. Table 4.17 shows that S5 was able to correctly identify all but one of his correct responses throughout all study sessions. The ability to discriminate between correct and incorrect responses was high, and comparable, among ASR and S5's self-monitoring.

S5 demonstrate a very strong bias towards judging his responses as correct, while that of the ASR was moderate.

Table 4. 17 Comparison between the performance of the ASR and the judgement of S5, relative to human raters

	TP	TN	FP	FN	a’	β’’	No Responses
Feedback	234	8	16	1	.83	-.96	42
No Feedback	190	25	7	0	.95	-1	39
ASR	224	16	8	11	.89	-.67	

DISCUSSION

This study followed up on the one described in the previous chapter — which focused on assessing the usability of the first version of the prototype of CBAT designed to allow people with aphasia to self-administer naming therapy from home — by a) assessing the usability of an improved version of the prototype; b) gathering information on how to further improve it; c) gathering participants' opinions on the accuracy and usefulness of the ASR; and d) continuing the quantitative assessment of how accurately the ASR can detect errors in naming responses, relative to human raters. Importantly, in addition to collecting data from participants who had no experience with the program (as in the prior study), this study also involved participants who used the prototype over an extended period of time (8 training sessions, wherein participants self-administered therapy from home).

Furthermore, this study extended the scope of the previous one by addressing novel questions about the feasibility, and effectiveness, of using the prototype of CBAT to self-administer therapy from home. In this regard, the main objectives were to assess: a) whether people with aphasia could benefit from using the current version of the prototype to self-administer therapy targeting word-finding and naming skills from their home; b) if, and to what extent, the provision of external feedback via an ASR resulted in additional benefits relative to a situation wherein users have to rely on their self-monitoring abilities, and whether the latter predict who benefits the most from external feedback; c) if there is a relationship between a person's perceptions on the usefulness of the external feedback and their ability to self-monitor their speech.

Insights on the Usability of the Prototype

In line with the previous study, participants expressed positive comments regarding the usability of the prototype. All participants, except for one, said that they would adopt the program to self-administer therapy from home. While only two of these participants met the criteria that we set to use the prototype over eight training sessions, this does not imply that the other participants – and more in general people with aphasia with a similar level of severity – would not have gained any benefit from using it. The inclusion criteria to take part in sessions 8-

17 were that a person would name at least 80 words incorrectly (out of 300, or 26.6%) during both sessions they were presented in. This choice was not dictated by considerations on how severe a person's naming deficit had to be in order to benefit from training; rather, the criteria was set only to satisfy the need to have four sets of words (one per condition), each with 20 items, which could potentially show improvement from baseline during training. This was a trade-off between the need to maximize the confidence in the results on the one hand, and minimizing the amount of training material on the other hand. Also, in this study we did not select the training material based on how relevant the items were to each participant — a feature that would be desirable to have in order to promote a person's motivation to adhere with the necessary frequency and intensity of exercise. However, the program as it is could easily — pending the integration of minor features — allow users to select a set of words of their choice to improve their ability to use them in conversation. This feature would be of use to a vast segment of users who might have milder symptoms than the two participants who could complete this study.

The Major Obstacles Affecting the Usability of the Prototype

We gauged the usability of the program using multiple, complementary metrics: the extent to which users could complete the various tasks independently (effectiveness), the number of mistakes that they committed while doing so (efficacy), their level of satisfaction, and their opinions on the strengths and weaknesses of the prototype.

Results pertaining to usability revealed two main commonalities with findings from the previous study. First, most (two-thirds) of participants could complete all of the tasks independently, or with minimal support. Of the remaining third, partners' support was required for two tasks, while the others could be completed independently. Second, those in need of more support were participants with lower AQs.

Broadly, this study highlighted four main aspects of the design that represented an obstacle to optimal interaction to some participants: inserting the credentials to login into the personal account; repeating a trial; need for a longer learning phase; and a user interface in the naming scene that has too many buttons.

Two participants could not login into their personal account, as they struggled with inserting their credentials. As discussed in the previous section, while this problem was experienced only by two participants, their limited upper-limb dexterity made it so difficult that they could not complete independently access their account. Since limited arm dexterity is a common concern among PwA, it is reasonable to expect that several potential users would encounter this problem. One immediate solution that could be adopted in future studies (wherein passwords will be set by a member of the research team) would be to further simplify the criteria for accessing a personal account, for example by adopting the password format *name.number (i.e.* paul.5*),* as opposed to *name_number* (i.e. paul_P5). This would require typing credentials without having to use a combination of keys (like that required to insert the '_' character on keyboards, or to search across different screens to find it on tablets) while preventing tablets from always forcing the first letter of a proper name to be uppercase. However, thinking beyond the context of research studies, in an hypothetical scenario wherein the program is publicly available to use, and where users are required to set their own personal password, one option to simplify the login process could be to add the option to login via an alternative method, such as fingerprint.

Another function that many participants found difficult to master immediately was repeating a trial. Clearly, despite the button associated with this function having a ‘replay’ icon that is commonly associated with it, this was not enough to make it noticeable to most users. This could be due to the fact that, as soon as the first attempt at naming a word was completed, the icon of the button changed from ‘play’ (the one needed to start a trial) to ‘replay’, and it is likely that users simply did not notice this change. Future versions of the prototype will change this arrangement, positioning the two buttons side-by-side, a few centimeters apart, and activating/deactivating them based on whether the user has to start a new trial or not.

The interviews conducted after participants engaged with the program revealed that some expressed the desire to have longer periods of familiarization with the program prior to the beginning of the usability session proper. These comments suggest that the short live demos that were delivered before participants were asked to engage with the program might not have been enough for some participants to get the gist of how the program worked. Other studies (Simic et al., 2016) assessing the usability of a CBAT dedicated up to 1 hour to training users before beginning the usability session proper. An alternative might be to dedicate a separate study session to user training, where the initial short live demo is followed-up by some time where users try the program and are allowed to request any sort of clarification, until everything is clear to them. The fact that most participants of this study took part only in one usability session, and therefore had the chance to gain very little experience with the program, might explain why some participants found the program not easy to use and expressed the need for support to successfully use it.

Finally, the semi-structured interviews also revealed that, to some participants, the user interface in the naming scene presented too many buttons, whose functions were at times difficult to remember. While this limitation had been pointed out in the previous study, no measures were taken to address it, except for hiding all buttons associated with video-tutorials under a single button which, upon being pressed on, would reveal all of them. A similar strategy could be used by having, by default, the three buttons associated with the hints hidden under a single button called 'Hints', and the two buttons associated with the functions that activate the audio files of the target word and of the recording of the last naming attempt hidden under a button called 'Audio'.

Participants' Perceived Accuracy and Usefulness of the ASR

In contrast with the previous study, wherein all participants commented positively on the usefulness of the ASR, two participants (S1 and S10) said that they did not find it to be useful. This could be explained by the fact that their ability to self-monitor was close to perfect (only one mismatch relative to human raters for S1, and none for S10), therefore the external feedback might have been non-informative to them.

Also, four participants did not comment positively on the accuracy of the ASR, with one of them explicitly saying that it was inaccurate. However, a direct comparison between participants' self-monitoring and the output of the ASR revealed that only in one case (S7) the ASR made several mistakes, while for the other three participants the ASR made only one (failing to detect one error).

This finding is relevant, as it suggests that some people's judgement on the reliability of the ASR is very sensitive to instances where there is a conflict with their judgement. It is important to remark that in this study participants' opinions were based on a very short experience with the program (10 naming attempts), and therefore it is possible that, using the program over an extended period of time, their opinions would have been less sensitive to few incongruences with their judgement. This raises a few questions about what would have happened had these participants taken part in the training phase of the study. Would their judgement on the accuracy of the ASR have been influenced by few instances of misalignment with their judgement? If yes, what strategy could be used to improve their perception of the credibility of the output of the ASR? While we currently do not have answers to these questions, we think it necessary to gather empirical evidence by having this segment of users take part in training studies. In this regard, a direct comparison between S2 and S5 self-monitoring and the output of the ASR provides preliminary (albeit weak) evidence in support of the conjecture that a longer period of engagement with the program would have made participants' judgment less sensitive to sporadic mismatches between these two measures. Such comparison showed that, despite the alignment being not complete, both participants considered the ASR to be both useful and accurate.

Accuracy of the ASR

Out of 681 naming responses that were pooled between all participants across the usability and training sessions (excluding errors of omission), the output of the ASR matched that of humans in 86% of the cases, a result that is line with other promising results obtained by recent attempts to test ASR on datasets of single-word utterances collected by PwA (Abad et al., 2013; Barbera et al., 2021; Le et al., 2018).

During the sessions dedicated to training naming and word-finding (sessions 8-15), the ASR demonstrated good ability to discriminate between correct and incorrect responses, as demonstrated by the a' equaling .89 for both participants S2 and S5.

However, there was a moderate bias towards judging naming responses as correct—in other words, false positives were more likely to occur than false negatives. While this result might be concerning, and points to the need to better characterize the performance of the ASR during extensive training sessions, there are two factors to take into account when interpreting it. The first is to consider how the ASR performed relative to participants' ability to detect their mistakes. A direct comparison revealed that, during all the sessions where the ASR was employed, participants would have gained a clear advantage from relying on its feedback instead of on their judgment. The second is to note that according to data presented in Table 4.17, for participant S5 the number of errors made throughout the sessions where ASR-based feedback was delivered was very low (24 errors). In such cases, a difference of a few units in the amount of false positives/true negatives can make a big difference when computing the bias score.

Considerations on Future Assessment of the Capabilities of the ASR

Future versions of the prototype developed as part of this thesis will hopefully be mature enough to scale up the training material in ways that would afford users the choice to decide by themselves which items to train. In such a situation, there would need to be some changes in the way the ASR is provided information about the target word, and its grammar. As of now, the grammar of a target word was set manually be the person conducting the experiments. This approach is not feasible as the number of users and training material scales up. One possible way to work around this issue is to integrate criteria for the selection of grammar words into the program. For example, the grammar can be composed of two (or more) words which must have the same length as the target, and must fall within a certain range of similarity across the phonological/orthographic or semantic dimensions. This information could be extracted automatically, for each target word, by exploiting well-established databases such as CELEX (https://www.ldc.upenn.edu/language-resources/tools).

Furthermore, moving forward it will be important to expand our understanding of the capabilities of the ASR. One promising avenue is to assess its ability to detect fine mistakes in speech, such phonological paraphasias, which are commonly observed in people with anomia (Dell et al., 1992, 1997; Levelt 1992). As discussed in detail in Appendix C, the ASR determines whether users' naming responses do show a satisfactory match with any of the three words that make up the space of possible responses. If so, the output out the ASR will be the word with the best match, otherwise it will be the word 'SILENCE'. In principle, this feature could be exploited to set up the space of possible responses in a way that includes the phonological neighbors of the target word. This way, if users make a mistake that is phonologically related to the target word (e.g. 'rat' in place of the target 'cat'), and the output of the ASR correctly matches with user's response ('rat'), a more fine-grained feedback on the nature of the mistake could be provided to users. For example, a prompt to 'Try Again' could be accompanied by the incorrect response and the target one, highlighting their difference (e.g. rat, cat). Drawing users' attention to the specific phoneme/syllable that they should correct would likely exert its benefit on a subsequent naming attempt in a way similar to how phonological cues do (i.e., priming the target phonemes by boosting their level of activation and increasing the chances that they will be selected as part of the output word during phonological encoding; Nardo et al., 2017). However, while promising, it is important to note that the higher the similarity among the utterances in the grammar (e.g., 'cat', 'can', 'car'), the harder it is to discriminate the target word from its phonological neighbors. Therefore, it is possible that employing grammars with the target file and its phonological neighbors might increase the number of both false negatives (in case a correct response is erroneously mistaken for a phonological neighbor) or false positives (in case a phonological paraphasia is erroneously mistaken for the target word). For this reason, a thorough investigation of the ASR capabilities in this scenario must be conducted before this feature is deployed.

Efficacy of Self-Administered Treatment

Overall, both participants benefited from self-administering naming therapy using the prototype that we developed. For S2, the gains were immediate, as demonstrated by the difference in naming performance observed between the first probe and the maximum observed during the baseline sessions. The very low and stable performance in the control condition suggests that no generalization or probing effect took place for S2. The large gains in naming performance observed between the first and last probes administered during the training phase suggest that S2 did strongly benefit from using the program. Importantly, these gains were largely maintained at 1-week and 1-month follow-up.

For S5 there were no immediate benefits, as in the very first attempt at naming each word during the first two training sessions naming performance equaled 30% and 35% of the maximum possible score in the no-feedback and feedback conditions, respectively. Performance on the probes and follow-up sessions associated with the control condition also revealed a large improvement relative to baseline. While the improvement in the control condition would normally justify considering generalization of training-induced gains to untrained items as a candidate explanation, the fact that improvements were observed *before* training had taken place (see third baseline scores for the experimental conditions in figure 4.5) goes counter to this hypothesis. A better explanation for the findings could be the well documented variability in item-level naming performance that characterizes impaired access to stored representations (Howard et al., 1985, 2015; Mirman & Britt, 2014). This phenomenon can introduce confounds in the results due to *regression to the mean* – whereby, when measuring behaviors that are known to be inconsistent from time to time, initial extreme performances get, on subsequent measurements, closer to a “mean” value that is characteristic to that behavior. Therefore, this study, in which words were selected based on the first two attempts at naming, was exposed to this problem. To prevent it, best practices recommend first administering a stimulus selection phase, employing two sessions to select, from a large set, words that are difficult to name. Subsequently, a baseline phase proper can begin, consisting of two or more sessions, wherein the selected words are further tested to get a better estimate of a participant’s expected performance (Howard et al., 2015). While the original plan for this study was to employ this method, we decided against it because of the already substantial number of study sessions. An alternative could have been to reduce the number of sessions dedicated to stimuli selection from five to two (sessions 2-3), administering 300 items per session instead of 120, selecting 80 stimuli based on naming performance and dedicating sessions 4-6 to assess the naming performance specifically on this restricted pool of words. Nonetheless, we decided against this option because it would have resulted in sessions 2-3 lasting about 3 hours, which many prospective participants and therapists advised against during the pilot phase. However, while these decisions were made to alleviate the burden on participants, this reduced the level of rigor and control that was necessary to conduct a quality study. In light of these constraints on the number and length of study sessions, a viable option that should be considered for future studies could be to reduce the size of the set of stimuli that is tested during the stimulus selection phase. This proposal is based on the consideration that, in this study, there was a striking gap between the naming performance of the three participants that were eligible to take part in the training phase, with a rate of errors made on the same word twice of at least 30%, and those who were not eligible, whose highest error rate was 8% (see Table 4.15). Assuming that, as in this study, 80 stimuli are needed to make up for the experimental and control conditions, and assuming 30% error rate as a lower threshold for participants to partake in the training phase, the size of the original set of stimuli could be reduced from 300 to 240. The 240 stimuli would have to be repeated twice during the stimuli selection phase, for a total of 480 trials. These could be distributed over three sessions (160/session, from sessions 2-4), which would lengthen each session by 25% of the time (approximately 75 minutes versus the original 60 minutes). Then, the selected 80 stimuli could be further tested over the course of three additional short sessions (sessions 5-7) which would last far less than 1 hour and which would result in the addition of only one extra session. We believe that it is reasonable to expect that most participants would tolerate these changes.

Despite the high performance observed in the control condition, naming performance in the set of words that were practiced via training demonstrated a consistent benefit relative to no-training, even during the follow-up sessions. These results are comparable to those reported in the study conducted by Ballard et al. (2019), although the two studies differed in some key respects. In fact, while both employed a training set of 20 words, the amount of training that participants received was comparable only for one of the participants (P4, who completed about 230 trials throughout the study, while S2 and S5 completed approximately 240 for any given set of words), while others in the Ballard study had the opportunity to practice between 4 and 10 times more than the two participants of this study. On the other hand, comparing these results with those of Palmer et al. (2019, 2020) presents some challenges. Similar to the work of Ballard et al., in Palmer et al. there was high variability in the extent to which participants engaged with the computerized therapy. In this study participants were recommended to engage with training for 20-30 minutes per day for a period of six months — for an expected total amount of time spent training of about 90 hours. Therefore, by default the study was meant to expose participants with a dosage of therapy that was one order of magnitude higher than in the study presented in this chapter. Despite most participants did not abide by the recommended levels of practice, as shown by the fact that the InterQuartile Range (IQR) for the time spent practicing was 4.9–49.7 hours, while the IQR for the number of times participants logged in to start a training session was 14–100, these results indicate that about 75% of participants engaged with training between 3 and 20 times more than in our study. Also, Palmer and colleagues employed 100 words in the training set, and reported an improvement of approximately 16.4% (16 words) over a baseline performance around 45%.

The Role of External Feedback and Self-Monitoring Abilities

Comparing the naming performance observed during the training and follow-up phases did not reveal any clear difference between the two experimental conditions (with external feedback, and without). One possible explanation for this result is that participants' ability to self-monitor their responses during the no-feedback condition was comparable to how accurately the ASR scored their responses in the feedback condition. The results of the signal detection analysis revealed that S2 and S5 had a moderate and strong bias towards judging their responses as correct, respectively. In other words, they tended to overestimate their performance. Both participants rarely judged one of their correct responses as wrong. This suggests that participants' deficit in self-monitoring was due almost entirely to a difficulty in detecting their errors, as opposed to incorrectly labelling correct responses as errors.

Interestingly, when engaging with training in the absence of external feedback, S2 showed a very similar ability to discriminate correct vs incorrect responses and a very similar bias relative to the ASR. This finding could explain the absence of a clear difference in outcome measures between the two experimental conditions. If the ability to detect mistakes is crucial to determine when a specific trial needs to be repeated, and therefore it positively impacts learning, relying on information with similar levels of discrimination accuracy and bias (self-assessment and ASR) should yield similar learning outcomes. In addition, it is important to note that, when comparing the bias of the ASR with that of S2's self-monitoring *within* the feedback condition, the latter was much higher. Therefore, it is likely that during the feedback condition the naming performance of S2 benefitted from receiving external feedback.

Regarding S5, the number of errors (excluding errors of omission) was low, oscillating between 4 and 10 throughout the training sessions. Therefore, the lack of a clear difference between the experimental conditions can be explained by the fact that there were not many opportunities for the external feedback to exert its benefits.

Despite the lack of any obvious gains in naming performance for the items that were trained with the provision of external feedback, both S2 and S5 thought that it was useful.

Lessons Learned and Next Steps

There are four main take-home messages from this study. First, by and large we validated the positive feedback that we received on the previous study, using a larger sample of participants. Most participants found the program easy to use and learn, and expressed an interest in using improved versions to self-administer naming therapy from home. We also gathered additional information on how to further refine the user interface, pointing to the need to: 1) make credentials more aphasia-friendly; 2) fix bugs in video-tutorials and make them more aphasia-friendly; 3) reduce the density of buttons in the naming scene; 4) allow participants a longer period of familiarization with the program before testing the usability. The latter point deserves particular attention, as it was brought about also in the previous study, and evidently not adequately addressed. The need to rethink how participants are trained on how to use the program was particularly evident when observing participants with moderate aphasia, who indeed relied on some form of support (mostly watching video tutorials) when engaging with most tasks. This could be due to the fact that they could not memorize all the information that was provided to them during the brief demonstration. While we do not have any data to suggest what might have caused difficulties in learning the various functions of the program (be it the presence of concomitant cognitive deficits, or difficulties in understanding long conversations, which was not tested as part of the WAB-R), the findings of this study suggest that future studies should devise and test ways to ensure that participants have grasped the most important features of the program before engaging with it for therapy. Aside from improving the quality of the video tutorials, one possibility to make them more effective at teaching how to use the program is to design interactive exercises to test users' ability to use specific functions when needed. To do so, participants could engage with a program that mimics all stages of the prototype. In each stage, a short video tutorial explaining the specific tasks and functions would be followed by instructions, provided by the program, to perform the target action (i.e. 'Press the button to start the trial', 'Press the button to repeat the trial'), along with feedback on whether the action was correct or not.

Second, while in the previous study participants engaged with the prototype for a very short period of time (as the idea was to capture the very first impression on what needed to be improved in the user interface, in line with other studies (Ballard et al., 2019; Simic et al., 2016)) we also investigated usability in the intended use scenario, involving participants who used the prototype to self-administer naming therapy over eight sessions. This way we could assess whether users can manage to effectively self-administer therapy, leveraging the appropriate types of support afforded by the program when necessary, and we could benefit from their much more informed opinions on the strengths and weaknesses of the program. Importantly, we observed that participants could easily learn to use the prototype and adopt it to effectively self-administer naming therapy at-home, and provided positive comments on the usability of the program.

Third, one of the key contributions of this thesis was to integrate an ASR — a technology that has been little used in the field of aphasia rehabilitation — into a computerized therapy for aphasia. Since this was one of the first attempts to adopt this technology in this context, chapters 3 and 4 of this thesis aimed to investigate not only how accurately the ASR scores users' naming attempts, but also users' perceptions about the usefulness and accuracy of the ASR. In this regard, by and large results were very promising: the ASR yielded high levels of discriminability and acceptable levels of bias when compared to human raters. Also, most participants, which importantly included the two who engaged with the prototype throughout the training sessions, reported that they found feedback to be accurate and useful.

Fourth, this study represented also a first attempt to assess whether participants could benefit from self-administering the prototype to improve their naming and word-finding abilities, and to explore the conditions under which automatically delivered feedback might or might not be needed. In this regard, the outcomes were positive: both participants demonstrated improvements while administering therapy, with gains largely maintained 1-month after the end of training. However, while results should be interpreted very cautiously —especially in light of the small sample size — one lesson learned from this thesis is that participants could make comparable improvements in naming accuracy when training with and without the support of external feedback. This does not imply that feedback was useless. In fact, both participants showed poorer ability to detect their mistakes in the conditions where feedback was provided (a finding that cannot be explained by an effect of feedback, as it was provided after subjects made a judgement on the accuracy of their response) suggesting that, in this specific case, had feedback not been provided, they might have shown poorer naming performance in one condition. Moving forward, to better illuminate the role of self-monitoring abilities in driving recovery in the absence of external feedback, future work should combine signal detection analysis, with an item-level analysis of how error detection (and correction) on a given trial affects naming accuracy during subsequent trials. In addition, it is important to consider how a person's confidence in their own judgement, along with the amount of trust they put on external feedback, affects the extent to which they benefit from external feedback. Moreover, it will be important to recruit not only more people who can complete the study, but also people with more severe naming and self-monitoring deficits than S2 and S5.

Finally, as part of future versions of the task targeting word-finding and naming abilities, reaction times (RTs) should be considered as an outcome measure, in addition to naming accuracy. While not frequently considered in studies assessing the efficacy of methods for remediating naming skills, RTs, as a measure of speed of processing, provide unique information on the integrity of a word's memory representation, over and above that provided by naming accuracy (Conroy et al., 2018; de Grosbois et al., 2023). For example, while in the absence of information on RTs two conditions yielding similar gains in accuracy would be considered to have similar efficacy, the presence of significant differences in RTs would suggest that the condition associated with faster retrieval might promote a more effective encoding and strengthening of a word's memory representation. Regarding the studies described in chapter 3 and 4 of this thesis, while they were not originally planned to consider RTs as an outcome measure, it could have been possible to perform this analysis retrospectively, had I not set up the task in such a way that, between the moment a picture appeared on the screen and the moment participants were prompted to speak, an interval of three seconds was inserted in order to give participants some time to prepare their response, releasing some pressure from the task. Therefore, analyzing the data collected so far would not provide reliable information on speed of processing, as this process was distorted by the way the task was structured. Future investigations can prevent this by instructing participants to provide their response as soon as they feel ready to, and by activating the microphone in correspondence of (or slightly before) the appearance of picture stimuli on the screen.

Limitations

The main limitation of this study was the very limited sample size, which did not allow us to elucidate the relationship between self-monitoring and the extent to which participants benefit from external feedback, a key objective of the study. Our plan for addressing this research question exhaustively was to test the hypothesis that the two variables are correlated. According to the power analysis, 28 participants were required to detect a big effect with a level of statistical power of 80%. As we were able to recruit only two participants, we could not address these two questions satisfactorily.

Aside from severely limiting the depth at which this study could address the research questions that it was conceived to tackle, the fact that, despite having been able to reach out to two vast groups of potential participants via our connections with charities, only two participants were eligible to complete the entire study might legitimately raise questions about the utility of such a product. However, there are reasons to believe that the low number of subjects was not due to attrition (an issue that will be addressed in the next chapter of this thesis as it has been documented in some commercially available CBATs, and therefore would likely be of extreme relevance for future versions of the product presented in this thesis), a lack of interest from prospective participants, or to the fact that the product as-is currently does not address any relevant need of PwA. Instead, we think the low sample size was due to the fact that a) we employed eligibility criteria that were unnecessarily stringent, and b) most prospective participants had very mild symptoms of aphasia (i.e. eight out of 12 participants showed error rates lower than 12%, one around 18%, over 600 naming attempts (see Table 4.15)). Regarding the first point, the inclusion and exclusion criteria can be relaxed without compromising the quality of a study. In fact, in hindsight having a criterion by which subjects were allowed to self-administer therapy only if their performance showed consistent item-specific difficulty over *only* two sessions might have been detrimental for the quality of the study. As anomia results from impaired access to words, as opposed to loss of the associated memory representations, retrieving a given word is not entirely precluded, rather it is made difficult by specific properties of the target word, such as word length, frequency, and abstraction: the longer, less frequent, more abstract a word, the harder it is to retrieve it. In other words, a word might be named correctly on a given session, and incorrectly in most others, and vice versa. On the other hand, what tends to stay constant across people with anomia is the overall error rate. Therefore, one commonly used strategy when selecting the training material is to include words that are named incorrectly at least twice over three or more sessions (Howard et al., 2015). This would have likely expanded the range of participants who could have been eligible to complete this study. Another viable way to multiply the pool of potential participants would be to establish a relationship akin to the one that was forged with the Aphasia Institute of Ontario with as many aphasia centers located in North America as it is possible.

While we consider it likely that the product currently does not address any relevant need of PwA — as in this study the training stimuli were selected solely on the basis of whether or not participants could name them during two baseline sessions, and probably had little to no relevance to participants — this did not contribute to the low sample size. In fact, all of the prospective participants who went through sessions 2-6 expressed an interest in completing the entire study, and if they did not it was only because they were actively excluded based on their baseline performance, not on them refusing to continue due to the low interest in the training material. However, we believe that it is all the more important to ensure that the real-life needs of prospective users are taken into account by allowing participants to indicate which stimuli should be introduced in the training material (Chapey et al., 2000; Kagan et al., 2008; Simmons-Mackie, 2017). This will have at least two benefits. *First*, it would ensure that any improvements in naming skills have a chance to translate into real-life scenarios (i.e. making specific requests, having conversations, etc.) in which participants would like to be more proficient. Relatedly, it is important to measure whether people with aphasia show any meaningful improvements in such scenarios (by administering tools for the functional assessment of communication abilities (Frattali et al., 2017)), and whether the treatment has had any effect on their subjective experience on their quality of life (Hilari et al., 2003). Second, and strictly related to the former point, adopting highly salient stimuli leads to increased engagement with the task, favouring acquisition and maintenance of gains (Kiran & Thompson, 2019), and incentivizes people to adhere with the frequency and levels of intensity that are required to maximize treatment-induced gains (Biel & Haley, 2023).

Another limitation was that the measurement of baseline performance did not follow best practices, and this prevented us from getting an accurate estimate of baseline naming performance for S5 before the beginning of the study. As discussed above, future studies should ensure that, if a set of words is selected based on participants' performance, it is subsequently tested at least three more times before the beginning of the training phase. This way, it will be possible to get an adequate estimate of the level, trend, and amount of variability characterizing the baseline naming performance, and to use these metrics as a reference to compare data from the other phase against.

Finally, related to best practices, a bug in the prototype did not allow us to collect data from the second control condition, which was designed to be administered only during the baseline and follow-up sessions. While, as discussed above, we have reasons to think that no generalization nor probing effect took place, had we encountered a participant showing some changes in level and trend in the control data. In such a case, only this additional measure would have ensured a level of control requested to characterize what phenomenon is responsible for the observed performance.

CONCLUSIONS

This study represented a step ahead in the process of testing the usability of the program. Results not only confirmed the positive findings of the previous study, but also suggest that people with aphasia could use the program in the intended use case scenario independently, or with minimal support. Furthermore, most participants expressed satisfaction with their experience, and said that they would use an improved version of the program to self-administer therapy.

This study also set out to assess whether people with aphasia benefit from using the computerized therapy, and to characterize in which cases, and to what extent, ASR-based feedback can be beneficial. Only two participants could complete the entire study: both of them benefited from using the program, with gains of 65% and 30% at 1-month after the end of therapy. However, no clear advantage from receiving the ASR-based feedback was observed, a result that can be explained by the fact that participants' ability to detect their mistakes was comparable to that of the ASR.

While these results might suggest that provision of external feedback might not always be necessary, a major limitation of this study was the inadequate sample size to address this question. Future studies must include larger pools of participants with a wide spectrum of self-monitoring and naming difficulties.

The next, final chapter of this thesis will elaborate further on the findings and limitations encountered throughout the studies presented in this dissertation, discussing them in relation to the wider literature and exploring possible ways to move ahead with the development of a computerized therapy for aphasia.

REFERENCES

Abad, A., Pompili, A., Costa, A., Trancoso, I., Fonseca, J., Leal, G., Farrajota, L., & Martins, I. P. (2013). Automatic word naming recognition for an on-line aphasia treatment system. *Computer Speech and Language*, *27*, 1235–1248.

Ballard, K., J, Etter, N., M., Shen, S., Monroe, P., & Tien Tan, C. (2019). Feasibility of Automatic Speech Recognition for Providing Feedback During Tablet-Based Treatment for Apraxia of Speech Plus Aphasia. *American Journal of Speech-Language Pathology*, *28*, 818–834.

Barbera, D. S., Huckvale, M., Fleming, V., Upton, E., Coley-Fisher, H., Doogan, C., Shaw, I., Latham, W., Leff, A. P., & Crinion, J. (2021). NUVA: A naming utterance verifier for aphasia treatment. *Computer Speech and Language*, *69*, 1–25. https://doi.org/10.1016/j.csl.2021.101221

Biel, M., & Haley, K. L. (2023). Motivation in Aphasia Treatment: Self-Determination Theory Applied to the FOURC Model. *American Journal of Speech-Language Pathology*, *32*(3), 1016–1036.

Braley, M., Sims Pierce, J., Saxena, S., De Oliveira, E., Taraboanta, L., Anantha, V., Lakhan, S. E., & Kiran, S. (2021). A Virtual, Randomized, Control Trial of a Digital Therapeutic for Speech, Language, and Cognitive Intervention in Post-stroke Persons With Aphasia. *Frontiers in Neurology*, *12*(626780). https://doi.org/10.3389/fneur.2021.626780

Brooke, J. (1986). *System Usability Scale (SUS)*.

Byiers, B. J., Reichle, J., & Symons, F. J. (2012). Single-Subject Experimental Design for Evidence-Based Practice. *American Journal of Speech-Language Pathology*, *21*(4), 397–414. https://doi.org/10.1044/1058-0360(2012/11-0036)

Chapey, R. (2008). *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders* (5th ed.). Lippincott Williams and Wilkins.

Chapey, R., Duchan, J., Elman, R., Garcia, L., Kagan, A., Lyon, J., & Simmons-Mackie, N. (2000). *Life Participation Approaches to Aphasia: A statement of values for the future.* (5; pp. 4–6). ASHA Leader.

Cheng, X., Schafer, G., & Akyürek, E. G. (2010). Name agreement in picture naming: An ERP study. *Official Journal of the International Organization of Psychophysiology*, *76*(3), 130–141.

Conroy, P., Sotiropoulou Drosopoulou, C., Humphreys, G. F., Halai, A. D., & Lambon Ralph, M. A. (2018). Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia. Brain, 141(6), 1815-1827

Constant Therapy. (2022, February). https://constanttherapyhealth.com/science/

Cuetos, F., Aguado, G., Izura, C., & Ellis, A. W. (2002). Aphasic naming in Spanish: Predictors and errors. *Brain and Language*, *82*(3), 344–365.

de Grosbois, J., Canthiya, L., Philipp-Muller, A. E., Hickey, N. K., Hodzic-Santor, B., Heleno, M. C., ... & Meltzer, J. A. (2023). Asynchronous, online spaced-repetition training alleviates word-finding difficulties in aphasia. Neuropsychological Rehabilitation, 33(10), 1672-1696.

Dallery, J., Cassidy, R. N., & Raiff, B. R. (2013). Single-Case Experimental Designs to Evaluate Novel Technology-Based Health Interventions. *Journal of Medical Internet Research*, *15*(2), 1–17. https://doi.org/10.2196/jmir.2227

Dell, G. S., & O'Seaghdha, P. (1992). Stages of lexical access in language production. *Cognition*, *42*, 287–314.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review*, *104*(4), 801–838.

Doesborgh, S. J. C., van de Sandt-Koenderman, M. W. M. E., Dippel, D. W. J., van Harskamp, F., Koudstaal, P. J., & Visch-Brink, E. G. (2004). Cues on request: The efficacy of Multicue, a computer program for word finding therapy. *Aphasiology*, *18*(3), 213–222. https://doi.org/10.1080/02687030344000580

Fink, R. B., Brecher, A., Schwartz, M. F., & Robey, R. R. (2002). A computer-implemented protocol for treatment of naming disorders: Evaluation of clinicianguided and partially self-guided instruction. *Aphasiology*, *16*, 1061–1086.

Frattali, C. M., Thompson, C. K., Holland, A. L., Wohl, C. B., Wenck, C. J., Slater, S. C., & Paul, D. (2017). *American Speech-Language-Hearing Association Functional Assessment of Communication Skills for Adults (ASHA FACS)*. American Speech-Language-Hearing Association.

Glen, S. (n.d.). Binomial Confidence Interval. *From StatisticsHowTo.Com: Elementary Statistics for the Rest of Us!* Retrieved September 27, 2023, from https://www.statisticshowto.com/binomial-confidence-interval/

Graham, J. E., Karmarkar, A. M., & Ottenbacher, K. J. (2012). Small Sample Research Designs for Evidence-based Rehabilitation: Issues and Methods. *Archives of Physical Medicine and Rehabilitation*, *93*(8), 111–116. https://doi.org/10.1016/j.apmr.2011.12.017

Harrison, M., Palmer, R., & Cooper, C. (2020). Factors Associated With Adherence to Self-Managed Aphasia Therapy Practice on a Computer—A Mixed Methods Study Alongside a Randomized Controlled Trial. *Frontiers in Neurology*, *11*.

Herbert, R., Webster, D., & Dyson, L. (2012). Effects of syntactic cueing therapy on picture naming and connected speech in acquired aphasia. *Neuropsychological Rehabilitation*, *22*(4), 609–633. https://doi.org/10.1080/09602011.2012.679030

Hilari, K., Byng, S., Lamping, D. L., & Smith, S. C. (2003). Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39) Evaluation of Acceptability, Reliability, and Validity. *Stroke*, *34*, 1944–1950. https://doi.org/10.1161/01.STR.0000081987.46660.ED

Howard, D., Best, W., & Nickels, L. (2015). Optimizing the design of intervention studies: Critiques and ways forward. *Aphasiology*, *29*(5), 526–562. https://doi.org/10.1080/02687038.2014.985884

Howard, D., Patterson, K. E., Franklin, S. E., Orchard-Lislie, V. M., & Morton, J. (1985). The treatment of word retrieval deficits in aphasia: A comparison of two therapy methods. *Brain*, *108*, 817–829. https://doi.org/10.1093/brain/108.4.817

Kagan, A., Simmons-Mackie, N., Rowland, A., Huijbregts, M., Shumway, E., McEwen, S., Threats, T., & Sharp, S. (2008). Counting what counts: A framework for capturing real-life outcomes of aphasia intervention. *Aphasiology*, *22*(3), 258–280. https://doi.org/10.1080/02687030701282595

Kiran, S., & Thompson, C. K. (2019). Neuroplasticity of language networks in aphasia: Advances, updates, and future challenges. *Frontiers in Neurology*, *10*, 295.

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology*, *25*(4), 463–492.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-Case Intervention Research Design Standards. *Remedial and Special Education*, *34*(1), 26–38. https://doi.org/10.1177/0741932512452794

Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the Scientific Credibility of Single-Case Intervention Research: Randomization to the Rescue. *Psychological Methods*, *15*(2), 124–144. https://doi.org/10.1037/a0017736

Laganaro, M., Di Pietro, M., & Schnider, A. (2003). Computerised treatment of anomia in chronic and acute aphasia: An exploratory study. *Aphasiology*, *17*(8), 709–721. https://doi.org/10.1080/02687030344000193

Lavoie, M., Macoir, J., & Bier, N. (2017). Effectiveness of technologies in the treatment of post-stroke anomia: A systematic review. *Journal of Communication Disorders*, *65*, 43–53.

Le, D., Licata, K., & Provost, E. M. (2018). Automatic quantitative analysis of spontaneous aphasic speech. *Speech Communication*, *100*, 1–12. https://doi.org/10.1016/j.specom.2018.04.001

Levelt, W. J. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, *42*(1–3), 1–22.

Manolov, R., & Moeyaert, M. (2017). How Can Single-Case Data Be Analyzed? Software Resources, Tutorial, and Reflections on Analysis. *Behavior Modification*, *41*(2), 179–228. https://doi.org/10.1177/0145445516664307

Marshall, R. C., Neuburger, S. I., & Phillips, D. S. (1994). Verbal self-correction and improvement in treated aphasic clients. *Aphasiology*, *8*(6), 535–547.

Marshall, R. C., & Tompkins, C. A. (1982). Verbal Self-Correction Behaviors of Fluent and Nonfluent Aphasic Subjects. *Brain and Language*, *15*, 292–306.

Middleton, E. L., Schwartz, M. F., Dell, G. S., & Brecher, A. (2022). Learning from errors: Exploration of the monitoring learning effect. *Cognition*, *224*, 1–16.

Mirman, D., & Britt, A. E. (2014). What we talk about when we talk about access deficits. *Philosophical Transaction of the Royal Society B*, *369*. https://royalsocietypublishing.org/doi/10.1098/rstb.2012.0388

Mortley, J., Wade, J., & Enderby, P. (2004). Superhighway to promoting a client-therapist partnership? Using the Internet to deliver word-retrieval computer therapy, monitored remotely with minimal speech and language therapy input. *Aphasiology*, *18*(3), 193–211. https://doi.org/10.1080/02687030344000553

Nardo, D., Holland, R., Leff, A. P., Price, C. J., & Crinion, J. T. (2017). Less is more: Neural mechanisms underlying anomia treatment in chronic aphasic patients. *Brain*, *140*(11), 3039–3054.

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

Nunn, K., & Vallila-Rohter, S. (2023). Theory-driven treatment modifications: A discussion on meeting the linguistic, cognitive, and psychosocial needs of individual clients with aphasia. *Journal of Communication Disorders*, *103*. https://doi.org/10.1016/j.jcomdis.2023.106327

Nunn, K., Vallila-Rohter, S., & Middleton, E. L. (2023). Errorless, Errorful, and Retrieval Practice for Naming Treatment in Aphasia: A Scoping Review of Learning Mechanisms and Treatment Ingredients. *Journal of Speech, Language, and Hearing Research*, *66*(2), 668–687. https://doi.org/10.1044/2022_JSLHR-22-00251

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M. C., Bowen, A., Latimer, N., Julious, S., Cross, E., Alshreef, A., Harrison, M., Bradley, E., Witts, H., & Chater, T. (2019). Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): A multicentre, single-blinded, randomised controlled trial. *Lancet Neurology*, *18*, 821–833.

Palmer, R., Dimairo, M., Latimer, N., Cross, E., Brady, M. C., Enderby, P., Bowen, A., Julious, S., Harrison, M., Alshreef, A., Bradley, E., Bhadhuri, A., Chater, T., Hughes, H., Witts, H., Herbert, E., & Cooper, C. (2020). Computerised speech and language therapy or attention control added to usual care for people with long-term post-stroke aphasia: The Big CACTUS three-arm RCT. *Health Technology Assessment*, *24*(19). https://doi.org/10.3310/hta24190

Palmer, R., Enderby, P., & Paterson, G. (2013). Using computers to enable self-management of aphasia therapy exercises for word finding: The patient and carer perspective. *International Journal of Language & Communication Disorders*, *48*(5), 508–521. https://doi.org/10.1111/1460-6984.12024

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of All Non-Overlapping Data (PAND): An Alternative to PND. *The Journal of Special Education*, *40*(4), 194–204.

Ramsberger, G., & Marie, B. (2007). Self-Administered Cued Naming Therapy: A Single-Participant Investigation of a Computer-Based Therapy Program Replicated in Four Cases. *Americal Journal of Speech-Language Pathology*, *16*, 343–358. https://doi.org/1058-0360/07/1604-0343

Schwartz, M. F., Middleton, E. L., Brecher, A., Gagliardi, M., & Garvey, K. (2016). Does naming accuracy improve through self-monitoring ofe rrors? *Neuropsychologia*, *84*, 272–281.

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing Single-Subject Research: Issues and Applications. *Behavior Modification*, *22*(3), 221–242.

Simic, T., Leonard, C., Laird, L., Cupit, J., Höbler, F., & Rochon, E. (2016). A Usability Study of Internet-Based Therapy for Naming Deficits in Aphasia. *American Journal of Speech-Language Pathology.*, *25*, 642–653.

Simmons-Mackie, N. (2017). *Aphasia in North America*. https://www.aphasiaaccess.org/white-papers/

Simmons-Mackie, N., Damico, J. S., & Damico, H. L. (1999). A Qualitative Study of Feedback in Aphasia Treatment. *Americal Journal of Speech-Language Pathology*, *8*(3), 218–230.

Sindelar, P. T., Rosenburg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children.*, *8*, 67–76.

Swales, M. A., Hill, A. J., & Finch, E. (2016). Feature rich, but user-friendly: Speech pathologists' preferences for computer-based aphasia therapy. *International Journal of Speech-Language Pathology*, *18*(4), 315–328. https://doi.org/10.3109/17549507.2015.1081283

Sze, W. P., Hameau, S., Warren, J., & Best, W. (2020). Identifying the Components of a Successful Spoken Naming Therapy: A Meta-Analysis of Word-Finding Interventions for Adults with Aphasia. *Aphasiology*, *35*(1), 33–72.

The International Picture-Naming Project. (n.d.). https://crl.ucsd.edu/experiments/ipnp/

Thompson, C. K. (2006). Single subject controlled experiments in aphasia: The science and the state of the science. *Journal of Communication Disorders*, *39*, 266–291.

Valente, A., Bürki, A., & Laganaro, M. (2014). ERP correlates of word production predictors in picture naming: A trial by trial multiple regression analysis from stimulus onset to response. *Frontiers in Neuroscience*, *8*(390), 1–13.

What Works Clearinghouse. (2023). https://ies.ed.gov/ncee/wwc/

Zheng, C., Lynch, L., & Taylor, N. (2016). Effect of computer therapy in aphasia: A systematic review. *Aphasiology*, *30*(2–3), 211–244. https://doi.org/10.1080/02687038.2014.996521

Chapter 5 – DISCUSSION

This dissertation is part of a broader project whose mission is to develop a CBAT that improves currently available applications, by focusing on usability and — critically — feedback based on automated speech recognition (ASR). The research presented here focused on the initial user-centered design of a CBAT prototype, usability testing, and preliminary testing of the efficacy of the ASR-based feedback.

In this context, the chapters of this thesis document the first steps taken towards the development of a product that harnesses recent improvements on state-of-the-art ASR to give users feedback on how well they speak. Specifically, chapter 2 provides a detailed characterization of the most common unmet needs of PwA. Furthermore, with an eye on informing the design of a prototype of CBAT that is aphasia-friendly, this chapter provides a detailed examination of the desirable features of CBATs, along with the barriers and facilitators that PwA experience when using the modern technologies subserving such applications. Chapter 3 describes the key features of the first prototype of CBAT featuring an ASR, along with an investigation of its usability conducted with PwA. Finally, chapter 4 describes a preliminary assessment of a) the feasibility of using it to enable PwA to self-administer naming therapy independently, from their home, and b) of the benefit that they reap from doing so.

In this chapter, I will first summarize the main findings of the research studies, and then consider the general learnings and future directions informed by this research.

Summary of Findings by Chapter

Chapter 2 — Informing the Design of At-Home Therapies for Aphasia: A Needs Finding Analysis

This study, which combined a literature review with semi-structured interviews conducted with PwA and speech-language therapists, revealed four main findings.

First, both groups of stakeholders confirmed that adopting an ASR to deliver real-time feedback and video games to promote engagement and motivation were desirable and promising innovations.

Second, analysis of users' opinions regarding their favorite devices revealed no clear preference at the group level, with laptops and tablets being both popular among PwA. This finding suggested that the best course of action was to develop a product whose user interface is designed to work optimally on these devices (i.e., be clearly visible on 9.7 inch screens, those of an average tablet, and be easy to use both via touchpad and mouse).

Third, a detailed analysis of the desirable features of CBATs highlighted: the importance of providing a variety of tasks and stimuli, and to ensure that they appeal to the personal tastes of each user; that tasks present the right level of challenge to users, by measuring their progress and adapting the level of difficulty accordingly; that information on item-level performance is used to provide real-time feedback, and to provide summary statistics that update users on their overall progress as it accrues through practice; that are user-friendly and easily accessible. Regarding the latter point, the application was designed to be accessible from the web, such that users could launch it by simply copy/pasting an URL on their web browser, as opposed to having to install and set up a program on their personal device.

Fourth, this study provided insights into how CBATs can contribute to transitioning to a patient self-management model (Nichol et al., 2019; Nichol, Rodriguez, et al., 2023; Palmer et al., 2019) by performing multiple functions throughout the continuum of care, over and above enabling the self-administration of therapy. These insights will be useful at later stages of development, once the feasibility and efficacy of self-administering therapy under the guidance of ASR-based feedback will be strongly tested and validated.

Chapter 3 — Usability Study of the First Prototype of a Computer-Based Therapy for Aphasia Featuring Automatic Speech-Recognition

As a first step towards the development of a CBAT, results from the previous study were used to inform the design of a minimum viable product featuring ASR and a simple, practice-based approach to remediate word-finding abilities. The goal of this study was to identify any design features that might negatively affect the usability of the first version of the prototype; to assess users' perceptions about the accuracy and usefulness of the ASR; and to assess the accuracy of ASR responses relative to that of human raters.

This study revealed that most users could learn to use the program and complete all tasks independently, albeit in some cases requiring minimal support from video tutorials. Importantly, there was a consensus among participants that the program was well designed, and that they would use it frequently. Nevertheless, it was also observed that while participants with milder symptoms expressed high levels of satisfaction with the usability of the prototype, those with moderate severity of aphasia indicated that they did not find it easy to remember all the key functions of the program following the brief demonstration that was provided before the beginning of the usability session. For this reason, this segment of users would most likely benefit from more extensive training on how the program works, before engaging with it.

Relatedly, the main recommendation gathered from users was to improve the quality of the video tutorials.

Finally, a comparison of the output of the ASR with the judgment of human raters on the vocal responses recorded during the usability session showed that the ASR proved to be excellent at scoring words, with high levels of discriminability. However, a moderate bias towards judging naming responses as correct was also documented. Importantly, all participants declared that the feedback from the ASR was useful and, except for two individuals, accurate.

Chapter 4 — Improved Word-Finding and Naming Abilities Following Self-Administered Computer-Based Therapy

This study extended the scope of the previous one by addressing novel questions about the feasibility, and effectiveness, of using the prototype of CBAT to self-administer therapy from home. In this regard, the main objectives were to assess: a) whether PwA could benefit from using the current version of the prototype to self-administer therapy targeting word-finding and naming skills from their home; b) if, and to what extent, the provision of external feedback via an ASR resulted in additional benefits relative to a situation wherein users have to rely on their self-monitoring abilities, and whether the latter predict who benefits the most from external feedback; c) if there is a relationship between a person's perceptions on the usefulness of the external feedback and their ability to self-monitor their speech.

This study had four main take-home messages. First, by and large results from the previous study were validated, as most participants found the program easy to use and learn, could use it independently or with minimal support, and expressed an interest in using improved versions to self-administer naming therapy from home. We also identified major aspects of the user interface that need to be refined: making credentials more aphasia-friendly; fix bugs in video tutorials and make them more aphasia-friendly; reduce the density of buttons in the naming scene; allow participants a longer period of familiarization with the program before testing the usability.

Second, participants who took part in the eight training sessions could manage to effectively self-administer therapy independently, leveraging the appropriate types of support afforded by the program when necessary, and provided positive comments on the usability of the program. This demonstrates the feasibility of using it to enable PwA to self-administer naming therapy independently, from their home.

Third, both participants benefited from self-administering therapy, with gains in naming and word-finding abilities largely maintained 1-month after the end of training.

Fourth, we observed no clear advantage for providing external feedback. This might be due to the fact that both participants demonstrated an ability to judge the accuracy of their naming attempts which was comparable to that of the ASR.

Limitations, Lessons Learned, and Future Work

Human Factors

Usability

The main lesson learned from the assessments of the usability of the prototype conducted in the studies described in chapter 3 and 4 is that most participants were able to use the program independently (albeit in some cases relying on the use of video tutorials to successfully complete some tasks), after only a brief demonstration. Importantly, the two participants who took part in the training phase were also able to self-administer therapy and complete sessions 8-17 fully independently, and demonstrated a clear mastery of the various functions that were available to support their naming (i.e., cueing, playing audio of another person, replaying their own voice).

These findings suggest that the current version of the prototype can serve as a basis for studies testing the efficacy of the application in a context that resembles more that in which users are expected to self-administer therapy (i.e. longer amount of training sessions, using stimuli and tasks that are salient, etc.).

Another important lesson learned was that, although participants could successfully complete all tasks on their own, those with moderate aphasia required support with most tasks. This could be due to the fact that they could not memorize all the information that was provided to them during the brief demonstration. While we do not have any data to suggest what might have caused difficulties in learning the various functions of the program (be it the presence of concomitant cognitive deficits, or difficulties in understanding long conversations, which was not tested as part of the WAB-R), the findings of this thesis suggest that future studies should devise and test ways to ensure that participants have grasped the most important features of the program before engaging with it for therapy. In this regard, a survey of the literature on usability suggests that a potentially effective strategy to improve the effectiveness of user training would be to use interactive walkthroughs, which essentially complement the guidance provided by video-tutorials with exercises to test users' ability to use specific functions when needed (Kelleher and Pausch, 2005; Higgins, 2021; User Onboarding: Principles and Guidelines, Cibin, KS; Megyeri and Szabó, 2024). Interactive walkthroughs embody the principle of learning by doing, whereby active exploration, supported with guidance, is a more effective learning strategy than simply attending to instructions. In practice, future versions of the product developed as part of this thesis could present, in each stage, a short video tutorial explaining the specific tasks and functions would be followed by instructions, provided by the program, to perform the target action (i.e. ‘Press the button to start the trial’, ‘Press the button to repeat the trial’), along with feedback on whether the action was correct or not.

Finally, most information regarding the factors limiting usability, and recommendations on how to improve the design of the user interface came from users with moderate symptoms. Therefore, future studies assessing the usability of improved versions of the prototype should ensure that people from this specific segment of PwA are well represented.

Involving Therapists in Product Design and Usability Testing

A major limitation of the usability studies conducted as part of this dissertation is that they engaged solely PwA, but not therapists. This was not to discount the role of therapists in the process of design and development. Conversely, as discussed at length in chapters 1 and 2, we consider them to be stakeholders and end-users of the application that has been developed as part of this dissertation. Rather, at this stage, the main reason for limiting recruitment to PwA was to quickly identify the main obstacles to usability and to ensure that the program featuring the ASR could be deployed with satisfactory performance, before bringing in therapists to provide more detailed advice on how to improve the product.

CBATs, on their own, are unlikely to address the problems of the current service delivery model. Theories of motivation (Biel et al., 2022; Biel & Haley, 2023; Ryan & Deci, 2017) and studies investigating the factors that promote or discourage adherence to treatment (Harrison et al., 2020) suggest that counselling on the right course of action, provision of education on aphasia, and psychological support are fundamental ingredients for a successful self-management approach (Nichol et al., 2019). Furthermore, studies suggest that the advantage of interacting with a therapist might extend beyond the mere provision of information and assistance throughout the therapeutic journey: a direct comparison between therapy delivered face-to-face versus remotely (via teleconferencing software) has shown that while training-induced benefits were comparable, in-person therapy led to higher increases in a person's confidence in their ability to communicate effectively than those who experienced therapy remotely, even when controlling for the amount of training-induced improvements (Meltzer et al., 2018). This finding is significant, as self-confidence is key to take the initiative to engage in social interactions, which in turn promote a reintegration into one's community — one of the ultimate goals that PwA aspire to achieve at the end of their rehabilitation journey (Berg et al., 2017; Pettit et al., 2017; Rohde et al., 2012; Wallace et al., 2017; Worrall et al., 2011).

Therefore, while attention to human factors related to user interface design and video game-based tasks are key to promote motivation, they can hardly completely replace the role of the therapist. Conversely, in line with the tenets of the patient self-management approach, CBATs can deliver their full potential only when integrated in a support system that includes therapists and carers. For this reason, work conducted on future versions of the prototype should also engage with speech therapists, first capturing their specific user needs (i.e., what features should be present in their UI, what data to collect, etc.), then directly involving them in the iterative process of usability assessment and product refinement.

Assessing Motivation and Adherence to Recommended Therapy Regime

One factor that should be front and center when designing a CBAT is user motivation, as it is key to supporting practice at the level of intensity and dosage that is required to optimize recovery. Therefore, not only is it important to assess whether PwA can benefit from training when using CBATs, it is also important to assess that, when put in full charge of deciding when to self-administer therapy, people actually abide by the frequent, high intensity and prolonged training regimes. While the importance of user motivation is widely recognized within the community of those developing CBATs, little has been done to assess how the design of these applications affects motivation, and ultimately adherence to treatment. In fact, as of now most studies investigating the usability or effectiveness of CBATs were conducted in relatively constrained contexts that do not resemble the conditions under which PwA would be employing a commercially available product to self-administer therapy. For one, these studies typically focus on assessing the efficacy of CBATs targeting a specific aspect of language (Lavoie et al., 2017; Zheng et al., 2016), which requires lower doses than therapy targeting multiple language functions that are generally affected by aphasia. Hence, when it comes to most CBATs that have been tested, little can be said about the extent to which PwA would adhere to longer, and/or more intensive training regimes. Moreover, in the context of research studies (like those that are part this thesis) members of the research team, or therapists, play an active role in ensuring that participants complete all the scheduled study sessions, thus masking the attrition that might have occurred had participants been left in charge of deciding for themselves when to engage with therapy.

At present, the few studies investigating the extent to which users adhere to recommended levels of intensity and dosage when engaging with commercially available apps have revealed high levels of variability (Cordella et al., 2022; Harrison et al., 2020; Liu et al., 2023). Importantly, only a minority of users adhered to recommended training regimes (Harrison et al., 2020; Palmer et al., 2019), a finding that might cast doubts over the real potential of CBATs to serve as an effective means to fill the accessibility gap. Nevertheless, one study went a step further and analyzed which factors distinguished high versus low adherers. Such work highlights the fact that it could be possible to induce desirable changes in behavior in those users that tend to use CBATs less than recommended, with proper innovations in place (Harrison et al., 2020). For example, among the strongest factors associated with low adherence to treatment there were a lack of knowledge about the nature of aphasia, of the critical role of therapy to drive recovery, and a low confidence in one's own capacity to improve via training — factors that typically reinforce one another. Unsurprisingly, as discussed in chapter 2 of this thesis, these factors were listed among the most important unmet needs of PwA and their carers. As proposed in the same chapter, CBATs could serve not only as a means to deliver evidence-based therapy, but also as a tool that can support therapists with providing education on aphasia (i.e. via video tutorials), or serve as a forum wherein people who are new to aphasia can connect with others who have already travelled down the path and can provide valuable information, moral support, and hope.

A further reason to be optimistic about CBATs is that recently the field of aphasia rehabilitation has experienced a surge in interest in understanding how to promote motivation in PwA engaged in self-management. This brings together concepts related to user-centered design, and insights from motivation theory to make therapeutic programs more enticing to PwA (Biel et al., 2022; Biel & Haley, 2023; Nichol et al., 2019, 2022, 2023). Therefore, while currently there is scant evidence suggesting that CBATs would be engaging enough to prevent attrition when used for extensive periods of time, there are many reasons to be hopeful about their future.

Furthermore, in order to ensure that tasks present an optimal level of challenge to each user as they progress through therapy, the training material should be organized according to how difficult it is to engage with (i.e. from shorter to longer words; from higher to lower frequency words; from concrete to abstract words; etc.), and algorithms that can monitor user changing skill level and adapt task difficulty accordingly should be implemented.

Once these exercises will be embedded into video games, we will be in a position to assess how well will the resulting CBAT fare in promoting the motivation required to self-administer therapy recommended by therapists.

Feasibility

As discussed in chapter 4, recruiting participants was very challenging. As such, some of the key objectives of that study (and of the thesis, that is, elucidating the relationship between self-monitoring and the extent to which participants benefit from external feedback) could not be attained. Although I forged relationships with three communities of PwA, which helped with recruitment by pre-screening their members and putting us in touch with those who met some of the eligibility criteria for the study, the number of people who completed all stages of the study was limited. While we were originally provided with contacts for 24 PwA to discuss the study, only 12 participants started the study, and only two were eligible to complete it.

The fact that such a small fraction of the prospective participants was eligible to complete the study might understandably raise some questions regarding the feasibility of continuing the development of a CBAT integrating an ASR, and about the real utility of such an effort if so few people could benefit. For one, at this stage, an integral part of product development is not only the continuous improvement of its usability, but also a thorough assessment of its efficacy (i.e., the immediate and long-term benefits that users experience as a result of using it). This requires testing the product with participants that show the potential to benefit from training and, based on the inclusion criteria that we established for the study described in chapter 4, only two participants satisfied our criteria. However, while doubts about feasibility and utility are legitimate, it is important to point out that the main reason for the low number of participants completing the study was not due to attrition, a lack of interest from prospective participants, or the fact that the product does not address any relevant need of PwA. Instead, there are reasons to think that the inclusion criteria that we set were excessively strict (i.e., selecting only items that were never named correctly over only two attempts). Had the criteria been relaxed enough to not compromise on study quality (such as including items that were named correctly once over multiple baseline sessions), more people would have been eligible to complete the study.

Connected to this point, the pool of PWA which could stand to benefit from using a CBAT integrating an ASR could be considerably expanded by targeting outcome measures in addition to naming accuracy, such as naming RTs. In fact, RTs — which reflect how fast a word's memory representation can be accessed and retrieved — provide unique information on the integrity of such representation, over and above that provided by naming accuracy (Conroy et al., 2018; de Grosbois et al., 2023). All else controlled (e.g. other factors that are known to influence RTs, such as the level of fatigue and alertness; Ingles et al., 1999), the longer the RT measured during naming tasks, the harder we can infer that word retrieval was (Levelt et al., 1999; Oldfield and Wingfield, 1965). As the strength of memory representations might change through time, for example as a result of memory decay due to lack of practice, observing a between-sessions trend towards longer RTs for a given word might signal the need to practice that specific item more often. Had our study been designed to target not only naming accuracy, but also speed, it would have probably satisfied the will of all the prospective participants who expressed an interest in taking part in the study due to their desire to improve upon their naming skills, but that could not ultimately do so. Based on these considerations, future studies should include both naming accuracy and reaction times as outcome measures.

Relaxing the criteria on stimuli selection and introducing RTs as an additional outcome measure would also reduce the number of stimuli that would need to be tested during the baseline phases, most likely to a fraction of the 300 that were employed in the study described in chapter 4. This would help address another challenge that we encountered during study design — that of finding a trade-off between the number and length of baseline sessions dedicated to assessing baseline naming performance to select training material. As discussed in chapter 4, to account for the fact that PwA might be particularly susceptible to fatigue and experience distress as a consequence, we opted to spread the two presentations of each of the 300 items over five study sessions. Since the study had already 17 sessions, and since it was expected to take approximately 2-3 months from start to finish, we decided to not add extra sessions. In this regard, the time saved by reducing the number of stimuli tested could be used to employ additional baseline sessions (without further increasing study length), which would allow for a better estimate of pre-training performance and for the adoption of adequate statistical techniques to analyze the data (Howard et al., 2015).

A further alternative to testing vast amounts of items, as was done in the last study presented in this thesis, would be to ask participants to indicate the categories of items that they find most relevant to them, or even to identify the individual words they wish to practice. This would drastically shorten the amount of trials required for each baseline session, while allowing for additional repetitions of each session.

The Efficacy of the Computer-Based Therapy

Another objective of this dissertation was to conduct a preliminary assessment of the extent to which users could benefit from self-administering therapy. Results were promising, with PwA showing improvements in their ability to name words that were not capable of naming before the beginning of training. Furthermore, while the extent to which the two participants who completed the study benefited from using the CBATs developed as part of this thesis differed (e.g., immediate gains observed only in one case; different magnitude of gains observed relative to the control condition), the improvements in naming ability attributable to training were clear (as measured 1 week after the end of training), and persisted 1 month after the end of training.

Importantly, the gains in naming performance documented in this thesis are similar to those that were reported in previous studies exploring the effectiveness of similar products that did (Ballard et al., 2019) and did not (Lavoie et al., 2017) integrate ASR. However, results concerning the product described here must be corroborated with a larger pool of intended users before drawing firm conclusions regarding its effectiveness.

The Role of ASR-Based Feedback in Driving Treatment-Induced Gains

The project presented in this thesis is one of the first to adopt an ASR to deliver feedback on naming accuracy. The choice to introduce this technology was based on both theoretical grounds and clinical recommendations: on the one hand, there is strong evidence for the key role of feedback in driving gains in naming ability (Sze et al, 2020); on the other hand, both therapists and PwA endorse and encourage the adoption and integration of ASR-based feedback in CBATs, as documented in the literature (Chapey, 2008; Finch & Hill, 2014) and confirmed by the findings of the study presented in chapter 2 of this thesis. However, as discussed in the introductory chapter of this thesis, the bulk of the evidence on the role of feedback comes from studies of face-to-face therapy, wherein therapists deliver feedback when they perceive that their clients are not aware of their mistakes, or that they need encouragement. This delivery differs substantially from the way feedback is delivered by CBATs integrating ASRs, wherein every naming attempt is followed by feedback on performance. This goes counter to not only the theoretical and clinical practices discussed above, but also to key tenets of motivation theory, which recommend limiting the provision of feedback to instances in which subjects demonstrate lack of awareness of their mistakes (Ryan & Deci, 2017; Biel et al., 2022; Biel & Haley, 2023). Such recommendations are grounded on the idea that, if feedback is provided even when a person is capable of evaluating their performance, it might be perceived as imposed upon them, which might either provoke negative emotions or it might favor a mindset of passivity.

However, it is not at all guaranteed that feedback is processed the same way by PwA in the two contexts discussed here. The discomfort that clients might experience when a therapist labels their responses as right or wrong, irrespective of whether this information is needed or not, might or might not occur when on the other side there is an inanimate object providing feedback. In order to help elucidate these aspects of human-computer interaction — which have important implication for usability, and therefore product utility and efficacy — this thesis set out to investigate the conditions in which feedback is most necessary, and the extent to which PwA can benefit from it, by looking at the relationship between treatment-induced gains with and without feedback, subjects' self-monitoring abilities, and the extent to which they thought feedback was useful to them or not.

In this regard, the findings presented in chapter 4 suggest that, in fact, training can lead to comparable gains with and without the delivery of feedback. Within the limited context of this study, this finding could be explained by the fact that the self-monitoring abilities of the two participants who completed the efficacy study were comparable to the ability of the ASR to detect correct and erroneous responses. Therefore, the feedback from ASR may not have provided any benefit to these users, above and beyond their own ability to judge the accuracy of their production. However, it is also important to mention that both participants expressed positive opinions about the accuracy and usefulness of the ASR, which implies that they trusted its outcome. As discussed in chapter 4, this was not the case for all users: in the usability study, some participants expressed negative opinions about the accuracy of the ASR, even though it turned out to outperform their self-monitoring abilities. Unfortunately, these participants were not eligible to partake in the study sessions where the benefits of feedback were investigated, thus nothing can be said regarding whether external feedback would have yielded any benefits to these subjects. This finding underscores the importance of considering individuals' perception on the accuracy and usefulness of the ASR when investigating what factors affect the extent to which people benefit from automatic feedback.

Aside from shedding light on the issues discussed above, future studies should also draw from motivation theory for insights into approaches that deserve to be explored. One such approach is to grant users choice over when to access feedback, as opposed to relying on constant, immediate feedback to determine whether to practice a given item more (Ryan & Deci, 2017), and encouraging them to constantly exercise their judgement when deciding whether they need to correct their naming or not, after a given trial, and to decide when they need feedback. In such a case, feedback on item-level performance would be provided (in a manner akin to how it was presented at Stage 5 of the prototype described in chapter 3) at the end of a study session, or at regular interval within a session, so that participants can see which items they named incorrectly, and they can try again next time. In addition to it, given that the ASR-based accuracy scores for each naming attempt will be stored in the database regardless of whether or not users will decide to view them immediately after naming, information on item-level performance through time can be used to determine how frequently to present a given item for naming — i.e., presenting items that users struggle the most with more frequently than others.

In short, as it is the case with many aspects of aphasia therapy, when and how to deliver automatic feedback might be a parameter that needs to be calibrated on an individual basis.

Determining whether this is the case requires a careful investigation of the effect of multiple factors, such as people's skill level (i.e. error detection), their subjective sense of how trustworthy the ASR is, and how much they value the possibility of choosing when they get feedback.

A Vision for Computer-Based Therapies

As discussed in the introductory chapter of this thesis, CBATs are poised to play a crucial role in the transition to a service delivery model that is more accessible, effective, and financially sustainable, empowering PwA to engage with the necessary amount of therapy from the comfort of their home, and therapists to serve more clients. Moreover, the large-scale deployment of CBATs might finally enable, or at least contribute to, the acquisition of data that is required to understand how to effectively personalize aphasia therapy.

Given the preliminary nature of the studies described in this dissertation, the focus was on first reaching a stage of development wherein PwA could use the application integrating an ASR independently from their home. In line with the principles to consider when designing CBATs that were synthesized in chapter 2, next steps of development should integrate the possibility for users to select the material that is most relevant to them, and expand the range of language processes that are targeted by adding more evidence-based tasks and including real-life scenarios that are relevant to PwA (i.e. asking for information, ordering at the restaurant, doing the groceries etc.).

Examples of additional evidence-based tasks targeting processes for which feedback on production accuracy could prove to be beneficial are semantic feature analysis (SFA) and constraint-induced aphasia therapy (CIAT). As discussed in the introductory chapter of the thesis, SFA was conceived to improve lexical-semantic processing, and it entails participants providing a description of specific features of the target item. This could be done by selecting the right option among a set of competing ones, or by providing a vocal response. While the former response modality can be easily implemented on CBATs that do not integrate an ASR, as users can communicate their choice via touch or by moving the mousepad, the latter would require voice recognition to determine the correctness of the response, and to provide feedback on accuracy. CIAT is based on the concept of preventing learned nonuse, a phenomenon that is observed frequently in patients with motor or language disorders which, when executing a task that would normally involve the use of the impaired limb/language function, would tend to compensate by employing alternative strategies (Pulvermüller et al., 2001; Berthier and Pulvermüller, 2011). For instance, people with impaired speech production would frequently resort to gesturing when having to make requests directed at objects/people located within their visual field. While this strategy can be useful insofar as it allows people to achieve their goals independently, it is not conducive to recovery of lost function. In this regard, constrained induced therapies are based on ensuring that patients are forced to make use of the impaired functionality, in order to promote its recovery. In the context of language therapy, CIAT commonly entails patients engaging in group activities wherein each of them handles a set of cards depicting objects, and has to request a card that only others have. As participants cannot see which cards another person has, they cannot perform the request by using gestures (which would otherwise come spontaneously to them) but they are forced to make a verbal request. Given that this task has proven to be effective at improving a range of language abilities, like naming, repetition, comprehension, reading and communication effectiveness (Berthier and Pulvermüller, 2011; Wang et al., 2020; Pulvermüller et al., 2001; Zhang et al., 2017), it has so far found a large application in group therapy contexts. However, due to its nature, it can be easily implemented on an ASR-based CBAT, allowing users to interact with its virtual companions via the ASR interface.

Aside from these possible outcomes, recent technological improvements are, for the first time, expanding the range of tasks that can be offered by CBATs to domains of communication that seemed far-fetched only a few years ago. So far, CBATs have provided a means to practice several relatively simple therapeutic exercises (like the one employed in this dissertation) targeting reading, writing, speaking and understanding, due to the ease with which they could be implemented in commercially available CBATs. Up until now, more complex tasks resembling real-life scenarios, such as having conversations, have stayed firmly outside of the realm of possibilities for CBATs. The relatively recent deployment of large language models (models that are capable of encoding the most likely sequence of words in response to a specific prompt, based on all the data that is publicly available on the internet), has opened the doors to conversational agents that can entertain open-ended conversations on a vast multitude of topics, sporting a level of fluency and coherence that, unless probed with special prompts, could easily fool people into believing that they are having a conversation with a real human. Another technology that has recently matured, currently best known for driving the (in)famous 'deepfakes' (whereby an algorithm can easily clone a person's voice and face and generate a realistic video of the person saying anything that is requested), can finally be used to generate virtual therapists, a concept that has long been experimented with but with relatively limited application due to technological limitations (Cherney & Van Vuuren, 2012). Combining these two technologies, one can envision a scenario in which, granted permission from a personal therapist, PwA can self-administer therapy under the guidance of a very realistic artificial therapist.

Even if or when such a technology can be developed, the main challenge to overcome will be to know how to interact specifically with PwA, especially in the context of a therapy session. For example, which words and tone to use to communicate that speech is not intelligible, or when a sentence makes little sense in the context of a conversation; how and when to deliver corrective feedback; and how and when to encourage a person. In this regard, the key to bring these applications about could be to expose them to vast amounts of scripts, or videos, of interactions between therapist and clients, and to use theories of motivation to inform how to best encourage and motivate clients in the pursuit of their objectives.

That being said, as repeatedly stressed throughout this dissertation, while CBATs are expected to become far more comprehensive in their offer, and sophisticated in the way they interact with PwA, these applications can really exert their full potential only when working in tandem with therapists and carers. Decades of research on the management of aphasia have unambiguously shown that what really motivates PwA is not only the prospect to return to a situation in which they can meaningfully contribute to their family, their community, and to society writ large, but also having the constant support of a person throughout their journey through rehabilitation. Thus, no matter how capable CBATs will be to emulate human interaction, I believe that it will be unlikely that they can replace the "human touch" and care that can make a huge difference to a person’s outlook on their condition and prospects for their future.

CONCLUSIONS

The work presented in this dissertation was dedicated to the development of a computerized therapy for aphasia that aims to innovate upon two main limitations of currently available applications, integrating an ASR to deliver feedback on naming accuracy in real-time and employing game-based tasks to make therapy more engaging and maximize adherence to treatment. As a first step towards these goals, this thesis started with a detailed analysis of the main unmet needs of PwA that can be addressed by technology (study 1), which subsequently inspired the development of a first prototype of CBAT integrating an ASR and its preliminary assessment of usability with a representative group of PwA (study 2); this prototype was then refined based on feedback from users gathered during study 2, and was used to conduct a first assessment of the feasibility of employing it to allow PwA to self-administer therapy from home.

By and large the studies conducted as part of this thesis showed promising results, as users reported positive opinions about the usability of the prototype CBAT, and two of them could easily self-administer therapy from home, showing improvements in their ability to name a selected set of words. Therefore, while still in the early stages of development and far from a version that can be deployed on a large scale, the current version of the prototype can be used as a basis for expanding the functionalities of the product to offer more evidence-based tasks, the opportunity to personalize the training material based on the needs and preferences of each user, and integrate video game-based activities to make long and intensive training more engaging.

Hopefully, the work presented in this dissertation will inform the development of next generation applications that will contribute to the current trend of moving towards a service delivery model that is more adequate to meet the needs of individuals with aphasia while being economically sustainable.

REFERENCES

Ballard, K., J, Etter, N., M., Shen, S., Monroe, P., & Tien Tan, C. (2019). Feasibility of Automatic Speech Recognition for Providing Feedback During Tablet-Based Treatment for Apraxia of Speech Plus Aphasia. *American Journal of Speech-Language Pathology*, *28*, 818–834.

Berg, K., Askim, T., Balandin, S., Armstrong, E., & By Rise, M. (2017). Experiences of participation in goal setting for people with stroke-induced aphasia in Norway. *Disability & Rehabilitation*, *39*(11), 1122–1130. https://doi.org/10.1080/09638288.2016.1185167

Berthier, M., Pulvermüller, F. Neuroscience insights improve neurorehabilitation of poststroke aphasia.Nat Rev Neurol 7, 86–97 (2011). https://doi.org/10.1038/nrneurol.2010.201

Biel, M., Enclade, H., Richardson, A., Guerrero, A., & Patterson, J. (2022). Motivation theory and practice in aphasia rehabilitation: A scoping review. *American Journal of Speech-Language Pathology*, *31*(5S), 2421–2443.

Biel, M., & Haley, K. L. (2023). Motivation in Aphasia Treatment: Self-Determination Theory Applied to the FOURC Model. *American Journal of Speech-Language Pathology*, *32*(3), 1016–1036.

Chapey, R. (2008). *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders* (5th ed.). Lippincott Williams and Wilkins.

Cherney, L. R., & Van Vuuren, S. (2012). Telerehabilitation, virtual therapists, and acquired neurologic speech and language disorders. *Seminars in Speech and Language*, *33*(03), 243–258.

Cordella, C., Munsell, M., Godlove, J., Anantha, V., Advani, M., & Kiran, S. (2022). Dosage frequency effects on treatment outcomes following self-managed digital therapy: Retrospective cohort study. *Journal of Medical Internet Research*, *24*(7), e36135.

Finch, E., & Hill, A. J. (2014). Computer use by People with Aphasia: A Survey Investigation. *Brain Impairment*, *15*(2), 107–119. https://doi.org/doi.org/10.1017/BrImp.2014.17

Harrison, M., Palmer, R., & Cooper, C. (2020). Factors Associated With Adherence to Self-Managed Aphasia Therapy Practice on a Computer—A Mixed Methods Study Alongside a Randomized Controlled Trial. *Frontiers in Neurology*, *11*.

Higgins, K. (2021). Better onboarding. A Book Apart.

Howard, D., Best, W., & Nickels, L. (2015). Optimizing the design of intervention studies: Critiques and ways forward. *Aphasiology*, *29*(5), 526–562. https://doi.org/10.1080/02687038.2014.985884

Ingles, J. L., Eskes, G. A., & Phillips, S. J. (1999). Fatigue after stroke. *Archives of physical medicine and rehabilitation*, *80*(2), 173-178.

Kelleher, C., & Pausch, R. (2005, April). Stencils-based tutorials: design and evaluation. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 541-550).

Lavoie, M., Macoir, J., & Bier, N. (2017). Effectiveness of technologies in the treatment of post-stroke anomia: A systematic review. *Journal of Communication Disorders*, *65*, 43–53.

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

Liu, H., Cordella, C., Ishwar, P., Betke, M., & Kiran, S. (2023). Consistent long-term practice leads to consistent improvement: Benefits of self-managed therapy for language and cognitive deficits using a digital therapeutic. *Frontiers in Digital Health*, *5*, 1095110.

Megyeri, M., & Szabó, B. (2024). Investigating the effectiveness of user onboarding solutions with eye tracking: a case study on paint 3D. Ergonomics in Design, 32(1), 48-55.

Meltzer, J. A., Baird, A. J., Steele, R. D., & Harvey, S. J. (2018). Computer-based treatment of poststroke language disorders: a non-inferiority study of telerehabilitation compared to in-person service delivery. Aphasiology, 32(3), 290-311.

Nichol, L., Hill, A. J., Wallace, S. J., Pitt, R., Baker, C., & Rodriguez, A. D. (2019). Self-management of aphasia: A scoping review. *Aphasiology*, *33*(8), 903–942.

Nichol, L., Rodriguez, A. D., Pitt, R., Wallace, S. J., & Hill, A. J. (2023). "Self-management has to be the way of the future": Exploring the perspectives of speech-language pathologists who work with people with aphasia. *International Journal of Speech-Language Pathology*, *25*(2), 327–341.

Nichol, L., Wallace, S. J., Pitt, R., Rodriguez, A. D., Diong, Z. Z., & Hill, A. J. (2022). People with aphasia share their views on self-management and the role of technology to support self-management of aphasia. *Disability and Rehabilitation*, *44*(24), 7399–7412.

Oldfield, R. C. & Wingfield, A. (1965) Response latencies in naming objects. Quarterly Journal of Experimental Psychology 17:273–81.

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M. C., Bowen, A., Latimer, N., Julious, S., Cross, E., Alshreef, A., Harrison, M., Bradley, E., Witts, H., & Chater, T. (2019). Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): A multicentre, single-blinded, randomised controlled trial. *Lancet Neurology*, *18*, 821–833.

Pettit, L. K., Tonsing, K. M., & Shakila, D. (2017). The perspectives of adults with aphasia and their team members regarding the importance of nine life areas for rehabilitation: A pilot investigation. *Topics in Stroke Rehabilitation*, *24*(2), 99–106. https://doi.org/10.1080/10749357.2016.1207148

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. Stroke, 32(7), 1621-1626

Rohde, A., Townley-O’Neill, K., Trendall, K., Worrall, L., & Cornwell, P. (2012). A comparison of client and therapist goals for people with aphasia: A qualitative exploratory study. *Aphasiology*, *26*(10), 1298–1315. https://doi.org/10.1080/02687038.2012.706799

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford publications.

Simic, T., Laird, L., Brisson, N., Moretti, K., Théorêt, J. L., Black, S. E., ... & Rochon, E. (2022). Cognitive training to enhance aphasia therapy (Co-TrEAT): A feasibility study. Frontiers in Rehabilitation Sciences, 3, 815780.

Sze, W. P., Hameau, S., Warren, J., & Best, W. (2020). Identifying the Components of a Successful Spoken Naming Therapy: A Meta-Analysis of Word-Finding Interventions for Adults with Aphasia. *Aphasiology*, *35*(1), 33–72.

User Onboarding: Principles and Guidelines (Cibin, KS). Retrieved June 04, 2024, from https://www.uxmatters.com/mt/archives/2022/01/user-onboarding-principles-and-guidelines.php

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

Wang, G., Ge, L., Zheng, Q., Huang, P., & Xiang, J. (2020). Constraint-induced aphasia therapy for patients with aphasia: a systematic review. International journal of nursing sciences, 7(3), 349-358.

Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersh, D., Ferguson, A., & Davidson, B. (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, *25*(3), 309–322. https://doi.org/10.1080/02687038.2010.508530

Zhang, J., Yu, J., Bao, Y., Xie, Q., Xu, Y., Zhang, J., & Wang, P. (2017). Constraint-induced aphasia therapy in post-stroke aphasia rehabilitation: A systematic review and meta-analysis of randomized controlled trials. PloS one, 12(8), e0183349

Zheng, C., Lynch, L., & Taylor, N. (2016). Effect of computer therapy in aphasia: A systematic review. *Aphasiology*, *30*(2–3), 211–244. https://doi.org/10.1080/02687038.2014.996521

BIBLIOGRAPHY

Abad, A., Pompili, A., Costa, A., Trancoso, I., Fonseca, J., Leal, G., Farrajota, L., & Martins, I. P. (2013). Automatic word naming recognition for an on-line aphasia treatment system. *Computer Speech and Language*, *27*, 1235–1248.

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

Allen, M., Mehta, S., McClure, J., & Teasell, R. (2012). Therapeutic interventions for aphasia initiated more than six months post stroke: A review of the evidence. *Topics in Stroke Rehabilitation*, *19*, 523–535.

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., Kong, E., Larraburo, Y., Rolle, C., Johnston, E., & others. (2013). Video game training enhances cognitive control in older adults. *Nature*, *501*(7465), 97–101.

Anguera, J., Brandes-Aitken, A., Rolle, C., Skinner, S., Desai, S., Bower, J., Martucci, W., Chung, W., Sherr, E., & Marco, E. (2016). Characterizing cognitive control abilities in children with 16p11. 2 deletion using adaptive ‘video game’ technology: A pilot study. *Translational Psychiatry*, *6*(9), e893–e893.

Arcury, T. A., Preisser, J. S., Gesler, W. M., & Powers, J. M. (2005). Access to transportation and health care utilization in a rural region. *The Journal of Rural Health*, *21*(1), 31–38.

Avent, J., Glista, S., Wallace, S., Jackson, J., Nishioka, J., & Yip, W. (2005). Family information needs about aphasia. *Aphasiology*, *19*(3–5), 365–374. https://doi.org/10.1080/02687030444000813

Baddeley, A., & Wilson, B. A. (1994). When implicit learning fails: Amnesia and the problem of error elimination. *Neuropsychologia*, *32*(1), 53–68.

Baker, J. C., LeBlanc, L. A., & Raetz, P. B. (2008). A behavioral conceptualization of aphasia. *The Analysis of Verbal Behavior*, *24*, 147–158.

Ballard, K., J, Etter, N., M., Shen, S., Monroe, P., & Tien Tan, C. (2019). Feasibility of Automatic Speech Recognition for Providing Feedback During Tablet-Based Treatment for Apraxia of Speech Plus Aphasia. *American Journal of Speech-Language Pathology*, *28*, 818–834.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459.

Barbera, D. S., Huckvale, M., Fleming, V., Upton, E., Coley-Fisher, H., Doogan, C., Shaw, I., Latham, W., Leff, A. P., & Crinion, J. (2021). NUVA: A naming utterance verifier for aphasia treatment. *Computer Speech and Language*, *69*, 1–25. https://doi.org/10.1016/j.csl.2021.101221

Berg, K., Askim, T., Balandin, S., Armstrong, E., & By Rise, M. (2017). Experiences of participation in goal setting for people with stroke-induced aphasia in Norway. *Disability & Rehabilitation*, *39*(11), 1122–1130. https://doi.org/10.1080/09638288.2016.1185167

Bersano, A., Burgio, F., Gattinoni, M., & Candelise, L. (2009). Aphasia burden to hospitalised acute stroke patients: Need for an early rehabilitation programme. *International Journal of Stroke*, *4*(6), 443–447.

Berthier, M. L., & Pulvermüller, F. (2011). Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nature Reviews Neurology*, *7*(2), 86–97.

Biel, M., Enclade, H., Richardson, A., Guerrero, A., & Patterson, J. (2022). Motivation theory and practice in aphasia rehabilitation: A scoping review. *American Journal of Speech-Language Pathology*, *31*(5S), 2421–2443.

Biel, M., & Haley, K. L. (2023). Motivation in Aphasia Treatment: Self-Determination Theory Applied to the FOURC Model. *American Journal of Speech-Language Pathology*, *32*(3), 1016–1036.

Blank, S., & Dorf, B. (2012). *The Startup Owner's Manual* (First). K&S Ranch Publishing Division.

Bove, R. M., Rush, G., Zhao, C., Rowles, W., Garcha, P., Morrissey, J., Schembri, A., Alailima, T., Langdon, D., Possin, K., & others. (2019). A videogame-based digital therapeutic to improve processing speed in people with multiple sclerosis: A feasibility study. *Neurology and Therapy*, *8*, 135–145.

Boyle, M. (2010). Semantic feature analysis treatment for aphasic word retrieval impairments: What's in a name? *Topics in Stroke Rehabilitation*, *17*(6), 411–422.

Brady, M. C., Ali, M., Berg, K. V., Williams, L. J., Williams, L. R., Abo, M., Becker, F., Bowen, A., Brandenburg, C., Breitenstein, C., & others. (2022). Complex speech-language therapy interventions for stroke-related aphasia: The RELEASE study incorporating a systematic review and individual participant data network meta-analysis. *Health and Social Care Delivery Research*, *10*(28).

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *Cochrane Database of Systematic Reviews*, *6*. https://doi.org/10.1002/14651858.CD000425.pub4.

Braley, M., Sims Pierce, J., Saxena, S., De Oliveira, E., Taraboanta, L., Anantha, V., Lakhan, S. E., & Kiran, S. (2021). A Virtual, Randomized, Control Trial of a Digital Therapeutic for Speech, Language, and Cognitive Intervention in Post-stroke Persons With Aphasia. *Frontiers in Neurology*, *12*(626780). https://doi.org/10.3389/fneur.2021.626780

Brandeburg, C., Worrall, L., Copland, D., & Rodriguez, A. D. (2017). Barriers and facilitators to using the CommFitTM smart phone app to measure talk time for people with aphasia. *Aphasiology*, *31*(8), 901–927. https://doi.org/10.1080/02687038.2016.1219016

Brandeburg, C., Worrall, L., Rodriguez, A. D., & Copland, D. (2013). Mobile computing technology and aphasia: An integrated review of accessibility and potential uses. *Aphasiology*, *27*(4), 444–461. https://doi.org/10.1080/02687038.2013.772293

Brodeur, M. (n.d.). *The BOSS dataset*. https://sites.google.com/site/bosstimuli/

Brodeur, M., Guerard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. *PlosOne*, *9*(9). https://doi.org/10.1371/ journal.pone.0106953

Brooke, J. (1986). *System Usability Scale (SUS)*.

Byiers, B. J., Reichle, J., & Symons, F. J. (2012). Single-Subject Experimental Design for Evidence-Based Practice. *American Journal of Speech-Language Pathology*, *21*(4), 397–414. https://doi.org/10.1044/1058-0360(2012/11-0036)

Cahana-Amitay, D., & Albert, M. (2015). *Redefining Recovery From Aphasia*. Oxford University Press.

Canadian Stroke Network. (2011). *The Quality Of Stroke Care in Canada*. https://www.canadianstroke.ca/sites/default/files/resources/QoSC-EN.pdf

Canadian Stroke Best Practices. (2024). https://www.strokebestpractices.ca/recommendations/cerebral-venous-thrombosis/2-acute-treatment-of-symptomatic-cerebral-venous-thrombosis

Caporali, A., & Basso, A. (2003). A survey of long-term outcome of aphasia and of chances of gainful employment. *Aphasiology*, *17*(9), 815–834.

Carmichael, S. T. (2016). The 3 Rs of stroke biology: Radial, relayed, and regenerative. *Neurotherapeutics*, *13*, 348–359.

Carrera, E., & Tononi, G. (2014). Diaschisis: Past, present, future. *Brain*, *137*(9), 2408–2422.

Caughlin, S., Mehta, S., Corriveau, H., Eng, J. J., Eskes, G., Kairy, D., Meltzer, J., Sakakibara, B. M., & Teasell, R. (2020). Implementing telerehabilitation after stroke: Lessons learned from Canadian trials. *Telemedicine and E-Health*, *26*(6), 710–719.

Chaiken, S., & Maheswaran, D. (1994). Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgment. *Journal of Personality and Social Psychology*, *66*(3), 460–473.

Chang, Y.-N., & Lambon Ralph, M. A. (2020). A unified neurocomputational bilateral model of spoken language production in healthy participants and recovery in poststroke aphasia. *Proceedings of the National Academy of Sciences*, *117*(51), 32779–32790.

Chapey, R. (2008). *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders* (5th ed.). Lippincott Williams and Wilkins.

Chapey, R., Duchan, J., Elman, R., Garcia, L., Kagan, A., Lyon, J., & Simmons-Mackie, N. (2000). *Life Participation Approaches to Aphasia: A statement of values for the future.* (5; pp. 4–6). ASHA Leader.

Chen, C. C., & Bode, R. K. (2011). Factors Influencing Therapists' Decision-Making in the Acceptance of New Technology Devices in Stroke Rehabilitation. *American Journal of Physical Medicine & Rehabilitation*, *90*, 415–425. https://doi.org/10.1097/PHM.0b013e318214f5d8

Cheng, X., Schafer, G., & Akyürek, E. G. (2010). Name agreement in picture naming: An ERP study. *Official Journal of the Internation Organization of Psychophysiology*, *76*(3), 130–141.

Cherney, L. R., & Van Vuuren, S. (2012). Telerehabilitation, virtual therapists, and acquired neurologic speech and language disorders. *Seminars in Speech and Language*, *33*(03), 243–258.

Choi, Y.-H., Park, H. K., & Paik, N.-J. (2016). A telerehabilitation approach for chronic aphasia following stroke. *Telemedicine and E-Health*, *22*(5), 434–440.

Christensen, J. M., & Anderson, J. D. (1989). Spouse adjustment to stroke: Aphasic versus nonaphasic partners. *Journal of Communication Disorders*, *22*(4), 225–231.

Code, C., & Heron, C. (2003). Services for aphasia, other acquired adult neurogenic communication and swallowing disorders in the United Kingdom, 2000. *Disability & Rehabilitation*, *25*(21), 1231–1237. https://doi.org/10.1080/0963828031000159996l

Code, Chris, & Petheram, B. (2011). Delivering for aphasia. *International Journal of Speech-Language Pathology*, *13*(1), 3–10. https://doi.org/10.3109/17549507.2010.520090

Conroy, P., Sotiropoulou Drosopoulou, C., Humphreys, G. F., Halai, A. D., & Lambon Ralph, M. A. (2018). Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia. *Brain*, *141*(6), 1815–1827.

Constant Therapy. (2022). [Computer software]. https://constanttherapyhealth.com/science/

Copland, D. (2003). The basal ganglia and semantic engagement: Potential insights from semantic priming in individuals with subcortical vascular lesions, Parkinson's disease, and cortical lesions. *Journal of the International Neuropsychological Society*, *9*(7), 1041–1052.

Cordella, C., Munsell, M., Godlove, J., Anantha, V., Advani, M., & Kiran, S. (2022). Dosage frequency effects on treatment outcomes following self-managed digital therapy: Retrospective cohort study. *Journal of Medical Internet Research*, *24*(7), e36135.

Cota, T. T., & Ishitani, L. (2015). Motivation and benefits of digital games for the elderly: A systematic literature review. *Revista Brasileira de Computacao Aplicado*, *7*(1), 2–16. https://doi.org/10.5335/rbca.2015.4190

Cramer, S. C. (2008). Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. *Annals of Neurology*, *63*(3), 272–287.

Criptographic hash function. (N.d.). (n.d.). https://en.wikipedia.org/wiki/Cryptographic_hash_function

Crosson, B., Benefield, H., Cato, M. A., Sadek, J. R., Moore, A. B., Wierenga, C. E., Gopinath, K., Soltysik, D., Bauer, R. M., Auerbach, E. J., & others. (2003). Left and right basal ganglia and frontal activity during language generation: Contributions to lexical, semantic, and phonological processes. *Journal of the International Neuropsychological Society*, *9*(7), 1061–1077.

Cuetos, F., Aguado, G., Izura, C., & Ellis, A. W. (2002). Aphasic naming in Spanish: Predictors and errors. *Brain and Language*, *82*(3), 344–365.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 318.

Dallery, J., Cassidy, R. N., & Raiff, B. R. (2013). Single-Case Experimental Designs to Evaluate Novel Technology-Based Health Interventions. *Journal of Medical Internet Research*, *15*(2), 1–17. https://doi.org/10.2196/jmir.2227

Damasio, A. (1998). Signs of aphasia. In *Acquired aphasia* (M.T.Sarno, pp. 25–41). Academic Press.

Davis, G. A. (2013). *Aphasia and related cognitive-communicative disorders*. Pearson Higher Ed.

de Grosbois, J., Canthiya, L., Philipp-Muller, A. E., Hickey, N. K., Hodzic-Santor, B., Heleno, M. C., Jokel, R., & Meltzer, J. A. (2023). Asynchronous, online spaced-repetition training alleviates word-finding difficulties in aphasia. *Neuropsychological Rehabilitation*, *33*(10), 1672–1696.

De Schutter, B., & Malliet, S. (2014). The older player of digital games: A classification based on perceived need satisfaction. *Communications*, *39*(1), 67–88. https://doi.org/10.1515/commun-2014-0005

de Silva, D. (2011). *Evidence: Helping people help themselves*.

Debray, T. P., Schuit, E., Efthimiou, O., Reitsma, J. B., Ioannidis, J. P., Salanti, G., Moons, K. G., & Workpackage, G. (2018). An overview of methods for network meta-analysis using individual participant data: When do benefits arise? *Statistical Methods in Medical Research*, *27*(5), 1351–1364.

DeLeon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., Lee, A., & Hillis, A. E. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain*, *130*(5), 1408–1422.

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*, *93*(3), 282–321.

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, *21*(2–4), 125–145.

Dell, G. S., Nozari, N., & Oppenheim, G. M. (2014). Word production: Behavioral and computational considerations. *The Oxford Handbook of Language Production*, 88–104.

Dell, G. S., & O'Seaghdha, P. (1992). Stages of lexical access in language production. *Cognition*, *42*, 287–314.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review*, *104*(4), 801–838.

Dell, G. S., Schwartz, M. F., Nozari, N., Faseyitan, O., & Coslett, H. B. (2013). Voxel-based lesion-parameter mapping: Identifying the neural correlates of a computational model of word production. *Cognition*, *128*(3), 380–396.

DesRoches, C. A., Balachandran, I., Ascenso, E. M., Tripodis, Y., & Kiran, S. (2015). Effectiveness of an impairment-based individualized rehabilitation program using an iPad-based software platform. *Frontiers In Human Neuroscience*, *8*(1015). https://doi.org/doi: 10.3389/fnhum.2014.01015

DesRoches, C. A., & Kiran, S. (2017). Technology-Based Rehabilitation to Improve Communication after Acquired Brain Injury. *Frontiers In Human Neuroscience*, *11*(382). https://doi.org/10.3389/fnins.2017.00382

Diaz-Orueta, U., Facal, D., Nap, H. H., & Ranga, M.-M. (2012). What Is the Key for Older People to Show Interest in Playing Digital Learning Games? Initial Qualitative Findings from the LEAGE Project on a Multicultural European Sample. *Games for Health Journal: Research, Develoment, and Clinical Applications*, *1*(2). https://doi.org/10.1089/g4h.2011.0024

Dickey, L., Kagan, A., Lindsay, M. P., Fang, J., Rowland, A., & Black, S. (2010). Incidence and profile of inpatient stroke-induced aphasia in Ontario, Canada. *Archives of Physical Medicine and Rehabilitation*, *91*(2), 196–202.

DigitalOcean. (N.d.). (n.d.). https://www.digitalocean.com/

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General, 121*(3), 275–277. https://doi.org/10.1037/0096-3445.121.3.275

Doesborgh, S. J. C., van de Sandt-Koenderman, M. W. M. E., Dippel, D. W. J., van Harskamp, F., Koudstaal, P. J., & Visch-Brink, E. G. (2004). Cues on request: The efficacy of Multicue, a computer program for wordfinding therapy. *Aphasiology*, *18*(3), 213–222. https://doi.org/10.1080/02687030344000580

Doesborgh, S. J., van de Sandt-Koenderman, M. W., Dippel, D. W., van Harskamp, F., Koudstaal, P. J., & Visch-Brink, E. G. (2004). Effects of semantic treatment on verbal communication and linguistic processing in aphasia after stroke: A randomized controlled trial. *Stroke*, *35*(1), 141–146.

Doroudian, A. (2019). *Design, Development, and Evaluation of an Online Escape Game for Older Adults*. Simor Fraser.

Dorze, G. L., & Brassard, C. (1995). A description of the consequences of aphasia on aphasic persons and their relatives and friends, based on the WHO model of chronic diseases. *Aphasiology*, *9*(3), 239–255.

Draper, P., & Brocklehurst, H. (2007). The impact of stroke on the well-being of the patient's spouse: An exploratory study. *Journal of Clinical Nursing*, 264-- 271. https://doi.org/10.1111/j.1365-2702.2006.01575.x

Efstratiadou, E. A., Papathanasiou, I., Holland, R., Archonti, A., & Hilari, K. (2018). A systematic review of semantic feature analysis therapy studies for aphasia. *Journal of Speech, Language, and Hearing Research*, *61*(5), 1261–1278.

Ellis, A. W., Miller, D., & Sin, G. (1983). Wernlcke's aphasia and normal languagp processing: A case study in cognitive neuropsychology. *Cognition*, *15*, 111–144.

Ellis, C., Simpson, A. N., Bonilha, H., Mauldin, P. D., & Simpson, K. N. (2012). The one-year attributable cost of poststroke aphasia. *Stroke*, *43*(5), 1429–1431.

Elman, R. (2016). Aphasia Centers and the Life Participation Approach to Aphasia. A Paradigm Shift. *Topics in Language Disorders*, *36*(2), 154–167. https://doi.org/10.1097/TLD.0000000000000087

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

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

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

Finch, E., & Hill, A. J. (2014). Computer use by People with Aphasia: A Survey Investigation. *Brain Impairment*, *15*(2), 107–119. https://doi.org/doi.org/10.1017/BrImp.2014.17

Fink, R. B., Brecher, A., Schwartz, M. F., & Robey, R. R. (2002). A computer-implemented protocol for treatment of naming disorders: Evaluation of clinicianguided and partially self-guided instruction. *Aphasiology*, *16*, 1061–1086.

Flowers, H., Skoretz, A., Silver, F., Rochon, E., Fang, J., Flamand-Roze, C., & Marinto, R. (2016). Poststroke aphasia frequency, recovery and outcomes: A systematic review and meta-analysis. *Archives of Physical Medicine and Rehabilitation*, *97*(12), 2188---2201.

Flynn, R. M., Colón-Acosta, N., Zhou, J., & Bower, J. (2019). A game-based repeated assessment for cognitive monitoring: Initial usability and adherence study in a summer camp setting. *Journal of Autism and Developmental Disorders*, *49*(5), 2003–2014.

Foley, N., McClure, J. A., Meyer, M., Salter, K., Bureau, Y., & Teasell, R. (2012). Inpatient rehabilitation following stroke: Amount of therapy received and associations with functional recovery. *Disability & Rehabilitation*, *34*(25), 2132–2138. https://doi.org/10.3109/09638288.2012.676145

Fonseca, J., Raposo, A., & Pavao Martins, I. (2018). Cognitive functioning in chronic poststroke aphasia. *Applied Neuropsychology: Adult*. https://doi.org/10.1080/23279095.2018.1429442

Foster, A., O'Halloran, R., Rose, M., & Worrall, L. (2016). "Communication is taking a back seat": Speech pathologists' perceptions of aphasia management in acute hospital settings. *Aphasiology*, *30*(4–6), 585–608.

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language*, *43*(2), 182–216.

Frattali, C. M., Thompson, C. K., Holland, A. L., Wohl, C. B., Wenck, C. J., Slater, S. C., & Paul, D. (2017). *American Speech-Language-Hearing Association Functional Assessment of Communication Skills for Adults (ASHA FACS)*. American Speech-Language-Hearing Association.

Fridriksson, J., den Ouden, D.-B., Hillis, A. E., Hickok, G., Rorden, C., Basilakos, A., Yourganov, G., & Bonilha, L. (2018). Anatomy of aphasia revisited. *Brain*, *141*(3), 848–862.

Fridriksson, J., & Hillis, A. E. (2021). Current approaches to the treatment of post-stroke aphasia. *Journal of Stroke*, *23*(2), 183.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 27–52.

Galliers, J. R., Wilson, S., Muscroft, S., Marshall, J., Roper, A., Cocks, N., & Pring, T. (2011). Accessibility of 3D Game Environments for People with Aphasia: An Exploratory Study. *Paper Presented at the 13th International ACM SIGACCESS Conference on Computers and Accessibility, 24-26 Oct, Dundee, Scotland*.

Gamberini, L., Alcaniz, M., Barresi, G., Fabregat, M., Ibanez, F., & Prontu, L. (2006). Cognition, technology and games for the elderly: An introduction to ELDERGAMES Project. *PsychNology Journal*, *4*(3), 285–308.

Ganzfried, E., & Hinckley, J. (2012). *The persistent needs of people living with aphasia: Results of a national survey*. Clinical Aphasiology Conference, Lake Tahoe, CA. http://aphasiology.pitt.edu/id/eprint/2399

Garrett, M. F. (1975). The analysis of sentence production. In *Psychology of learning and motivation* (Vol. 9, pp. 133–177). Elsevier.

Geschwind, N. (1970). The Organization of Language and the Brain: Language disorders after brain damage help in elucidating the neural basis of verbal behavior. *Science*, *170*(3961), 940–944.

Gleichman, A. J., & Carmichael, S. T. (2014). Astrocytic therapies for neuronal repair in stroke. *Neuroscience Letters*, *565*, 47–52.

Glen, S. (n.d.). Binomial Confidence Interval. *From StatisticsHowTo.Com: Elementary Statistics for the Rest of Us!* Retrieved September 27, 2023, from https://www.statisticshowto.com/binomial-confidence-interval/

Godecke, E., Armstrong, E., Hersh, D., & Bernhardt, J. (2014). Missed opportunities: Communicative interactions in early stroke recovery. *International Journal of Stroke*, *9*(48).

Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*, *21*(6), 649–683.

Goodglass, H., & Wingfield, A. (1997a). Anatomical and theoretical considerations in anomia. In *Anomia: Neuroanatomical and cognitive correlates* (H. Goodglass&A. Wingfield, pp. 20–29). Academic Press.

Goodglass, H., & Wingfield, A. (1997b). *Anomia: Neuroanatomical and Cognitive Correlates*. Academic Press.

Graham, J. E., Karmarkar, A. M., & Ottenbacher, K. J. (2012). Small Sample Research Designs for Evidence-based Rehabilitation: Issues and Methods. *Archives of Physical Medicine and Rehabilitation*, *93*(8), 111–116. https://doi.org/10.1016/j.apmr.2011.12.017

Graham, J. R., Pereira, S., & Teasell, R. (2011). Aphasia and return to work in younger stroke survivors. *Aphasiology*, *25*(8), 952–960.

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*, 105–112. https://doi.org/doi:10.1016/j.nedt.2003.10.001

Graves, A. (2012). Long Short-Term Memory. *Supervised Sequence Labelling with Recurrent Neural Network*, *385*, 37–45.

Grefkes, C., & Fink, G. R. (2020). Recovery from stroke: Current concepts and future perspectives. *Neurological Research and Practice*, *2*(1), 1–10.

Greig, C. A., Harper, R., Hirst, T., Howe, T., & Davidson, B. (2008). Barriers and Facilitators to Mobile Phone Use for People with Aphasia. *Topics in Stroke Rehabilitation*, *15*(4), 307–324. https://doi.org/10.1310/tsr1504-307

Grossman, M. (2014). Anomia. In *Encyclopedia of the Neurological Sciences* (M. J. Aminoff&R. B. Daroff (Eds.), Vol. 1, pp. 194–197). Academic Press.

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

Halai, A. D., Woollams, A. M., & Ralph, M. A. L. (2018). Triangulation of language-cognitive impairments, naming errors and their neural bases post-stroke. *NeuroImage: Clinical*, *17*, 465–473.

Harmon, T. G., Hardy, L., & Haley, K. L. (2018). Proactive social validation of methods and procedures used for training speech production in aphasia. *Aphasiology*, *32*(8), 922–943.

Harrison, M., Palmer, R., & Cooper, C. (2020). Factors associated with adherence to self-managed aphasia therapy practice on a computer—A mixed methods study alongside a randomized controlled trial. *Frontiers in Neurology*, *11*, 582328.

Hartwell, J. (2015). *Access Denied: An Investigation of the Impact of Aphasia on Social Inclusion in Long-Term Care Facilities* [University of Louisiana at Lafayette]. https://www.proquest.com/openview/866791505f5dae3b6c9368461ae49702/1?pq-origsite=gscholar&cbl=18750

Heart & Stroke. (2017). *Stroke report 2017*. Heart and Stroke Foundation of Canada. https://www.heartandstroke.ca/-/media/pdf-files/canada/stroke-report/strokereport2017en.ashx?rev=48bbc6fe37d84c618687e259acc69920

Heart & Stroke. (2020). *Where every dollar goes*. Heart and Stroke. https://www.heartandstroke.ca/what-we-do/finances/where-every-dollar-goes

Hebert, D., Lindsay, M., McIntyre, A., Kirton, A., Rumney, P., … Bagg, S. (2016). Canadian Stroke Best Practice Recommendations: Stroke Rehabilitation Practice Guidelines. *International Journal of Stroke*, *11*(4), 459–484.

Herbert, R., Webster, D., & Dyson, L. (2012). Effects of syntactic cueing therapy on picture naming and connected speech in acquired aphasia. *Neuropsychological Rehabilitation*, *22*(4), 609–633. https://doi.org/10.1080/09602011.2012.679030

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*(2), 135–145.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402.

Higgins, K. (2021). Better onboarding. A Book Apart.

Hilari, K. (2011). The impact of stroke: Are people with aphasia different to those without? *Disability & Rehabilitation*, *33*(3), 211–218.

Hilari, K., Byng, S., Lamping, D. L., & Smith, S. C. (2003). Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39) Evaluation of Acceptability, Reliability, and Validity. *Stroke*, *34*, 1944–1950. https://doi.org/10.1161/01.STR.0000081987.46660.ED

Hilari, K., Northcott, S., Roy, P., Marshall, J., Wiggins, R., Chataway, J., & Ames, D. (2010). Psychological distress after stroke and aphasia: The first six months. *Clinical Rehabilitation*, *24*(2), 181–190.

Hill, A. J., & Breslin, H. M. (2016). Refining an Asynchronous Telerehabilitation Platform for Speech-Language Pathology: Engaging End-Users in the Process. *Frontiers In Human Neuroscience*, *10*(640). https://doi.org/10.3389/fnhum.2016.00640

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780.

Holland, A., Fromm, D., Forbes, M., & MacWhinney, B. (2017). Long-term recovery in stroke accompanied by aphasia: A reconsideration. *Aphasiology*, *31*(2), 152–165.

Howard, D., Best, W., & Nickels, L. (2015). Optimizing the design of intervention studies: Critiques and ways forward. *Aphasiology*, *29*(5), 526–562. https://doi.org/10.1080/02687038.2014.985884

Howard, D., Patterson, K. E., Franklin, S. E., Orchard-Lislie, V. M., & Morton, J. (1985). The treatment of word retrieval deficits in aphasia: A comparison of two therapy methods. *Brain*, *108*, 817–829. https://doi.org/10.1093/brain/108.4.817

Howard, D., Patterson, K., Franklin, S., Orchard-Lisle, V., & Morton, J. (1985). Treatment of word retrieval deficits in aphasia. *Brain*, *108*(8), 17–829.

Howard, G., & Goff, D. C. (2012). Population shifts and the future of stroke: Forecasts of the future burden of stroke. *Annals of the New York Academy of Sciences*, *1268*(1), 14–20.

Howe, T., Davidson, B., Worrall, L., Hersh, D., Ferguson, A., Sherratt, S., & Gilbert, J. (2012). `You needed to rehab. . . Families as well': Family members' own goals for aphasia rehabilitation. *International Journal of Language & Communication Disorders*, *47*(5), 511–521. https://doi.org/10.1111/j.1460-6984.2012.00159.x

Huang, Q., Shu, T., Xu, W., & Chen, Z. (2022). Alerting, Orienting, Executive Control and Language Function in Post-stroke Aphasia Patients. *Rehabilitation Science*, *7*(1), 5–15. https://doi.org/10.11648/j.rs.20220701.12

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*, 255.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144.

Ingles, J. L., Eskes, G. A., & Phillips, S. J. (1999). Fatigue after stroke. *Archives of physical medicine and rehabilitation*, *80*(2), 173-178.

International Standard Organization. (2016). *Application of Usability Engineering to Medical Devices*.

Jamal, N., Shanta, S., Mahmud, F., & Sha'abani, M. (2017). Automatic Speech Recognition (ASR) based Approach for Speech Therapy of Aphasic Patients: A Review. *AIP Conference Proceedings*, *1883*. https://doi.org/10.1063/1.5002046

Jarso, S., Li, M., Faria, A., Davis, C., Leigh, R., Sebastian, R., Tsapkini, K., Mori, S., & Hillis, A. E. (2013). Distinct mechanisms and timing of language recovery after stroke. *Cognitive Neuropsychology*, *30*(7–8), 454–475.

Jones, F., & Riazi, A. (2011). Self-efficacy and self-management after stroke: A systematic review. *Disability and Rehabilitation*, *33*(10), 797–810.

Joy, M. T., & Carmichael, S. T. (2021). Encouraging an excitable brain state: Mechanisms of brain repair in stroke. *Nature Reviews Neuroscience*, *22*(1), 38–53.

Kagan, A., Simmons-Mackie, N., Kapral, N., Hoch, J., Graves, E., Mecredy, G., Whitehurst, D., & Chan, M. (2017, June 3). *Health care utilization and costs in people with stroke with and without aphasia in Ontario, Canada* [Poster presentation at the Clinical Aphasiology Conference].

Kagan, A., Simmons-Mackie, N., Rowland, A., Huijbregts, M., Shumway, E., McEwen, S., Threats, T., & Sharp, S. (2008). Counting what counts: A framework for capturing real-life outcomes of aphasia intervention. *Aphasiology*, *22*(3), 258–280.

Kaufman, D., Ma, M., Sauvè, L., Renaud, L., & Duplàa, E. (2019). Benefits of Digital Gameplay for Older Adults: Does Game Type Make a Difference? *Internation Journal of Aging Research*, *2*(43).

Kauhanen, M., Korpelainen, J., Hiltunen, P., Määttä, R., Mononen, H., Brusin, E., & Myllylä, V. (2000). Aphasia, depression, and non-verbal cognitive impairment in ischaemic stroke. *Cerebrovascular Diseases*, *10*(6), 455–461.

Kaye. (1997). *Disability watch: The status of people with disabilities in the United States.* Disability Statistics Center, Institute for Health and Aging, University of California at San Francisco.

Kelleher, C., & Pausch, R. (2005, April). Stencils-based tutorials: design and evaluation. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 541-550).

Kerr, E., Ivanova, B., & Strijkers, K. (2022). Lexical Access in Speech Production. Psycho- and neurolinguistic perspectives on the spatiotemporal dynamics. In *Current issues in the Psychology of Language.* Routledge Press (Taylor and Francis).

Kertesz, A. (2007). *Western Aphasia Battery: Revised*. https://books.google.com/books/about/Western_Aphasia_Battery.html?hl=&id=bD3DxQEACAAJ

Kim, E. S., Suleman, S., & Hopper, T. (2020). Decision Making by People With Aphasia: A Comparison of Linguistic and Nonlinguistic Measures. *Journal of Speech, Language, and Hearing Research*, *63*(6), 1845–1860. https://doi.org/10.1044/2020_JSLHR-19-00182

Kiran, S., Meier, E. L., & Johnson, J. P. (2019). Neuroplasticity in aphasia: A proposed framework of language recovery. *Journal of Speech, Language, and Hearing Research*, *62*(11), 3973–3985.

Kiran, S., & Thompson, C. K. (2003). *The role of semantic complexity in treatment of naming deficits*.

Kiran, S., & Thompson, C. K. (2019). Neuroplasticity of language networks in aphasia: Advances, updates, and future challenges. *Frontiers in Neurology*, *10*, 295.

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology*, *25*(4), 463–492.

Kleim, J. A. (2011). Neural plasticity and neurorehabilitation: Teaching the new brain old tricks. *Journal of Communication Disorders*, *44*(5), 521–528.

Kleim, J. A., & Jones, T. A. (2008). *Principles of experience-dependent neural plasticity: Implications for rehabilitation after brain damage*.

Krakauer, J. W., & Carmichael, S. T. (2022). *Broken movement: The neurobiology of motor recovery after stroke*. MIT Press.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-Case Intervention Research Design Standards. *Remedial and Special Education*, *34*(1), 26–38. https://doi.org/10.1177/0741932512452794

Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the Scientific Credibility of Single-Case Intervention Research: Randomization to the Rescue. *Psychological Methods*, *15*(2), 124–144. https://doi.org/10.1037/a0017736

Kristinsson, S. (2021). *Towards Understanding Therapy Response in Chronic Aphasia* [PhD Thesis]. University of South Carolina.

Kristinsson, S., Basilakos, A., Elm, J., Spell, L. A., Bonilha, L., Rorden, C., Den Ouden, D. B., Cassarly, C., Sen, S., Hillis, A., & others. (2021). Individualized response to semantic versus phonological aphasia therapies in stroke. *Brain Communications*, *3*(3), fcab174.

Kümmerer, D., Hartwigsen, G., Kellmeyer, P., Glauche, V., Mader, I., Klöppel, S., Suchan, J., Karnath, H.-O., Weiller, C., & Saur, D. (2013). Damage to ventral and dorsal language pathways in acute aphasia. *Brain*, *136*(2), 619–629.

La Guardia, J. G. (2017). *Self-determination theory in practice: How to create an optimally supportive health care environment*. éditeur inconnu.

Laganaro, M., Di Pietro, M., & Schnider, A. (2003). Computerised treatment of anomia in chronic and acute aphasia: An exploratory study. *Aphasiology*, *17*(8), 709–721. https://doi.org/10.1080/02687030344000193

Laine, M., & Martin, N. (2006). *Anomia. Theoretical and Clinical Aspects* (1st ed.). Psychology Press.

Laliberte, M., Gaureau, C., & LeDorze, G. (2016). A pilot study on how speech-language pathologists include social participation in aphasia rehabilitation. *Aphasiology*, *30*(10), 1117–1133.

Lam, J. M., & Wodchis, W. P. (2010). The relationship of 60 disease diagnoses and 15 conditions to preference-based health-related quality of life in Ontario hospital-based long-term care residents. *Medical Care*, 380–387.

Laska, A. C., Hellblom, A., Murray, V., Kahan, T., & Von Arbin, M. (2001). Aphasia in acute stroke and relation to outcome. *Journal of Internal Medicine*, *249*(5), 413–422.

Lavoie, M., Bier, N., & Macoir, J. (2019). Efficacy of a self-administered treatment using a smart tablet to improve functional vocabulary in post-stroke aphasia: A case-series study. *International Journal of Language & Communication Disorders*, *54*(2), 249–264.

Lavoie, M., Macoir, J., & Bier, N. (2017). Effectiveness of technologies in the treatment of post-stroke anomia: A systematic review. *Journal of Communication Disorders*, *65*, 43–53.

Lazar, R. M., Minzer, B., Antoniello, D., Festa, J. R., Krakauer, J. W., & Marshall, R. S. (2010). Improvement in aphasia scores after stroke is well predicted by initial severity. *Stroke*, *41*(7), 1485–1488.

Le Belle, J. E., Orozco, N. M., Paucar, A. A., Saxe, J. P., Mottahedeh, J., Pyle, A. D., Wu, H., & Kornblum, H. I. (2011). Proliferative neural stem cells have high endogenous ROS levels that regulate self-renewal and neurogenesis in a PI3K/Akt-dependant manner. *Cell Stem Cell*, *8*(1), 59–71.

Le, D., Licata, K., & Provost, E. M. (2018). Automatic quantitative analysis of spontaneous aphasic speech. *Speech Communication*, *100*, 1–12. https://doi.org/10.1016/j.specom.2018.04.001

Lee, J. B., Kocherginsky, M., & Cherney, L. R. (2020). Attention in individuals with aphasia: Performance on the Conners' Continuous Performance Test – 2nd edition. *Neuropsychological Rehabilitation*, *30*(2), 249–265. https://doi.org/10.1080/09602011.2018.1460852

Leonard, C., Rochon, E., & Laird, L. (2008). Treating naming impairments in aphasia: Findings from a phonological components analysis treatment. *Aphasiology*, *22*(9), 923–947.

Levelt, W. J. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, *42*(1–3), 1–22.

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

Li, S., Nie, E. H., Yin, Y., Benowitz, L. I., Tung, S., Vinters, H. V., Bahjat, F. R., Stenzel-Poore, M. P., Kawaguchi, R., Coppola, G., & others. (2015). GDF10 is a signal for axonal sprouting and functional recovery after stroke. *Nature Neuroscience*, *18*(12), 1737–1745.

Linebarger, M. C., & Romania, J. F. (n.d.). *Aphasia Therapy System* (U.S. Patent and Trademark Office Patent 7203649).

Lingraphica. (n.d.). *Virtual connections: An entire community at your fingertips*. https://www.aphasia.com/virtual-connections-info/

Liu, H., Cordella, C., Ishwar, P., Betke, M., & Kiran, S. (2023). Consistent long-term practice leads to consistent improvement: Benefits of self-managed therapy for language and cognitive deficits using a digital therapeutic. *Frontiers in Digital Health*, *5*, 1095110.

Lorenz, A., & Ziegler, W. (2009). Semantic vs. Word-form specific techniques in anomia treatment: A multiple single-case study. *Journal of Neurolinguistics*, *22*(6), 515–537.

Madden, E. B., Robinson, R. M., & Kendall, D. L. (2017). Phonological treatment approaches for spoken word production in aphasia. *Seminars in Speech and Language*, *38*(01), 062–074.

Maher, L. M., Gonzales Rothi, L. J., & Heilman, K. M. (1994). Lack of error awareness in an aphasia patient with relatively preserved auditory comprehension. *Brain and Language*, *46*, 402–418.

Makowski, D. (2018, March 29). Compute Signal Detection Theory Indices with R. *The PsychoBlog. Learn and Improve your R skills for Psychology.* https://neuropsychology.github.io/psycho.R/2018/03/29/SDT.html

Manolov, R., & Moeyaert, M. (2017). How Can Single-Case Data Be Analyzed? Software Resources, Tutorial, and Reflections on Analysis. *Behavior Modification*, *41*(2), 179–228. https://doi.org/10.1177/0145445516664307

Marshall, R. C., Neuburger, S. I., & Phillips, D. S. (1994). Verbal self-correction and improvement in treated aphasic clients. *Aphasiology*, *8*(6), 535–547.

Marshall, R. C., Rappaport, B. Z., & Garcia-Bunuel, L. (1985). Self-monitoring behavior in a case of severe auditory agnosia with aphasia. *Brain and Language*, *24*, 297–313.

Marshall, R. C., & Tompkins, C. A. (1982). Verbal Self-Correction Behaviors of Fluent and Nonfluent Aphasic Subjects. *Brain and Language*, *15*, 292–306.

Martin, N., & Dell, G. S. (2019). Maintenance versus transmission deficits: The effect of delay on naming performance in aphasia. *Frontiers in Human Neuroscience*, *13*, 406.

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, *41*(1), 3–29.

Mayo, N. E., Anderson, S., Barclay, Cameron, Desrosiers, J., Eng, J. J., Huijbregts, M., Kagan, A., MacKay-Lyons, M., Moriello, C., Richards, C. L., Salbach, N. M., Scott, Teasell, R., & Bayley, M. (2015). Getting on with the rest of your life following stroke: A randomized trial of a complex intervention aimed at enhancing life participation post stroke. *Clinical Rehabilitation*, *29*(12), 1198–1211.

Megyeri, M., & Szabó, B. (2024). Investigating the effectiveness of user onboarding solutions with eye tracking: a case study on paint 3D. Ergonomics in Design, 32(1), 48-55.

Mel-frequency cepstrum. (N.d.). (n.d.). Retrieved March 26, 2023, from https://en.wikipedia.org/wiki/Mel-frequency_cepstrum

Meltzer, J. A., Baird, A. J., Steele, R. D., & Harvey, S. J. (2018). Computer-based treatment of poststroke language disorders: A non-inferiority study of telerehabilitation compared to in-person service delivery. *Aphasiology*, *32*(3), 290–311.

Middleton, E. L., Chen, Q., & Verkuilen, J. (2015). Friends and foes in the lexicon: Homophone naming in aphasia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 77.

Middleton, E. L., Rawson, K. A., & Verkuilen, J. (2019). Retrieval practice and spacing effects in multi-session treatment of naming impairment in aphasia. *Cortex*, *119*, 386–400.

Middleton, E. L., Schuchard, J., & Rawson, K. A. (2020). A review of the application of distributed practice principles to naming treatment in aphasia. *Topics in Language Disorders*, *40*(1), 36–53.

Middleton, E. L., Schwartz, M. F., Dell, G. S., & Brecher, A. (2022). Learning from errors: Exploration of the monitoring learning effect. *Cognition*, *224*, 1–16.

Middleton, E. L., Schwartz, M. F., Rawson, K. A., & Garvey, K. (2015). Test-Enhanced Learning Versus Errorless Learning in Aphasia Rehabilitation: Testing Competing Psychological Principles. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, *41*(4), 1253–1261.

Middleton, E. L., Schwartz, M. F., Rawson, K. A., Traut, H., & Verkuilen, J. (2016). Towards a Theory of Learning for Naming Rehabilitation: Retrieval Practice and Spacing Effects. *Journal of Speech, Language, and Hearing Research*, *59*, 1111–1122. https://doi.org/10.1044/2016_JSLHR-L-15-0303

Miozzo, M., Pulvermüller, F., & Hauk, O. (2015). Early parallel activation of semantics and phonology in picture naming: Evidence from a multiple linear regression MEG study. *Cerebral Cortex*, *25*(10), 3343–3355.

Mirman, D., & Britt, A. E. (2014). What we talk about when we talk about access deficits. *Philosophical Transaction of the Royal Society B*, *369*. https://royalsocietypublishing.org/doi/10.1098/rstb.2012.0388

Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O. K., Coslett, H. B., & Schwartz, M. F. (2015). Neural organization of spoken language revealed by lesion–symptom mapping. *Nature Communications*, *6*(1), 6762.

Morgan, D. L. (1993). Qualitative Content Analysis: A Guide to Paths Not Taken. *Qualitative Health Research*, *3*(1), 112–121.

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

Mortenson, W. B., Sixsmith, A., & Kaufman, D. (2017). Non-Digital Game Playing by Older Adults. *Canadian Journal of Aging*, *36*(3), 342–350. https://doi.org/10.1017/S0714980817000162

Mortley, J., Wade, J., & Enderby, P. (2004). Superhighway to promoting a client-therapist partnership? Using the Internet to deliver word-retrieval computer therapy, monitored remotely with minimal speech and language therapy input. *Aphasiology*, *18*(3), 193–211. https://doi.org/10.1080/02687030344000553

Mostany, R., Chowdhury, T. G., Johnston, D. G., Portonovo, S. A., Carmichael, S. T., & Portera-Cailliau, C. (2010). Local hemodynamics dictate long-term dendritic plasticity in peri-infarct cortex. *Journal of Neuroscience*, *30*(42), 14116–14126.

Mozzafarian, D. (2016). Heart Disease and Stroke Statistics—2016 Update A Report From the American Heart Association. *Circulation*, *133*(4), e38–e48.

Nap, H. H., de Kort, Y. A. W., & Ijsselsteijn, W. A. (2009). Senior gamers: Preferences, motivations and needs. *Gerontechnology*, *8*(4), 247–262. https://doi.org/10.4017/gt.2009.08.04.003.00

Nardo, D., Holland, R., Leff, A. P., Price, C. J., & Crinion, J. T. (2017). Less is more: Neural mechanisms underlying anomia treatment in chronic aphasic patients. *Brain*, *140*(11), 3039–3054.

National Aphasia Association. (2020). *2020 Aphasia Awareness Survey*. https://www.aphasia.org/wp-content/uploads/2020/05/Aphasia-2020-Survey-Results.pdf

National Aphasia Association. (2022). *2022 Aphasia Awareness Survey*. https://www.aphasia.org/2022-aphasia-awareness-survey/

National Center for Injury Prevention and Control; Division of Unintentional Injury Prevention. (2015). *Report to Congress on Traumatic Brain Injury in the United States: Epidemiology and Rehabilitation*.

Nichol, L., Hill, A. J., Wallace, S. J., Pitt, R., Baker, C., & Rodriguez, A. D. (2019). Self-management of aphasia: A scoping review. *Aphasiology*, *33*(8), 903–942.

Nichol, L., Pitt, R., Wallace, S. J., Rodriguez, A. D., & Hill, A. J. (2023). "There are endless areas that they can use it for": Speech-language pathologist perspectives of technology support for aphasia self-management. *Disability and Rehabilitation: Assistive Technology*, *18*(8), 1473–1488.

Nichol, L., Rodriguez, A. D., Pitt, R., Wallace, S. J., & Hill, A. J. (2023). “Self-management has to be the way of the future”: Exploring the perspectives of speech-language pathologists who work with people with aphasia. *International Journal of Speech-Language Pathology*, *25*(2), 327–341.

Nichol, L., Wallace, S. J., Pitt, R., Rodriguez, A. D., Diong, Z. Z., & Hill, A. J. (2022). People with aphasia share their views on self-management and the role of technology to support self-management of aphasia. *Disability and Rehabilitation*, *44*(24), 7399–7412.

Nichol, L., Wallace, S. J., Pitt, R., Rodriguez, A. D., & Hill, A. J. (2022). Communication partner perspectives of aphasia self-management and the role of technology: An in-depth qualitative exploration. *Disability and Rehabilitation*, *44*(23), 7199–7216.

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology*, *16*(10–11), 935–979.

Nickels, L., & Best, W. (1996). Therapy for naming disorders (Part II): Specifics, surprises and suggestions. *Aphasiology*, *10*(2), 109–136.

Nickels, L., & Howard, D. (1995). Phonological errors in aphasic naming: Comprehension, monitoring and lexicality. *Cortex*, *31*, 209–237.

Nozari, N., & Dell, G. S. (2013). How damaged brains repeat words: A computational approach. *Brain and Language*, *126*(3), 327–337.

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

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

Nozari, N., & Pinet, S. (2020). A critical review of the behavioral, neuroimaging, and electrophysiological studies of co-activation of representations during word production. *Journal of Neurolinguistics*, *53*, 100875.

Nunn, K., & Vallila-Rohter, S. (2023). Theory-driven treatment modifications: A discussion on meeting the linguistic, cognitive, and psychosocial needs of individual clients with aphasia. *Journal of Communication Disorders*, *103*. https://doi.org/10.1016/j.jcomdis.2023.106327

Nunn, K., Vallila-Rohter, S., & Middleton, E. L. (2023). Errorless, Errorful, and Retrieval Practice for Naming Treatment in Aphasia: A Scoping Review of Learning Mechanisms and Treatment Ingredients. *Journal of Speech, Language, and Hearing Research*, *66*(2), 668–687. https://doi.org/10.1044/2022_JSLHR-22-00251

NVIDIA Accelerates Real Time Speech to Text Transcription 3500x with Kaldi. (n.d.). Retrieved March 23, 2023, from https://developer.nvidia.com/blog/nvidia-accelerates-speech-text-transcription-3500x-kaldi/

Ovbiagele, B., Goldstein, L. B., Higashida, R. T., Howard, V. J., Johnston, S. C., Khavjou, O. A., Lackland, D. T., Lichtman, J. H., Mohl, S., Sacco, R. L., & others. (2013). Forecasting the future of stroke in the United States: A policy statement from the American Heart Association and American Stroke Association. *Stroke*, *44*(8), 2361–2375.

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M. C., Bowen, A., Latimer, N., Julious, S., Cross, E., Alshreef, A., Harrison, M., Bradley, E., Witts, H., & Chater, T. (2019). Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): A multicentre, single-blinded, randomised controlled trial. *Lancet Neurology*, *18*, 821–833.

Palmer, R., Dimairo, M., Latimer, N., Cross, E., Brady, M. C., Enderby, P., Bowen, A., Julious, S., Harrison, M., Alshreef, A., Bradley, E., Bhadhuri, A., Chater, T., Hughes, H., Witts, H., Herbert, E., & Cooper, C. (2020). Computerised speech and language therapy or attention control added to usual care for people with long-term post-stroke aphasia: The Big CACTUS three-arm RCT. *Health Technology Assessment*, *24*(19). https://doi.org/10.3310/hta24190

Palmer, R., Enderby, P., & Paterson, G. (2013). Using computers to enable self-management of aphasia therapy exercises for word finding: The patient and carer perspective. *International Journal of Language & Communication Disorders*, *48*(5), 508–521. https://doi.org/10.1111/1460-6984.12024

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of All Non-Overlapping Data (PAND): An Alternative to PND. *The Journal of Special Education*, *40*(4), 194–204.

Parr, S. (2007). Living with Sevee Aphasia: Tracking Social Exclusion. *Aphasiology*, *21*, 98–123.

Patterson, R., Robert, A., Berry, R., Cain, M., Iqbal, M., Code, C., Rochon, E., & Leonard, C. (2015). Raising public awareness about aphasia in southern Ontario, Canada: A survey. *International Journal of Speech-Language Pathology*, *17*(2), 121–126.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Sage Publications, Inc.

Pedersen, P. M., Stig Jørgensen, H., Nakayama, H., Raaschou, H. O., & Olsen, T. S. (1995). Aphasia in acute stroke: Incidence, determinants, and recovery. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *38*(4), 659–666.

Pedersen, P. M., Vinter, K., & Olsen, T. S. (2003). Aphasia after stroke: Type, severity and prognosis: The Copenhagen aphasia study. *Cerebrovascular Diseases*, *17*(1), 35–43.

Peeva, M. G., Guenther, F. H., Tourville, J. A., Nieto-Castanon, A., Anton, J.-L., Nazarian, B., & Alario, F.-X. (2010). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *Neuroimage*, *50*(2), 626–638.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 539.

Pettit, L. K., Tonsing, K. M., & Shakila, D. (2017). The perspectives of adults with aphasia and their team members regarding the importance of nine life areas for rehabilitation: A pilot investigation. *Topics in Stroke Rehabilitation*, *24*(2), 99–106. https://doi.org/10.1080/10749357.2016.1207148

Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hanneman, M., Motlicek, P., Qian, Y., Schwarz, P., Silovsk, J., Stemmer, G., & Vesel, K. (2011). The Kaldi Speech Recognition Toolkit. *Workshop on Automatic Speech Recognition and Understanding*. IEEE, Hilton Waikoloa Village, Big Island, Hawaii, US.

Pulvermüller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., Koebbel, P., & Taub, E. (2001). Constraint-induced therapy of chronic aphasia after stroke. *Stroke*, *32*(7), 1621–1626.

Ralph, M. A. L., & Fillingham, J. K. (2007). The importance of memory and executive function in aphasia: Evidence from the treatment of anomia using errorless and errorful learning. In *Automaticity and control in language processing* (pp. 211–234). Psychology Press.

Ramsberger, G., & Marie, B. (2007). Self-Administered Cued Naming Therapy: A Single-Participant Investigation of a Computer-Based Therapy Program Replicated in Four Cases. *Americal Journal of Speech-Language Pathology*, *16*, 343–358. https://doi.org/1058-0360/07/1604-0343

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460.

Raymer, A. M., Beeson, P., Holland, A., Kendall, D., Maher, L. M., Martin, N., Murray, L., Rose, M., Thompson, C. K., Turkstra, L., & others. (2008). *Translational research in aphasia: From neuroscience to neurorehabilitation*.

Raymer, A., Thompson, C. K., Jacobs, B., & Le Grand, H. (1993). Phonological treatment of naming deficits in aphasia: Model-based generalization analysis. *Aphasiology*, *7*(1), 27–53.

Reeves, B., & Nass, C. (1996). *The media Equation: How People Treat Computers, Television, and New media like Real People and Places.* Cambridge University Press.

Roberts, S., Bruce, R. M., Lim, L., Woodgate, H., Ledingham, K., Anderson, S., Lorca-Puls, D. L., Gajardo-Vidal, A., Leff, A. P., Hope, T. M., & others. (2022). Better long-term speech outcomes in stroke survivors who received early clinical speech and language therapy: What's driving recovery? *Neuropsychological Rehabilitation*, *32*(9), 2319–2341.

Robey, R. R. (1998). A meta-analysis of clinical outcomes in the treatment of aphasia. *Journal of Speech, Language, and Hearing Research*, *41*(1), 172–187.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1–3), 107–142.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*(3), 249–284.

Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: The WEAVER++/ARC model. *Cortex*, *59*.

Roelofs, A., & Ferreira, V. S. (2019). The architecture of speaking. *Human Language: From Genes and Brains to Behavior*, 35–50.

Rohde, A., Townley-O'Neill, K., Trendall, K., Worrall, L., & Cornwell, P. (2012). A comparison of client and therapist goals for people with aphasia: A qualitative exploratory study. *Aphasiology*, *26*(10), 1298–1315. https://doi.org/10.1080/02687038.2012.706799

Rose, M., Ferguson, A., Power, E., Togher, L., & Worrall, L. (2013). Aphasia rehabilitation in Australia: Current practices, challenges and future directions. *International Journal of Speech-Language Pathology*, *16*(2), 169–180. https://doi.org/10.3109/17549507.2013.794474

Ross, K., LaPointe, L., & Katz, R. (2010). *Risk factors for depression in aphasia: Clinical implications*. Clinical Aphasiology Conference, Isle of Palms, SC.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford publications.

Salmon, J. P., Dolan, S. M., Drake, R. S., Wilson, G. C., Klein, R. M., & Eskes, G. A. (2017). A survey of video game preferences in adults: Building better games for older adults. *Entertainment Computing*, *21*, 45–64. https://doi.org/10.1016/j.entcom.2017.04.006

Sampson, M., & Faroqi-Shah, Y. (2011). Investigation of self-monitoring in fluent aphasia with jargon. *Aphasiology*, *25*(4), 505–528.

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., & others. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, *105*(46), 18035–18040.

Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., & Weiller, C. (2006). Dynamics of language reorganization after stroke. *Brain*, *129*(6), 1371–1384.

Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., Kümmerer, D., Klöppel, S., Glauche, V., Lange, R., & others. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *Neuroimage*, *49*(4), 3187–3197.

Schuchard, J., & Middleton, E. L. (2018a). The Roles of Retrieval Practice Versus Errorless Learning in Strengthening Lexical Access in Aphasia. *Journal of Speech, Language, and Hearing Research*, *61*, 1700–1717.

Schuchard, J., & Middleton, E. L. (2018b). Word repetition and retrieval practice effects in aphasia: Evidence for use-dependent learning in lexical access. *Cognitive Neuropsychology*, *35*(5–6), 271–287.

Schuchard, J., Rawson, K. A., & Middleton, E. L. (2020). Effects of distributed practice and criterion level on word retrieval in aphasia. *Cognition*, *198*, 104216.

Schutter, D. (2011). Never Too Old to Play: The Appeal of Digital Games to an Older Audience. *Games and Culture*, *6*(2), 155–170. https://doi.org/10.1177/1555412010364978

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*(2), 228–264.

Schwartz, M. F., Middleton, E. L., Brecher, A., Gagliardi, M., & Garvey, K. (2016). Does naming accuracy improve through self-monitoring ofe rrors? *Neuropsychologia*, *84*, 272–281.

Scruggs, T. E., & Mastropieri, M. A. (1998). Summarizing Single-Subject Research: Issues and Applications. *Behavior Modification*, *22*(3), 221–242.

Seah, E. T. W., Kaufman, D., Sauvè, L., & Zhang, F. (2018). Play, Learn, Connect: Older Adults' Experience With a Multiplayer, Educational, Digital Bingo Game. *Journal of Educational Computing Research*, *56*(5), 675–700. https://doi.org/10.1177/0735633117722329

Simic, T., Leonard, C., Laird, L., Cupit, J., Höbler, F., & Rochon, E. (2016). A Usability Study of Internet-Based Therapy for Naming Deficits in Aphasia. *American Journal of Speech-Language Pathology.*, *25*, 642–653.

Simmons-Mackie, N. (2008). Social Approaches to Aphasia Intervention. In *Language Intervention Strategies in Aphasia and Related Neurogenic Communication Disorders* (5th ed., p. 1091). Lippincott Williams and Wilkins.

Simmons-Mackie, N. (2017). *Aphasia in North America*. https://www.aphasiaaccess.org/white-papers/

Simmons-Mackie, N., Code, C., Armstrong, E., Stiegler, L., & Elman, R. (2002). What is aphasia? An international survey of public awareness of aphasia. *Aphasiology*, *16*, 837–848.

Simmons-Mackie, N., Damico, J. S., & Damico, H. L. (1999). A Qualitative Study of Feedback in Aphasia Treatment. *Americal Journal of Speech-Language Pathology*, *8*(3), 218–230.

Simmons-Mackie, N., & Holland, A. L. (2011). Aphasia Centers in North America: A Survey. *Seminars in Speech and Language*, *32*(3), 3–105. https://doi.org/10.1055/s-0031-1286175

Simmons-Mackie, N., Kagan, A., Christie, O., Huijbregts, M., McEwen, S., & Willems, J. (2007). Communicative access and decision making for people with aphasia: Implementing sustainable health care systems change. *Aphasiology*, *21*(1), 39–66.

Sindelar, P. T., Rosenburg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children.*, *8*, 67–76.

Stefaniak, J. D., Halai, A. D., & Lambon Ralph, M. A. (2020). The neural and neurocomputational bases of recovery from post-stroke aphasia. *Nature Reviews Neurology*, *16*(1), 43–55.

StepByStep Aphasia Therapy. (N.d.). (n.d.). [Computer software]. https://aphasia-software.com/

Stockert, A., Kuemmerer, D., & Saur, D. (2016). Insights into early language recovery: From basic principles to practical applications. *Aphasiology*, *30*(5), 517–541.

Strijkers, K., Costa, A., & Pulvermüller, F. (2017). The cortical dynamics of speaking: Lexical and phonological knowledge simultaneously recruit the frontal and temporal cortex within 200 ms. *NeuroImage*, *163*, 206–219.

Swales, M. A., Hill, A. J., & Finch, E. (2016). Feature rich, but user-friendly: Speech pathologists' preferences for computer-based aphasia therapy. *International Journal of Speech-Language Pathology*, *18*(4), 315–328. https://doi.org/10.3109/17549507.2015.1081283

Sze, W. P., Hameau, S., Warren, J., & Best, W. (2020). Identifying the Components of a Successful Spoken Naming Therapy: A Meta-Analysis of Word-Finding Interventions for Adults with Aphasia. *Aphasiology*, *35*(1), 33–72.

The International Picture-Naming Project. (n.d.). https://crl.ucsd.edu/experiments/ipnp/

The RELEASE Collaborators, Brady, M. C., Ali, M., VandenBerg, K., Williams, L. J., Williams, L. R., Abo, M., Becker, F., Bowen, A., Brandenburg, C., & others. (2022). Precision rehabilitation for aphasia by patient age, sex, aphasia severity, and time since stroke? A prespecified, systematic review-based, individual participant data, network, subgroup meta-analysis. *International Journal of Stroke*, *17*(10), 1067–1077.

Thompson, C. K. (2006). Single subject controlled experiments in aphasia: The science and the state of the science. *Journal of Communication Disorders*, *39*, 266–291.

Townend, E., Brady, M. C., & McLaughlan, K. (2007). Exclusion and inclusion criteria for people with aphasia in studies of depression after stroke: A systematic review and future recommendations. *Neuroepidemiology*, *29*(1–2), 1–17. https://doi.org/10.1159/000108913.

Ueno, T., Saito, S., Rogers, T. T., & Ralph, M. A. L. (2011). Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*, *72*(2), 385–396.

Unity. (n.d.). [Computer software]. https://unity.com/

User Onboarding: Principles and Guidelines (Cibin, KS). Retrieved June 04, 2024, from https://www.uxmatters.com/mt/archives/2022/01/user-onboarding-principles-and-guidelines.php

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing and Health Sciences*, *15*, 398–405. https://doi.org/doi: 10.1111/nhs.12048

Valente, A., Bürki, A., & Laganaro, M. (2014). ERP correlates of word production predictors in picture naming: A trial by trial multiple regression analysis from stimulus onset to response. *Frontiers in Neuroscience*, *8*(390), 1–13.

van Hees, S., Angwin, A., McMahon, K., & Copland, D. (2013). A comparison of semantic feature analysis and phonological components analysis for the treatment of naming impairments in aphasia. *Neuropsychological Rehabilitation*, *23*(1), 102–132.

van Hees, S., McMahon, K., Angwin, A., de Zubicaray, G., & Copland, D. A. (2014). Neural activity associated with semantic versus phonological anomia treatments in aphasia. *Brain and Language*, *129*, 47–57.

Varkanitsa, M., Godecke, E., & Kiran, S. (2023). How Much Attention Do We Pay to Attention Deficits in Poststroke Aphasia? *Stroke*, *54*, 55–66. https://doi.org/10.1161/STROKEAHA.122.037936

Vestling, M., Tufvesson, B., & Iwarsson, S. (2003). Indicators for return to work after stroke and the importance of work for subjective well-being and life satisfaction. *Journal of Rehabilitation Medicine*, *35*(3), 127–131.

Wafa, H. A., Wolfe, C. D. A., Emmett, E., Roth, G. A., Johnson, C. O., & Wang, Y. (2020). Burden of Stroke in Europe. Thirty-Year Projections of Incidence, Prevalence, Deaths, and Disability-Adjusted Life Years. *Stroke*, *51*. https://doi.org/10.1161/STROKEAHA.120.029606

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

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

Wang, G., Ge, L., Zheng, Q., Huang, P., & Xiang, J. (2020). Constraint-induced aphasia therapy for patients with aphasia: A systematic review. *International Journal of Nursing Sciences*, *7*(3), 349–358.

Warner, G., Packer, T., Villeneuve, M., Audulv, A., & Versnel, J. (2015). A systematic review of the effectiveness of stroke self-management programs for improving function and participation outcomes: Self-management programs for stroke survivors. *Disability and Rehabilitation*, *37*(23), 2141–2163.

What Works Clearinghouse. (2023). https://ies.ed.gov/ncee/wwc/

Wienrich, C., Reitelbach, C., & Carolus, A. (2021). Assistants in the Context of Healthcare Investigating the Effect of Perceived Expertise on the Trustworthiness of Voice Assistants, Providers, Data Receivers, and Automatic Speech Recognition. *Frontiers in Computer Science*, *3*, 1–12. https://doi.org/10.3389/fcomp.2021.685250

Williams, S. E., & Freer, C. A. (1986). Aphasia: Its effect on marital relationships. *Archives of Physical Medicine and Rehabilitation*, *67*(4), 250–252.

Wilson, S. M., & Schneck, S. M. (2020). Neuroplasticity in post-stroke aphasia: A systematic review and meta-analysis of functional imaging studies of reorganization of language processing. *Neurobiology of Language*, *2*(1), 22–82.

Winstein, C., Stein, J., Arena, R., Bates, B., Cherney, L., & Cramer, S. (2016). Guidelines for adult stroke rehabilitation and recovery: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, *47*, e98–e169.

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

Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersh, D., Ferguson, A., & Davidson, B. (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, *25*(3), 309–322. https://doi.org/10.1080/02687038.2010.508530

Yourganov, G., Smith, K. G., Fridriksson, J., & Rorden, C. (2015). Predicting aphasia type from brain damage measured with structural MRI. *Cortex*, *73*, 203–215.

Zhang, J., Yu, J., Bao, Y., Xie, Q., Xu, Y., Zhang, J., & Wang, P. (2017). Constraint-induced aphasia therapy in post-stroke aphasia rehabilitation: A systematic review and meta-analysis of randomized controlled trials. *PloS One*, *12*(8), e0183349.

Zheng, C., Lynch, L., & Taylor, N. (2016). Effect of computer therapy in aphasia: A systematic review. *Aphasiology*, *30*(2–3), 211–244. https://doi.org/10.1080/02687038.2014.996521

APPENDIX A

Qualitative Content Analysis: Example of How an Excerpt from the Interview was Analyzed to Extract Relevant Themes

Phase 1: Familiarization and Transcription

- [Interviewer] How often do you attend therapy sessions?

  [Participant] I go for speech therapy one hour per week, and it's something good...but I would like to try...but there's nobody, there's no place to go. [Name of therapist] took me for so long after [the stroke], and then said "Okay, that's it, that's enough, I'm not making you any better", and she put me out. I still go to her once a week [*here referring to a group of people with aphasia lead by the former therapist, not to formal therapy*], and that's good. I want get better. It wasn't ever boring or anything, she would make me, like how we take a book, and I would read, a bit of her book, and then after write down a synopsis...what the book was. And it was very hard for me. And she saw that. And I tried, I couldn't do it. That was a hard one. But then we played with games or whatever, where you'd guess a fruit, whatever - that was easy.

- [Interviewer: You still want to do therapy, and not doing any. Why?

  [Participant] I don't know. I've gotten good, and then they say "That's it, I'm not making you any better".

- [Interviewer: What do you think are strategies, exercises or activities that are helping you?]

  [Participant] I don't know...she has a little board, and she says "Okay, write as many green foods as you can". So you go, and they give you 2 minutes...and that's good for me, I think. The stuff like that is, it's good, it makes me, it makes me - try...You make me want to read this, I can read maybe four or five lines - the rest of it - it's all just jumbled. I'll sit down and I'll take half an hour or so and read that whole thing, no problem. But, until I have a half hour to read that whole thing, it's just jumbled. And when I get it, then okay, I can answer a little bit myself.

Phase 2: Search for Meaning Units

- [Interviewer] How often do you attend therapy sessions?

  [Participant] I go for speech therapy one hour per week, and it's something good...but I would like to try...but there's nobody, there's no place to go. [Name of therapist] took me for so long after [the stroke], and then said "Okay, that's it, that's enough, I'm not making you any better", and she put me out. I still go to her once a week [*here referring to a group of people with aphasia lead by the former therapist, not to formal therapy*], and that's good. I want get better. – **Meaning Unit 1**

  It wasn't ever boring or anything, she would make me, like how we take a book, and I would read, a bit of her book, and then after write down a synopsis...what the book was. And it was very hard for me. And she saw that. And I tried, I couldn't do it. That was a hard one. But then we played with games or whatever, where you'd guess a fruit, whatever - that was easy. **– Meaning Unit 2**

- [Interviewer: You still want to do therapy, and not doing any. Why?

  [Participant] I don't know. I've gotten good, and then they say "That's it, I'm not making you any better". – **Meaning Unit 3**

- [Interviewer: What do you think are strategies, exercises or activities that are helping you?]

  [Participant] I don't know...she has a little board, and she says "Okay, write as many green foods as you can". So you go, and they give you 2 minutes...and that's good for me, I think. The stuff like that is, it's good, it makes me, it makes me - try...You make me want to read this, I can read maybe four or five lines - the rest of it - it's all just jumbled. I'll sit down and I'll take half an hour or so and read that whole thing, no problem. But, until I have a half hour to read that whole thing, it's just jumbled. And when I get it, then okay, I can answer a little bit myself. **– Meaning Unit 4**

Phase 3: Search for Themes

- [Interviewer] How often do you attend therapy sessions?

  [Participant] I go for speech therapy one hour per week, and it's something good...but I would like to try...but there's nobody, there's no place to go. [Name of therapist] took me for so long after [the stroke], and then said "Okay, that's it, that's enough, I'm not making you any better", and she put me out. I still go to her once a week [*here referring to a group of people with aphasia lead by the former therapist, not to formal therapy*], and that's good. I want get better. – **Meaning Unit 1**

  Inability to find a therapist that could help improve communicative skills past a certain level – **Theme 1**

  It wasn't ever boring or anything, she would make me, like how we take a book, and I would read, a bit of her book, and then after write down a synopsis...what the book was. And it was very hard for me. And she saw that. And I tried, I couldn't do it. That was a hard one. But then we played with games or whatever, where you'd guess a fruit, whatever - that was easy. – **Meaning Unit 2**

  Writing synopsis of books and playing games as strategies employed during therapy that were deemed helpful – **Theme 2**

- [Interviewer: You still want to do therapy, and not doing any. Why?

  [Participant] I don't know. I've gotten good, and then they say "That's it, I'm not making you any better". – **Meaning Unit 3**

  Inability to find a therapist that could help improve communicative skills past a certain level – **Theme 1**

- [Interviewer: What do you think are strategies, exercises or activities that are helping you?]

  [Participant] I don't know...she has a little board, and she says "Okay, write as many green foods as you can". So you go, and they give you 2 minutes...and that's good for me, I think. The stuff like that is, it's good, it makes me, it makes me - try...You make me want to read this, I can read maybe four or five lines - the rest of it - it's all just jumbled. I'll sit down and I'll take half an hour or so and read that whole thing, no problem. But, until I have a half hour to read that whole thing, it's just jumbled. And when I get it, then okay, I can answer a little bit myself. – **Meaning Unit 4**

  Time-constrained exercises as a strategy to promote engagement – **Theme 2**

Phase 5 (after themes have been reviewed for consistency with the meaning units they have been derived from): Assign Name to Themes

Theme 1: Barriers to access therapy

Inability to find a therapist that could help improve communicative skills past a certain level

Theme 2: Factors making therapy engaging

Writing synopsis of books and playing games as strategies employed during therapy that were deemed helpful
Time-constrained exercises as a strategy to promote engagement

Table A1: Demographics

Subject	Age	Sex	Time Since Stroke	Type of Aphasia	Education	Language		Experience with Therapy	
						L1	L2	Past	Present
S1	59	F	14y	Receptive; anomia	Diploma (X-Ray technologist)	English	-	8y of 1-to-1 2d/w	5y group therapy 3d/w
S2	67	M	9y	Expressive; reading deficits	BSc. (Civil Eng.)	English	-	8y of 1-to-1 1h/w	Group + 1-to-1 2h/w
S3	61	M	8y	Expressive	Bachelor (Insurance)	English	French	2y of 1-to-1 2d/w	-
S4	61	M	4y	Mild expressive	High School	English	-	ND	ND
S5	51	F	7y	Expressive	Bachelor (Education)	English	-	1-to-1 1d/month	-
S6	-	M	4y	-	Radio College of Canada	English	-	1-to-1 2d/w	Group therapy
S7	69	M	7y	Expressive; reading and writing	High School	English	-	1.5y of 1-to-1, 1h/w	-
S8	59	F	20y	Receptive; expressive	Diploma (Hairdresser)	French	English	3y 1-to-1	1-to-1 2d/month
S9	-	M	4y	Mild expressive	-	English	-	-	1-to-1

Table A.2: Consequences of aphasia

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S3	The first thing, there was nothing. I looked like a blithering idiot for...that was three weeks of my stroke is gone. There was nothing I remember, and first week it took me a long time to say my own name. So that was tough. Really tough. I went to rehab centre.	Lack of memory in the first 3 weeks after stroke occurred	Inability to communicate in the immediate aftermath of stroke
	I took me a year and a half to get it back [ability to communicate]		
S4	It's a bit hard to separate the aphasia from the rest of it, because I was completely paralyzed in the right side of my body...and I couldn't talk. So I guess it improved, and I learned to speak all over again. I had to start up with words, sounds (like 'm' sound, the 'l' sound). Now I don't know how you differentiate that from the aphasia. Can you?	Hemiplegia and inability to talk in the immediate aftermath of stroke ability to express any idea in the immediate aftermath of stroke	
	In my mind aphasia comes a little bit after, when I have lost the words. At the beginning it was just the sounds and the repeating, sounds and learning to talk. I know the words, I know what I want to say, and I can't say it, and aphasia makes me not say it. Does it make sense?	Inability to express any idea in the immediate aftermath of stroke	
S8	I wasn't talking at all [soon after aphasia occurred]	Inability to express any idea in the immediate aftermath of stroke	
S1	It took me several years to learn the English language. My speech is more fluent than it used to be, but I still have to find words. I do speak differently than I used to because before I could just say what I want to say, but now, I can't.	Difficulty finding the right	Long-lasting physical, cognitive and linguistic deficits caused by stroke/aphasia
	Phone conversations can be difficult. If I have to go through menus [referring to automatic voice presenting multiple options to direct the call to] … people don't generally answer the phone right now. That can be a challenge.	Challenges experienced having conversations on the phone	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S1	I guess it's different situations really that can.. cause trouble.. with numbers especially. It's recalling them back, especially if it's on the phone, like phone numbers because if they say it quickly, I won't get it. Some numbers switch fours and fives, and there's a bunch of the same numbers, they turn into another number, which is the wrong one…You have to think it's like different stages of memory when you're listening to what people are saying. It's not just one step. It's understanding what they are, what it is, and what it is they want you to do, so it's like three or four steps. Of course, they're still talking, and you know you might not be on the same page.	Difficulty holding information about numbers in memory	
	I think with noise, sometimes you really can't do much about it	Struggle dealing with background noise	
S2	[Interviewer: How has aphasia affected your life?] Walking, reading, communication [difficulties with]		
S4	[Interviewer: What part is most difficult for you?] Noise, any rumor, in the crowd/party or something like that. It's kind of like, I don't get anything, just noise. I would have to pull one person and concentrate on a person's voice and talk to one person, so I don't get anything from anybody else. That's hard.	Struggle with understanding in the presence of white noise	
	I could read every email message. But when I try to respond, what the symbol I'm going to say OK, I know that's what I want to do. So I just say "O", but I did not know what letter to put for "O", and try every letter, but none did. So that's how I found I could not write. It's better now, but I still have hard time. In short messages I make all sorts of spelling errors, sometimes not even being close that the program can guess what I'm trying to say, not even close.	Struggle with writing	
S5	I have problems with reading and writing and talking. I can talk to you and we can converse quite freely but if I were to explain something to you even though I knew I couldn't explain it to you...that is still hard after 7 years.	Struggle recovering communication skills	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S6	I look at my life at Rogers [where the interviewee worked]... I would get a big long email from somebody, and I have to...it takes me an hour or so to read something that I used to get in five minutes	Struggle with reading	
S7	[Difficulties with] getting the point out, reading - I can only read 2 or 3 words/lines, writing, listening	Difficulties with communication	
S8	Yes [about suffering deficits in understanding]. Because I can talk and that, but it doesn't mean I know what he's saying. Or she's saying. Sometimes, I can - like some words are coming out, little by little...I got to think about it.	Difficulties understanding	
	Yes I have [problems reading]. I'm trying to read now...She [the therapist] gives me a book, a little book, with little sections that I can read. I have to read and read, and then it'll come to me.	Difficulties with reading	
	I can't use my right hand		
	I can get things, like long things to say. I've got a thing if it's harder, like keep going harder and harder. I don't know if it's like, how can I put it, I'm not sure. I can say some things and some things I can't.	Difficulties getting the point across	
S9	Before my stroke, with publications...it takes me about 50% longer to read publications..	Difficulties reading	
S1	I was not able to go back to work after I had the stroke.	Inability to hold previous job or getting a new one	Difficulties holding a job or getting a new one due to stroke/aphasia
	The problem that I have still, it's better than it was before, is understanding what people say. So really, an employer wants you to know what people are saying. I have tried several times to go back to work, but it wasn't very successful.	Difficulty holding a job or getting a new one due to problems with communication	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S3	I think it took me about a year to admit I was...I wasn't going back to work. After that it was learning how to retire.	Inability to hold job due to aphasia	
S5	I am working small jobs, I am not teaching	Inability to hold previous job	
S6	I was a manager at Rogers for ten years, and then this thing [stroke occurred]	Loss of job due to consequences of stroke	
S1	Trying to get the main point of what they're saying. Sometimes you do just have to repeat them and ask them to repeat. I have been more quick to just talk to people if it's just one-on-one because I've missed usually just the main word. So, I'm faster to tell them it's just this word because they like to repeat the whole sentence. Which they don't have to.	Asking people to repeat as a way to facilitate understanding	Miscellanea about strategies used to help with aphasia
S2	[Partner] He can't get it out...then he moves on. Sometimes if we're home I ask him if he can write it, even if it's just the first letter.	Writing/spelling as a way to help getting the message across	
S5	I use technology all the time and if I have problems explaining something I will write a note in an email and send it off because I can talk to my phone and I use to have it dictate but I had to talk in order for it to type. It is a long road. I use my Iphone now and talk and it types and it's been a blessing.	Using the iPhone to help getting the message across by writing notes or using speech-to-text functionality to facilitate writing	
S7	Keywords, pointing [to-be named item], use gestures	Strategies used to help getting the point across	
	I read slower, small amounts to help. Same for writing.	Strategy used to help understanding	
S9	Before my stroke, with publications...it takes me about 50% longer to read publications, and with that I would need to have a dictionary on my computer at that point just to double-check whether the word is correct or not.	Difficulties reading and strategies to help with understanding and writing	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
	[Interviewer: Previously you mentioned you use dictionary, pen and paper to draw/write down ideas. When you are fact-to-face, what strategies do you use to help with communication?] When I'm talking to people, I think just to keep things flowing with conversation...if I have a word that I stumble on, to keep things flowing I ask "can you help me...you know the word "verb...", can you say after me so that I can hopefully repeat it", and then it would continue, that type of thing...as long as the person knows that I have aphasia right, that's one thing…	Asking for help when the right word does not come out	
S9	If I'm at the store, at a bank, especially if it's something like...if I have to go at a bank, and I have to talk about opening an account, or something like that, I will try to go through in mind a list of words that I would use, and then I go and talk to them. Sometimes there would be some questions, it might be a word that I didn't think about, but I'll try to get it through, one way or another, maybe I'll try to use a different type of word, but those are the ways that I deal with it.	Preparing a list of keywords to use before starting a conversation as a way to facilitate communication	
S1	Just about everything [changed after the onset of aphasia] really. When I was working, I was busy with that. I used to be on a call quite often. I would be called in so that was time-consuming. I used to curl, but I really stopped curling because of the job, because it took so much time really…. I was divorced. I had to move. And I had to relearn how to speak. So, yeah. Everything was changed.	Disruption in everyday life due to stroke-induced communication difficulties	Miscellanea about changes in lifestyle and habits caused by aphasia, or embraced to positively adapt to it
S3	It changed my every, my routine every day. While the easiest way to tell it is my wife still works, I'm staying home and cleaning house, making the meals [and] stuff like that. Taking me a little longer than normal, but I do it.	Finding ways to adapt to aphasia and contribute to wellbeing of family	
	Another thing has changed is my appetite, my smoking. I don't smoke anymore. And my coffee intake, cut it in about, one eight?	Quit smoking and reduce coffee intake as positive lifestyle changes adopted after onset of aphasia	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S3	What you gotta do is admit to yourself that you are going to retire. Once you, once you admit to yourself that you are, it's easy. Like if it rains or snows, I stay inside.	Adapting lifestyle to live with aphasia	
	I was reading about four books per week.. [it was about] mystery, history, little political, sports, and I don't read as much as I used to. But every morning that I stay first help on the computer, CBC, NBC, about like I check all these internet stuff, too, to take a look around the world.	Drastic reduction in time spent reading. News-check on internet as new habit developed after onset of aphasia	
S1	If it's...you know even just white noise is a challenge too because I may not understand what people are saying for that reason. Most people do not really, they don't know what aphasia is. So if I make a mistake, I hesitate now to tell them because some people will say, "well, you look fine."	Frustration due to lack of awareness about aphasia	
S3	Normally everybody can understand me, eventually. Some days are better than other ones. Normally it takes me a little longer to explain myself to anybody that there's no, first it's [to] meet me. That's the only thing that I hate about the stroke, because I was born a salesman.	Difficulties with communication and the effect on personal identity	

Table A.3: *Barriers to access therapy*

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S1	I think there was a couple of years that I used to go twice a week, because initially you would go to learn at Hearing and Speech, which MSI pays for, but there's a waitlist. There's always a waitlist. But I used to go to a private one at the same time, so when I was doing both, I would go twice…. You would often be put on a waitlist again for Hearing and Speech. So that I would only have the private one.	Long waitlist as a barrier to access publicly supported therapy services	Long waitlist as a barrier to access publicly supported therapy services
S2	[Partner: With the rehab we had to wait at least 6 months and that's not good. We left hospital and we had to wait 6 months. Fortunately we have coverage for private specialist; if he had waited 6 months doing nothing, I think doing speech not much help, you know from the government.]	Long waitlist as a barrier in accessing publicly supported therapy services	
S3	First 6 months, I sat at home by myself staring at the computer figuring out how to handle it...because well I have to think about it. Then 3 months in rehab, maybe 4, and then another 2 months, 3 months by myself. Takes you a long time to get somebody there to help you unless you're like myself, had an insurance plan in place because the average person can't afford to get it.	Long waitlist as a barrier in accessing publicly supported therapy services	
S9	Yes [about having had problems with access to therapy] but that was through Hearing and Speech. If I was looking for a provider no problem from there, as long as you pay. There's no issue there. But for Hearing and Speech it's been on and off. So basically in the first 4 months it was there, then for the next 6 months I could not have access, than I was able again...I've been on and off...3 months ago I started again.	Long waitlist as a barrier to access publicly supported therapy services	
S1	Probably you could travel, you could go to other centers and other cities, but it's usually very costly. The access for financial help within government, within the MSI framework is unequal...for aphasia, when you're discharged from a hospital.. you're on your own. You're on a waitlist to go Hearing and Speech. I think it might be a little different now. I think it's a little bit better, but it's what is was like when I had a stroke.	Financial barriers to access therapy services	Miscellanea about barriers to access to therapy
S5	I use the aphasia applications but we have to pay for it to get advanced.. I can do all the things no problem but I have to pay for it to get the advanced version and I think it's a [incomprehensible], I don't know but I would love to find one that I do not have to pay for and that is on my level…	Desire to get applications that are affordable and that offer the right level of challenge	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S5	I was in the hospital for about three months rehab...and I received excellent service at all places and I was warned that you have to set up services before you go home... I was in the hospital for about a week longer than I needed because the services were not set up for me, if I had gone home a week before I would have gone to the bottom of the list, do you know what I mean?... I stayed the extra week and got the services all lined up, I had no idea. My sister was a nurse and she said "wait for the services to be lined up or you're screwed"	Poor communication and assistance about how to enroll for publicly-funded therapy	
S6	I go for speech therapy one hour per week, and it's something good...but I would like to try...but there's nobody, there's no place to go. [Name of therapist] took me for so long after [the stroke], and then said "Okay, that's it, that's enough, I'm not making you any better", and she put me out. I still go to her once a week [here referring to a group of people with aphasia lead by the former therapist, not to formal therapy], and that's good. I want get better. [Interviewer: You still want to do therapy, and not doing any. Why?] I don't know. I've gotten good, and then they say "That's it, I'm not making you any better".	Difficulty finding a therapist that could help improve communication skills past a certain level	
S7	[Interviewer: Are there any challenges with receiving more therapy? Is it expensive?] Yes	Financial barriers to access adequate amount of therapy	
S8	[Interviewer: Would you like to do more than two times a month?] She [the therapist] does all my stuff. But she used to have another one that used to be at the hospital.. After a while I got hit in the head, and all my arm and up.. I was doing with her once a week...But now it isn't - I've been calling her, and I've never got her. I would like to do more..	Difficulty getting adequate access to therapy service	

Table A.4: Opinions of people with aphasia about positive aspects of therapy

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S5	[Interviewer: Was there something that you would have liked to be different in the therapy?] Looking back we never did much technology and I would have liked more technology and less going over words. She gave me a binder with words and I had to go home and say words but...that was pretty boring but it was great too. [Interviewer: Was therapy helpful?] Yes	Desire to engage more with technology to help with recovery	Miscellanea about aspects of therapy that could be improved
	[Interviewer: Would you have liked part of the time to be dedicated to the use of some sort of device?] Using different ways to access or use the same things I use the aphasia applications but we have to pay for it to get advanced.. I can do all the things no problem but I have to pay for it to get the advanced version and I think it's a [incomprehensible], I don't know but I would love to find one that I do not have to pay for and that is on my level…	Desire to get access to applications that are affordable and that offer right level of challenge	
S1	...That helped a lot [referring to a rehabilitation program the interviewee was referred to] because it was the sounds of the words that I really had a lot of trouble, and that's one of the things that they kind of focused on, so that helped a lot really. And they said that it can do well on most of the speech therapy, and these continue on. So, that did, I think well for quite a long time…. And reading too. We read a book out loud, which helped with fluency and also understanding...I think I had to do some essays too, which was also trouble for me…	Training sound-to-letter mapping and reading book aloud as a way to ameliorate reading skills	Miscellanea about strategies employed during therapy that are considered particularly helpful
	They are all different [referring to book club] because the students have different themes each day. I mean it's similar because games or different strategies that these students...it's like a Jeopardy thing. It's like word-finding and sometimes there's just conversation within the other groups too...They do different things each week really, which would be part of one of their practice.	Games used to stimulate comprehension and word-finding skills at book club sessions	
	[Interviewer: So one of the reasons you think therapy is engaging is because it is motivating, you had the motivation to improve. Are there any other aspects you like?] Yes, it usually is. It's interesting because you now see what you can remember or how you can improve from the last session, to try improve.	Therapy promotes motivation by providing a measure of one's own current skill level and of day-by-day progress	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S3	She [the therapist] started with A,B,C, all the way through the [alphabet], and then I was released and went home...and then I got this stuff at the province that took me about six months that was going to start at Cobequid...every week I used to see what gave me a problem, and I'd go. I'd write it all myself, and then I had to explain it this way is.	Identifying and reporting about problems encountered during the week as an approach to help with aphasia	
	[Interviewer: Did you find therapy engaging?] Oh yes, because she [the therapist] was tough, tough on me because that's what makes the program work...I don't know it would've helped if it hadn't been tough.	Therapy perceived as engaging because therapist was tough	
S5	[Interviewer: When you did therapy in the past, did you enjoy it?] Yes, very much...the speech language therapist would give me different things so I wouldn't get bored, and I used to be from a farm and she would get things that I liked - she was amazing.	Finding the use of new and relevant material a feature of therapy that made it stimulating	
S6	It wasn't ever boring or anything, she would make me, like how we take a book, and I would read, a bit of her book, and then after write down a synopsis...what the book was. And it was very hard for me. And she saw that. And I tried, I couldn't do it. That was a hard one. But then we played with games or whatever, where you'd guess a fruit, whatever - that was easy.	Writing synopsis of books and playing games as strategies employed during therapy that were deemed helpful	
	[Interviewer: What do you think are strategies, exercises or activities that are helping you?] I don't know...she has a little board, and she says "Okay, write as many green foods as you can". So you go, and they give you 2 minutes...and that's good for me, I think. The stuff like that is, it's good, it makes me, it makes me - try...You make me want to read this, I can read maybe four or five lines - the rest of it - it's all just jumbled. I'll sit down and I'll take half an hour or so and read that whole thing, no problem. But, until I have a half hour to read that whole thing, it's just jumbled. And when I get it, then okay, I can answer a little bit myself.	Time-constrained exercises as a strategy to promote engagement	
S7	[Interviewer: What do you enjoy about therapy?] Just talk. [Interviewer: Interacting with people?] Yes	Talking and interacting with people as aspects of therapy that promote engagement	
S8	[Interviewer: Of the different things that you try with your therapist, which do you think is the one that helps you most?] I like reading. That's where I'm starting to read, a little bit...I just gotta go-go-go, I can't stand staying in the house...I'm always going and I'm always talking and talking	Reading as a task that is perceived as very helpful	

Subject	Meaning	Condensed Meaning Unit	Theme
S9	[Interviewer: Do you enjoy therapy?] Yes. For me a lot of it is the ability just to talk. It helps me. And I use different words than I would normally do. Knowing that she's a Speech Language Therapist and she will be able to help, and give me ideas on how to deal with type of words, readings, along that line. It's not really anything that we done, with the therapist, that I have enjoyed.	Working in tandem with the therapist provides a sense of security that encourages the patient pushing his limits	
	For me it works well. She'll work through a couple of questions, and that type of things, and then she'll send things home with me so that I work on them... Most of them are more with memory and speech, basically you got four or three words, and you have to try to remember all four, and the other thing too is "there's something similar about a couple of words?"	Homework based on memory and speech exercises	
S1	Initially because you're trying to regain some of your skills, it's horrifying because you start...with word findings and identifying common shapes, which you a few days ago would have known automatically. It's just things that you just, they're gone. Or you call them another name. It's lost. So, you have to get it back.	Distress experienced when struggling with tasks that were once performed with no	
S5	[Interviewer: Why did you stop doing therapy?] The Speech Language pathologist is awesome, but I just felt like living my life was good enough...Putting myself out there was good enough...I could go to a speech therapist for an hour but I would rather spend an hour out there talking to people, and putting what she teach me in place.	Interrupting therapy when it was felt that it could not add value over conducting an active social life	
S7	[Interviewer: Did you liked more group therapy or one-on-one?] I really don't know...if I had two people...missing something there…[Interviewer: With more people it's more difficult to understand the conversations?] Yeah but you don't get enough... there are two..[Partner: You don't get enough time to express yourself?] Yeah	Difficulties engaging in conversations with more than one person	
S8	[Interviewer: Where there any moments where you needed therapy, but you couldn't get it for any reason?] Yes I did. I went with it for 13 years, I couldn't do it...because I was making no progress...I just didn't do it...and then...I had my little girl a year later...I had to get her to school...she did it right through high school... and she used to tell me "Why don't you start talking again?", and I said "Okay, I'll start again", 16 years later.		

Table A.5: SLTs views on the common challenges and unmet needs experienced by people with people

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT1	[Interviewer: What challenges do you face in delivering therapy?] Everybody's gonna have some different challenges. The big challenge in my area is transportation, access to the service. It might not seem too concerning in bigger centers where there is public transportation system; however, in smaller, rural communities, even just being able to come to an outpatient office might require a week or so of planning from the caregiver, or if the patient does not have any family member or the family members all work, then there have been times where people just don't have the service at all because they can't come to outpatient.	Barrier to access therapy in rural areas due to difficulties with transportation	Difficulties with transportation as main barriers to access to therapy
SLT3	...transportation can be an issue in terms of attending the 1-1 group therapy...It could be about cost - having to pay for transportation, a cab, parking, bus - could be a deterrent.	Barrier to access therapy due to difficulty with transportation	
	...they are not always able to make it because of where they live...	Barriers to access therapy due to difficulty with transportation	
SLT3	[Interviewer: What kind of challenges did you find in delivering therapy?] I would say [there is a challenge in] meeting that intensity due to time constraints or [lack of] funding...it could be that they are sharing a caseload with pediatric). So even if PWA are willing to be seen there might be a funding issue.	Financial barriers to access therapy	Financial barriers to access therapy
	When we are able to offer intensive programs - so we've offered constraint induced [therapy], this kind of thing - they are not always able to make it because of where they live, or cost, or maybe family, or other support.		
SLT1	I think something that should also be considered is the repercussions in the family dynamic. There should be a connection immediately with the family or caregivers, so that they know where to go, and if they have questions because of course they are going to be affected as well in a different way…	Inadequate involvement of family/caregivers when defining therapy outcomes	Inadequate service delivery model

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT2	Most significant barrier would be having a support network for that client...to help with the therapy, so whether that's a spouse or another family member who is capable of participating in the training.	Lack of social network to tandem with therapist to support recovery	
	[Interviewer: What are the greatest unmet needs of people with aphasia?] Community support, so there needs to be something. [Interviewer: Stronger network?] Like therapy- maintenance I guess you could call it, so they are at a point where there is a plateau and a discharge from clinical programs necessary but there are functional goals that can still be achieved in a supportive community environment and whether that same kind of accessible support group or whether that's community therapy in house support. [Interviewer: Okay, so you think these are also the possible solutions - to strengthen this network making it more organized and more efficient] Yes.	Lack of community support that can provide necessary assistance to continue improve when patients are discharged from therapy	
	[Interviewer: What are the biggest challenges that you as a therapist face in delivering treatment to people with aphasia?] The challenges here in this province would be staffing. So each staff would go around to do the intensive therapy in the acute stage and then again it's financial - to be able to provide community service or outreach on a team... so working within the service delivery model that we have and with the restricted resources there are significant wait times for the outpatient portion which should be community based, but really isn't.	Lack of personnel along with an inadequate service delivery model as a challenge in meeting the needs of individuals with aphasia	
SLT3	...I think as well if there are not that sort of supportive family or friends to help with carryover etc.,		
SLT2	[Interviewer: What are the features of speech and language therapy that are most important for aphasia recovery?] It certainly depends on the stage of recovery whether you're in acute care or whether - you've moved into rehab or even home-based but...when you're in acute care, you need intensive therapy and as you move outside of that, you really need speech interactive therapy where they're setting functional goals individually.	Intensity and a focus on functionally significant tasks/outcomes as factors that promote recovery	Intensity and functionally meaningful goals as factors that promote recovery

Subject	Meaning Unit	Condensed Meaning unit	Theme
SLT3	[Interviewer: What do you think are the most important factors in SLT that help with recovery?] I would say intensity, and functionality (so making goal materials functional). Having family members active during goal setting and throughout the therapy process, as well as the opportunity to meet other people living with aphasia, and encouraging other disciplines (recreation therapy, involving them in them the process and using their communication in activities that they did enjoy or find new activities to practice)	Intensity and a focus on functionally significant tasks/outcomes as factors that promote recovery	
SLT1	[Interviewer: What are the biggest challenges that people with aphasia face when they are in the process of doing speech-language therapy?] Number one, it is quite common that there are mental health issues that emerge following stroke...we are seeing a lot more patients who are younger ...and definitely mental health is a factor for a lot of these individuals, simply because they're still in the prime of their life, and all of a sudden they are aware that they can't do what they used to do or don't have the communication skills to be able to return to work.	Mental health issues (depression) as an additional consequence of aphasia	Mental health as an underestimated factor affecting quality of life of people with aphasia
	So there is definitely an element of depression, which I really think needs to be addressed. I don't think it's being taken as seriously, maybe with some of the family traditions. I think that one remedy is to say "Well I'll prescribe a pill for you", but it's not about prescribing the pill. It's about providing mental health, actual treatment to coincide with other therapy, and I think it should be a huge part of the recovery process.	Lack of adequate understanding and support for people with aphasia struggling with mental health issues	
SLT3	[Interviewer: For those who instead start the program, have you ever noticed a particular challenge for those clients?] Yes absolutely. I would say that oftentimes PWA are also doing other forms of therapy - physiotherapy, occupational therapy - so even if it's able to be offered [they have to] squeeze it in and prioritizing.	Being enrolled in multiple therapy programs as a barrier to adequately attend Language Therapy	Miscellanea on negative effects of being enrolled in other therapies
	I would say the other challenge... is issues that they have with other therapies, or accessing services and that can deter the traditional Speech Language sessions, but it's more so spending a lot of time accessing those other services, advocating for the clients - that can deter the traditional Speech Language Therapy.	Difficulties experienced with other therapies negatively impact Speech Language Therapy	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT1	[Interviewer: What challenges do people with aphasia face once they have completed therapy?] I think it depends if there's other areas that have been affected, like clients who are ambulatory, and who prior to their stroke were quite active. I find that their recovery back into their routine seems to have been less problematic than those who continue to present with hemiplegia or hemiparesis… So, for example, in the past, we have had clients who have relied on either a wheelchair or they use a walker or a cane, and I found that there seems to be more challenges in their was their expressive language. There's that component when you think of, it depends on also being independently mobile...being able to go where you want, which is what even all my current adult stroke group members are able to do. One of them has a significant motor speech issue, yet because he is ambulatory and he drives, I find that, he is more willing, even though he's got a significant motor-speech issue where he has to use some compensatory strategies. I find that he appears to be more motivated to try strategies of psychotherapy than maybe a client who has to rely on something else or if there's…	Presence of motor disorders affecting ability to walk/drive independently as detrimental for motivation to engage with language therapy	Miscellanea
	One of the biggest indicators of success we know is if our clients are following up with recommendations or strategies that they can practice in succession and they come back, and they're not seeing any sort of progress that they just don't seem to sort of understand that their recovery is specifically based on how often they see me; it's everything that's happening when they are not here. I think a lot of the times, one of the first discussions that I have with a lot of my location clients is what the expectations are because, of course, they wanna know "Am I going to be able to talk again, am I going to be able to go back to work, how long will I be coming for therapy?". Concern for clients who have been followed for a long time and you are planning on discharging is "I don't think I'm ready, I need to come and still see you." So, they don't see the progress, sometimes, that they are making, and they feel that they are kind of being let loose, so it's kind of a double-edged sword. You sort of have those clients that want the service but don't realize that they're involved in the recovery side of the therapy.	Lack of active engagement with therapy as a barrier to successful recovery	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
	...on the other side of the spectrum, those that are self-motivated and so eager that when you let them know that "I think we've kind of reached the end", they start to panic, and they can't sort of see themselves being independent communicators without having a lifeline connected to you, even though you know that they would be able to manage.	Difficulty embracing independency when discharged from therapy	
	Something else to consider too is that [in] the communities that I have there is a literacy issue. You now have to put into consideration a lot of the therapy resources that at most clinicians just have available basically assuming that the patient has relatively adequate literacy skills, which is not the case. There is going to be a lot of people who will not admit that they have literacy issues for fear of stigma or embarrassment. So, that's another, that's something else that we have to really be considerate of as well very sensitive towards that because that assumption could also kind of blend into a person's perspective of the trust towards the clinician too.	Illiteracy as an underestimated factor that needs to be accounted when planning therapy	
SLT2	[Interviewer: What challenges do people with aphasia experience after they have completed therapy?] What I hear from clients when sometimes we touch-base for like a year is that their social networks have shrunk, or a lot smaller, so the visitors that came in the acute stage stopped coming, and they really have a much smaller circle.	Drastic reduction in social network as a common thing people with aphasia experience	
	I'm thinking functionally people seem to adjust to the other limitations, in terms of having other people take on daily living tasks for them like budgeting or chores or I don't know, groceries...those kind of things don't seem to be the complaints that they make…	Delegation of some daily tasks as a way to cope	
SLT3	...In the larger context of their lives - returning to activities, isolation that can exist, esteem to comeback to groups and networks with other disciplines to help support (like I said) those things that can impact their quality of life…	Greatest challenges brought about by aphasia	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
	[Interviewer: In your opinion, what are the main challenges that PWA, in general, still face - even after they finish therapy?] I would say that oftentimes it is rare to return to work, so possibly not return to work. Social isolation, change in relationship status (from partner to caregiver); the fact that even after therapy is done they might not be able to return home. So, those are the challenges that, even whatever stage of therapy is complete, individuals with aphasia still face at home.	Common permanent life changes experienced after therapy is terminated - join with clients	
	...In the larger context of their lives - returning to activities, isolation that can exist, esteem to comeback to groups and networks with other disciplines to help support (like I said) those things that can impact their quality of life…	Returning to activities, social isolation, low self-esteem as greatest challenges brought about by aphasia	

Table A.6*: SLTs views on appropriate approach to treatment*

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT1	The second thing is I think when I started out as a clinician there was a lot of... experience, but there seemed to be a lot of promotion towards a lot of therapy resources like the activity books and therapy material, which are all fine. I certainly have no objection to therapy books and… activities. But what I'm finding over the last few years where I really changed how I provide my service is I'm going directly for functional. So, finding tasks, communication goals that are meaningful, practical, and functional for the client because there doesn't seem to be as much progress in the patient's outcome in recovery when a lot of their time and energy is devoted to tasks that don't consequently relate to day-to-day activities.	Putting an emphasis on engaging clients on activities that are functionally meaningful	Client-centered approach
	[Interviewer: Do you assign some sort of homework that you have to do, between therapy sessions?] No, there's only been a handful of times where I've had to review a patient's rationale for coming to the appointments they weren't willing to follow-up, and with the home programming that I provide again I'm not sending home sheets of papers of those photocopied booklet. They have to identify what their goal is when they start therapy with me, so the first thing we do after an assessment, I ask them "What are your goals? What do you need to do?" And so, I kind of see myself as facilitating how they're going to achieve that goal. That may sound like I'm cruel, but if they're not doing the work, there's nothing more that I can really do for them. So with clients who have come to all of their appointments and who follow, these are the ones you know outside the therapy room are practicing and are trying, and there's like a few clients off the top of my head that have just created phenomenal augmentative communication aid based around their hobbies and interests. So, I really think that the effort put aside in the therapy context is really dependent on the individual's goal.	Emphasis on therapist as helping clients reach goals that are meaningful to them	
	They decide what they want to work on...it is [approach to therapy] directed on what the patient's goals are, not what my goals for them are...		

Subject	Meaning Unit	Condensed Meaning Unit	Theme
	I think the movement now is towards client-centered service as opposed to clinician-centered or clinician-focused service. I think you're going to be seeing that in a lot of different disciplines... the client has to go to the rehab and back to the home community.	Shift of focus to clients' needs with an emphasis on the role of community support	
SLT2	[Interviewer: Could you describe what criteria you use to assign homework?] It very much depends on whether they're needing to accomplish this independently or they have a supportive partner that can allow them to perform just a little, much higher than they could on their own, so it can be super simple or it can be a daily program which allows them to generally improve much quicker.	Criteria used to assign homework	
	[Interviewer: Do you think in general your clients do adhere to the homework you prescribe or just depends from client to client?] Yes, because I believe I put a lot of thought into preparing something that they're able to complete with their social situation and their skill level...but if it was only a homework package that everyone received, then I can see that a large number of clients would not complete it because of the barriers that they have in their situations...So I guess it's hard to assign homework.	Adherence to homework is contingent on how well it is tailored to a client's needs	
	So with the client and the family it can take a session or two in terms of what they would like to work on and break that down into something dependable. Then we look at targeting activities to reach those goals and that does the very best for clients, and we try to include a whole programming component so even if they are inpatient in restoring care unit, they have things to do when they are not with me. We also have technology that I can use in sessions, I have an iPad and an android tablet that have some therapy apps on and for restorative care, they have some iPads that patients can use in their rooms so they can use that as their whole program. I would say I don't know like fifty, it's broad range, but 50% to 70% percent of clients now have access to some kind of technology in tablet form, which makes the whole programming a lot easier… So the harder part is getting the clients focused on a measurable goal.	Setting up goals with clients and their caregivers, and using technology to help with	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT3	[Interviewer: In general, how well do people adhere to the prescribed homework?] Yes, people do it here, it is well tailored to their needs and interests and if they have the support, somebody to help, if not, it is more difficult. [For example] if they are able to use apps etc. that they are able to do independently and get that feedback, it works better; but if they are not a great therapy candidate, for whatever reason, we also use volunteers if there is not support system at home, so they are able to practice.	Adherence to homework is contingent on how well it is tailored to a client's needs	
SLT1	[Interviewer: What do you think are the most important aspects of speech-language therapy that help to promote aphasia recovery?] I think number one is… support from and collaborative team-up, which is huge because a lot of success in therapy and recovery kind of starts out at the hospital level. That's number one. You want to make sure you have collaborative services provided. You also need to provide education in training to any caregiver who is going to be in the client's care. So that could be a family member [or] a nurse. That is exceptionally critical.	Importance of adopting a collaborative service delivery model that considers caregivers and nurses as active members of the process right from the onset of aphasia	Collaborative service model as most appropriate to deliver treatment
	[Interviewer: Is there any strategy that you consistently try first with someone with aphasia before giving them the kind of the therapy?] Generally, when we do caregiver training, we need to make sure that the caregivers are involved early on in the treatment process because you need to make sure that the client has a communication partner who knows how to implement strategies to support communication. So, you need to look at caregiver training and education. Hopefully, there is a caregiver, whether it's a spouse, a brother, whatever, that's going to be involved in the client's care.	Importance of having the caregivers being involved early on in the treatment process	
	Then you need a service delivery model that allows you to do community sports and to include community partners or family partners in the process and not to just have an in-office kind of approach.	Collaborative service delivery model to promote sustained and meaningful recovery	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT1	The first thing that I do is self-anchored rating scale, with the clients and communication support (if they have one). Then through the self-anchored scale we develop functional goals and then I would assess those goals - i.e. if the goal is to read a book I would assess their reading skills, and then develop a plan based on that. That's typically what I do when I first meet a client with aphasia. Before meeting a client, we also try to offer a communication support person, so that would be the first step, and then I would do what I just described.	Approach adopted to plan therapy: through the self-anchored scale we develop functional goals and then I would assess those goals	Specific strategies adopted to plan/ deliver therapy
SLT2	Number two, before we even go into the specific therapy, you want to be able to determine what strategy the patient is able to use. So, we look at multimodal communication. So, we look at verbal, or "Are they able to gesture? Do they have any written communication intact? Written communication skills? Are they able to scan a letter board? Do they have a history of being familiar with technology? If so, do they own a smartphone or an iPad? Have they retained any of the skills to navigate those?". So, you're sort of becoming a little bit of a detective at first to find out which strategy.. because if communication breaks down, if they're trying to speak verbally and the message isn't getting across, you need to teach compensatory strategy they can go to. So, that's kind of a big part. And then, you do some training. If maybe they have some gestural skills, or very general gestures, maybe one of the goals might be to teach them how they can use more specific gestures to augment their verbal, so they were relaying the message to the class following their communication partner.	You want to be able to determine what strategy the patient is able to use	
SLT3	[Interviewer: Are there any standard treatment approaches that you would try first with people with aphasia?] It depends on the level of severity. I would always try, naming first. But it really would depend on the goal that we develop through the self-anchored scale, the initial session, to see what I would target, but the treatment that I would do are based on the goal that they would have. If I had somebody with severe fluent aphasia I would do first awareness, really targeting the awareness of their error first. It just depends, I would not always try Melodic Intonation with somebody with a more nonfluent type.	Criteria adopted to decide what treatment to try first	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT3	So, basically, our mandate is to focus on more functional type goals, but depending on what their goal is then I would try specific methods - so I could use parallel or reading, or I could use a card method if spelling was there. There is no package treatment that I would do with expressive vs receptive, just because everyone is so different.		

Table A.7: SLTs views on common goals of people with aphasia

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT2	[Interviewer: Any pattern in the type of goals that people with aphasia typically have for the treatment?] They have really generally things like “I want to talk again” and sometimes it can get specific like “I want to talk again on the phone to a certain person occasionally”, not often, people target reading and writing and they may be able to pinpoint certain things they want to understand like “I want to be able to listen to the news again”...“Or listen to more than one person talk at a time” so like “have a chat over coffee with three people or- ”...“Go to a family party and not be completely overwhelmed”... that’s the level usually.	Common goals of clients	
SLT3	Yes, it’s to talk better, usually it’s the verbal, I would say the main goal that individuals with aphasia have is to gather their verbal expression; the other main one that would be there is auditory comprehension. Those are the ones that would impact their quality of life the most.	Speaking and understanding as most common goals	
	Returning to reading, whether it is newspaper or book, reading is quite important for some people, or being able to read directions - those kinds of things that are required to return to drive. And then I would say that for writing, that could be emails, or texting, or if it’s higher levels it could be writing reports if their return to work is possible. So those are the main goals I would say.	Reading and writing as most common goals	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT2	[Interviewer: Have you found some factors like best motivators for people with aphasia to engage and persist in speech and language therapy?] In general people with aphasia come with their own motivation...It doesn't require a lot of tuning but what people sometimes underestimate [is] what they will be able to accomplish and maybe give up on themselves a little sooner that I might see. In terms of motivators, I find exposing themselves to different situations which they have specific facilitations to perform at a higher level, whether that's created in therapy sessions with a spouse or a son or daughter or whoever is in their family and facilitating a discussion in which they get to be a contributing network member and to show off how competent they are…That generally motivates them to somewhat participate and more hopefully in the therapy activities, it gives them the "this is why I'm doing this because I can do better".	Strategies to promote motivation in people aphasia	
	The other situation that tends to motivate people greatly is group sessions...They see other people struggling with the same things they are with various levels of competence and generally it can make them proud of what they are doing and also see that they can grow further. [Interviewer: So it's basically just being more confident about your skills and more aware of the fact that they can be partially at least in control of how well they can do...Not just a passive recipient of therapy but taking an active part in the process and empowering them..] Absolutely, that goal is possible and that they have the control over that..	The other situation that tends to motivate people greatly is group sessions… They see other people struggling with the same things they are with various levels of competence and generally it can make them proud of what they are doing and also see that they can grow further.	
SLT3	Active participation in the goal setting, and then re-evaluating through the course of Speech Language Therapy; and then tailoring activities in their life according to their interests, to make them as personal as possible while not being extremely time-consuming but something they can fit in their typical daily activities at home or in patient settings. Also, engaging family and friends is a motivator as well.	Approach to goal setting	

Table A.8: SLTs opinions on video games and ASR

Subject	Meaning Unit	Condensed Meaning Unit	Theme
SLT1	Technology would have to be readily accessible	Desirable features of technologies targeting aphasia	
	I think they have to be very, very easy to navigate. Very few kinds of buttons and whistles and what's not involved with it.	Desirable features of technologies targeting aphasia	
	The app has to be meaningful for individuals. So, I think it would have to be an app that can be individualized. Very similar to what the story is for my client. You've got that the foundation of what you put into it is your own	It would have to be an app that can be individualized	
	[Interviewer: What do you think of using a speech-recognition software in a videogame for speech language therapy?] If they're meaningful words and phrases to the individual, then possibly [it is beneficial]. If they're not I would question their relevance and use… For an individual who is working on improving their verbal expression, energy and time devoted to their therapy should be focusing on meaningful words and phrases that have application to their life… because what's relevant for someone who lives in the city where calling and finding out what the bus route schedule is or how much the bus fare is or something that's more relevant to urban setting might be completely different than someone who lives in a rural setting.		
SLT2	In terms of the feedback that clients are getting back on their production, in most apps, it's recording the production, playing it back and giving them a self-evaluation or having someone else evaluate it. The tablet itself is not evaluating the correctness of the production [Interviewer: Do you think it is beneficial if there was a speech recognition engine that can provide reliable and really accurate feedback eventually?] It will be more reliable and accurate.. the more independent we can make the clients the better, but they are reliable.	Automatic voice recognition systems should be reliable and accurate... the more independent we can make the clients the better	
	So the game would have to be a more every day kind of game, like Sims		

Subject	Meaning Unit	Condensed Meaning Unit	Theme
	[Interviewer: What would you not like about a videogame that could help with speech-language therapy?] I guess in terms of when I think of video games, like the shooting ones that my teenage sons play, other ones like that card games and candy crush, that people are engaging in and I guess among those all the things I'm seeing that would fit is possibly the alternative reality kind of games but it would have to be aphasia friendly and the task you are performing to advance in the game are things that are improving your language skills.	It would have to be aphasia friendly and the task you are performing to advance in the game are things that are improving your language	
SLT3	[Interviewer: Is there anything that you don't like about video games/apps?] Yes, I think it depends, it could lead to even more isolation, so you are not engaging with others. I think that if it wasn't based on their performance, so sort of adjusting accordingly, that would be a deterrent I guess. Or it would be a deterrent if they were continuing to try but were not advancing as it would impact self-esteem. There are definitely pros and cons, I don't think it should be the only form of therapy [video games] but it could definitely be in conjunction with other things.	Feedback and ability to adapt level of difficulty according to clients' changing skill level as a desirable feature of apps targeting aphasia	
	The apps that we use are the Tactus apps (comprehension, reading, naming). We've used Constant Therapy, that's more personalized. BrainHQ. So those would be the main treatment apps for language. They are more based for motor speech...I like that it is something that people can do independently; I like that they show how performance is going, especially for those that have issues with awareness; I like that, for Constant Therapy, it is more personalized and adjusts based on performance; and I like that you can adjust the settings.	Feedback, adaptivity, and ability of clients to be independent as positive features of apps used for therapy	
	[Interviewer: What are features that you would wish apps would improve on?] Tactus is based on either the Speech-Language Pathologist or the family member to adjust the level of difficulty, so it's not based on performance [it does not adjust automatically]..I don't like that feature.	Feedback and ability to adapt level of difficulty according to clients' changing skill level as a desirable feature of apps targeting aphasia	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
	So, the Tactus app was certainly more impairment-based, targeting naming. So even if they are using circumlocutions because there is no speech recognition, while I would say that Constant Therapy is more functional and tailored to activities of daily living and there is some speech recognition, so it is more representative of what I would do in traditional language therapy. But I think there is still a long way to go. What I would like to see is more integration with speech-recognition software, more real life scenarios in a game setting, for example: if you are practicing ordering...Also, in terms of receptive language, a lot of these programs are doing those sort of "path, directions" whereas there could be more voice messages, having to work on more functional auditory processing, comprehension tasks that are closer to real life. Those people with fluent aphasia do much better with tasks that are more functional, and I think that that's the biggest challenge, so I hope to see improvements.	What I would like to see is more integration with speech-recognition software, more real life scenarios in a game setting… having to work on more functional auditory processing, comprehension tasks that are closer to real life	
SLT1	[Interviewer: Do you think that a person with aphasia would play a video game if it could help with their speech-language therapy?] It's really based on their previous experience with technology… there's some ninety-year olds who have way better tech skills than me [laughs]. But you know again for the majority of clients that I see, there's really just a handful that seem to rely on using those devices. Interviewer: Do you think that a person with aphasia would play a video game if it this could help them with their recovery?] I think as the population ages you will have a generation that comes up where that would appeal to. Currently I don't have a fair number of clients that do not have technology in their lives, and we will engage with tablets in therapy sessions and even having technology become a that huge part of life, I don't know that the game format would appeal to population... I think it would still be a creative take to the drill, it would be more context and it could be more embedded in the context if you were doing comprehension instead of just finding words in a grid, and you look at the grid then you move on to producing words that are also in that topic, you might be talking about a realm of movies or something so that there's more context... I think that the apps that are out there now could have more context, you could embed the drill, there could be more conversation that could be used and I suppose you could define that as a video game.	Familiarity with technology is important to embrace computer-based therapy for aphasia	

Table A.9: Main uses of devices

Subject	Device	Main Use
S1	PC	Email, Apps (BrainHQ, Lumosity)
S2		News, Email, Skype, Games, documentaries
S3		Browsing the web for news and online shopping; Store important documents
S5		Skype
S6		Online banking
S9		Read and write reports; games; google calendar; browsing web; social media
S1	Tablet	Email, Browsing Newspapers, games
S2		Browse web, email, skype
S3		Emails, games, keep up-to-date with sports
S5		Note-taking; browse web for news; email; games
S7		Social media; browse web; games; apps; youtube; skype
S8		Games; reading news; browsing web
S1	Smartphone	Email, calls, texting
S5		Speech-to-text; note-taking; browse web for news; email; games
S6		Email; Facebook; news; games; youtube
S9		Google calendar; emails; texting
S9	Wii	Games

Table A.10: Aspects that facilitate use of technologies

Subject	Device	Meaning Unit	Condensed Meaning Unit	Theme
S1	PC	I liked the computer before...now there's trouble with the computer...I mean I like the computer. [Interviewer: Is it your favourite?] I guess it, probably, [because] you can do more things because I liked computer games. You know BrainHQ...and I had another program too, which was on the computer, can't be on the iPad.	Possibility to perform more activities on computer	Miscellanea about positive aspects of computers
S2		[Interviewer: You mainly use computer and tablet. Which one do you prefer in general?] Computer. [Partner: I think because he used it more before his stroke and he is more familiar]	Familiarity	
S3		[Interviewer: What's your favorite of the devices you use?] I guess this [pointing at computer]. It's faster, and bigger, [and has more] capacity	Processing power	
S9		Now in the past two years I need to use my smartphone and computer for that scheduling piece [setup basketball matches], I've gotta be able to converse with people on their devices and I now use my smartphone to access google calendars, and then at that point our website is set up such that our schedule is live. It makes it much easier for me as well as for the basketball community.		
		When I have email or something like that...I would use many folders, basically it might be a folder for basketball for 2015, 2016… I have a folder for family. If you look at my email, whether it's at home or at work, I have many folders. So my favorite one is my computer rather than my smartphone, because [with] my phone is hard for me to move or make a new folder. I can do it, but it's hard because it takes many clicks, whereas when I do it online with my computer is much easier. But I use both. Through the week, I would say 50% of my work is with my computer.	Computer is very user-friendly compared to smartphone when the user has to create or navigate nested folders	

Subject	Device	Meaning Unit	Condensed Meaning Unit	Theme
S1	Smartphone	[Interviewer: Do you find the smartphone user-friendly?] Yes, more than the one I had before	User-friendly	Miscellanea about positive aspects of smartphone
		The phone helps a lot because they have call or display. Even with the cellphone, which you know it's extra, but I need it really for people to call back. I use that. And the answering machine too, because you can replay.	Usefulness of callback function	
S5		I use technology all the time and if I have problems explaining something I will write a note in an email and send it off because I can talk to my phone and I have to dictate...It is a long road. I use my iPhone now and talk and it types and it's been a blessing.	Use text-to-speech to overcome difficulties with writing	
		I have an iPhone and I was concerned about getting it because I didn't know about technology and my kids say "Don't get it because it's too expensive", but I got it and it is my saving grace. I can put my appointments in there and talk to text, I talk instead of write my texts. And whenever I am in a group and I have to text someone I do but it's broke, I don't know how to explain that. I love the talk to write my text.	Smartphone useful to take notes and send email through text-to-speech functionality	
		...and at a moments notice I can make an appointment and I can use it as a phone and it's...easy to use.	User-friendly	
		[Interviewer: What is your favourite device?] iPhone. I love it. Because it is me all the time and I can look up things...	Portable	
S6		[Interviewer: Of the devices that you use, which one do you prefer?] Smartphone. Because it's small, it comes with me, I've got it all the time. My daughter in Toronto, she'll send me a text, and look! it's there..	Portable	

Subject	Device	Meaning Unit	Condensed Meaning Unit	Theme
S9		Now in the past two years I need to use my smartphone and computer for that scheduling piece [setup basketball matches], I've gotta be able to converse with people on their devices and I now use my smartphone to access google calendars, and then at that point our website is set up such that our schedule is live. It makes it much easier for me as well as for the basketball community.	Smartphone easy to use for scheduling events	
		[Interviewer: Which device would you use to help with your aphasia?] I would say my smartphone would be the better solution because I could use whenever I am.	Portable	

Table A.11: Barriers to use of specific technologies

Subject	Device	Meaning Unit	Condensed Meaning Unit	Theme
S1	PC	I have a couple programs, one of which is BrainHQ and [the other] Lumosity, which [is] on the computer but now it's because I have trouble with that computer...		Miscellanea about barriers to use of computers
		I have a lot of trouble with the computer. How to figure out how to do things and sometimes if it's something that's...I'm not familiar with how to do what I want to do on the computer... I may even call it some of the problems too, but.. who knows if it's trouble that may have been initially caused by me, but I don't know.	Lack of familiarity with the computer	
		[Interviewer: Do you think the interfaces of the programs can be confusing?] They can be. They're unfamiliar, to follow the steps it's sometimes depending, I mean simple things I can do if I can read the directions, but if I have to figure out sometimes with that is complex on the computer, maybe not. And it might be just Windows. I might get a Mac… might be better because I have the iPad, so I find that better than the computer.	Difficulty executing complex tasks on computer	
		It's also the complex steps, which would be part of the aphasia too, which I probably could have figured out a lot more things.	Learning issues due to aphasia	
		[Interviewer: Do you find using the mouse helpful?] No.	Hard to control mouse	
S2		Difficulty...mouse. [Partner: He did at first [experience difficulty] but then our son got him a cordless one. With the cord there was no problem… cordless is good.]	Mouse as a barrier to use. Cordless mouse as an adaptation	
S2	Tablet	[Partner: It was difficult at first, touching it too fast or touching it too long. It's very sensitive]	High sensitivity of touchscreen	Miscellanea about barriers to use of tablets
		[Partner: This [referring to tablet] is newer, even I have trouble with the tablet, swiping too fast.]	High sensitivity of touchscreen	
S3		These iPods, iPads are really, really slow .		

Subject	Device	Meaning Unit	Condensed Meaning Unit	Theme
S7		[Interviewer: Do you find the iPad easy to use because of touchscreen?] Yes, but hard to open apps and navigate around. It's hard to type	Difficulty navigating and typewriting	
S1	Smartphone	[Interviewer: Is there something that makes the smartphone hard to use?] I think sometimes I just don't know how to do things...I think is just...me.	Learning issues due to aphasia	Miscellanea about barriers to use of smartphone
S8		Sometimes I can't touch it because, nothing comes up [referring to phone]. So I wait till [her daughter] gets off work, and tells me how to do my thing… but she has to come over… sometimes I click, and some button I did wrong.	Low-level mastery of device	
S6		[Interviewer: Do you prefer using the smartphone?] Yes, unless I have to pin-out.	Difficulty inserting numbers on smartphone	
S9		When I have email or something like that...I would use many folders, basically it might be a folder for basketball for 2015, 2016...I have a folder for family. If you look at my email, whether it's at home or at work, I have many folders. So my favorite one is my computer rather than my smartphone, because [with] my phone is hard for me to move or make a new folder. I can do it, but it's hard because it takes many clicks, whereas when I do it online with my computer is much easier. But I use both. Through the week, I would say 50% of my work is with my computer.	Hard to navigate nested folders	
S2		Difficulty...mouse. [Partner: He did at first [experience difficulty] but then our son got him a cordless one. With the cord there was no problem… cordless is good.]		
S6		[Interviewer: Is there anything you do to help with your reading? i.e. adapting size of the font..] I've got it [the font] biggest I can get...and I go like that [flips screen to enlarge text]	Strategy used to help reading on smartphone	

Subject	Device	Meaning Unit	Condensed Meaning Unit	Theme
S3		[Interviewer: Of the devices that you use (PC, laptop, tablet) do you find them difficult to use? Do you think there is some aspect that makes them difficult to use?] "The only thing I got is these addresses - .com, .ca, .net. Because most of these companies put on their website, and it is American. Everything is money because you gotta exchange every Canadian dollar to US dollar...they don't allow them to come to Canada. It took me about three weeks to find this stuff on this Canadian based stuff...the only thing about this technology, it takes me a little longer to understand it."	Difficulties with online shopping through foreign websites	Miscellanea
S5		To voice-type in a crowd it's not pretty, and I usually don't.	Crowded environment not conducive to use of speech-to-text function	
S8		[The therapist] she got me.. how to use the...Google! I can't use it now, I don't know what happened to it.	Struggle figuring out how to navigate web browser	
		[Interviewer: Is there any aspect of these devices that makes them difficult to use for you?] I have all kinds of problems… I always call my daughter.	Struggle learning how to use devices	
		If nothing else I use technology more than in the past - online dictionary, spell checkers - whereas before I was a very good speller, so there was very little that I had to do, whether it was reading on word or smartphone, but now I start a word knowing that if it's one syllable, is one thing, but if is 3 or 4 syllables that I'm not correct or sure about, I'll start it but then I look at the spell checker to help.	Online dictionary and spell-checker used as away to self-monitor	
S9		If there was a part of hardware or software that was not intuitive, than I would need to read before using it.		
		Chances are that if it was 3 or 4 pages or more, I probably think of trying to find a different app or software		

Table A.12: Activities stopped or reduced because of stroke/aphasia

Subject	Activity	Meaning Unit	Condensed Meaning Unit	Theme
S1	Curling	One of the co workers asked me if I wanted to go, so she would put me on her team. I didn't go because they do signals, you know. They ask you to sweep, and it was difficult before, but I could always figure out who was talking and who wasn't, but I wouldn't this time. And even though I was on my friend's team, it's different because you really have to know what they're saying. And people get cranky if they do...I kind of [lost interest in curling] because I would still have trouble with the strategy. I do watch it when it's on TV. I enjoy watching it. But sometimes, I can't figure out first what they say, what they're going to do. I do have trouble with the strategy... plan the game.	You really have to know what they're saying...I kind of lost interest in curling because I would still have trouble with the strategy... plan the game.	
S2	Golf	I used to golf a bit, but I'm really not that good at it...		Motor deficits as a barrier to the pursuit of past hobbies
	Video games Photography Fishing, Camping, Driving			
S6	Golf	I went golfing twice this year and… terrible! I don't have the swing.		
	Woodworking	[Referring to woodworking] I don't have the [mimes hammer motion], I can't hit hard. It used to be bam! and the nail is gone. Now it's [mimes hitting hammer multiple times]		
S7	Hunting			

Table A.13: Current hobbies and pastimes

Subject	Activity	Meaning Unit
S1	Weight-lifting	I learned a lot of the other weight- lifting and things like that, so I do different things now because of the stuff she [personal trainer] taught me, but I have been able to go to more classes now, so I go there a lot.
	Reading	I used to read a lot. I read again now but not as fast. I actually enjoy reading again because I didn't stop… it took so long to read a book. It would have to be really interesting, otherwise I wouldn't read it. So, I read more. I don't need to work, which I won't be doing anyway.
	Yoga; body flow; biking; swimming	
S2	Walking; Exercise; Running	
S3	Gardening; Woodworking	
S5	Running; Kayaking; Outdoor activities	I love to run and I love to go outside, even shoveling my driveway
S6	Gardening; Snow shoveling; Woodworking; Car cleaning; Drawing with charcoal	
S7	Playing cards (solitaire, cribbage, auction 45); listening to music; facebook; walking; swimming; kayak; fishing; monopoly; skip all; exercises at stroke club and Expressive Cafe (aphasia stroke club)	
S9	Woodworking	Before I had my stroke there was a fair amount of woodworking, and probably 6 months after it I started to work again with wood, I'm working quite well - I got my confidence back - because there was no physical issue with me.

Table A.14: Most popular video games

Subject	Video games
S1	BrainHQ; Lumosity; Pet Rescue; Candy Crush; Pyramid
S3	Deer Hunter 2014, Car Racing, Solitaire, Bubbles
S5	Sudoku; Memory Games; Scramble; Word Games
S9	Fortune

Table A.15: Familiarity with video games and opinion about desirable features

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S1	I actually like them a lot...I like that one [Lumosity]. Some of them are number-based. Some of them were difficult because I haven't used them for a bit because of that…if you get it wrong they let you know quickly. Sometimes with BrainHQ, I think they don't. You don't always know where you went wrong, but I think you get an opportunity again to try again. But I like to know if it's kind of where you went wrong, so you can fix it...you can fix it because you can do them over and over, but there's no point if you're going to be wrong all the time. But you don't know where, why you're wrong. You need to know where you're wrong.	Positive opinion of computer-based program offering real-time feedback on performance	Desirable features of video games targeting aphasia: adaptability, provision of feedback, variety of tasks/ challenges, being entertaining
	[Interviewer: Do you think the apps you are currently using are helping you?] Yes. For BrainHQ I think it's the variety. I think it has scientific proof that it actually does work, and I'm sure it does, really because I had noticed after I had been using it for a bit, I just really seem to be a little bit sharper with things. I think it really helped a lot really...I think it's engaging because you have to pay attention.	Programs designed to help with recovery of cognitive function should be challenging and offer a wide variety of tasks	
	I was in another study...and they have a brain exercise… it was challenging, it went up in levels and you went down if you got things wrong. And they had two different strategies...I found it very good	video games targeting aphasia should provide the right level of challenge and use multiple strategies to engage the user	
S3	Other than that, the harder it is, the better it is. You know? Gotta have some fun too.	Desirable feature of game targeting aphasia - right level of challenge, fun	
S5	..they keep track of it and it increases your levels automatically, if you get through something one time they increase it and you don't have to do anything, you don't have to even increase it yourself.	Desirable feature of video game targeting aphasia - adjust difficulty level according to performance	
	I use aphasia applications but we have to pay for it to get the advanced. I can do all the things no problem but I have to pay for it to get the advanced version and I think it's a [incomprehensible], I don't know but I would love to find one that I do not have to pay for that is on my level.	Desire to find application that allows to adapt level of difficulty	

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S9	We have a game called Fortune, I play it with my son and my wife, it works for me because I have to think about the words, and sound them out. I'm getting better at it, and because it's something you can do with family members, but you can also do it on your own.	Positive aspects of video game – play in group with family; engages with a task that has right level of challenging for a person with aphasia	
S5	I do a lot of games like memory games, I [played] bridge before I had my stroke. It requires a lot of memory. I'm working on that and sudoku, I didn't ever play sudoku before the stroke but I find it helps with my numbers, but the games that are really for the aphasia, it's quite expensive and I don't know if I, I had a game on there and I used it's really good, but it's expensive to renew and my situation is different because, I have three kids that are in college and I have no money. I use to be a teacher and it's cut in half...I have to get download free applications.	The games that are really for aphasia ...they are expensive... I have to download free applications.	Financial barriers to access applications to help with aphasia
	I use aphasia applications but we have to pay for itto get the advanced. I can do all the things no problem but I have to pay for it to get the advanced version and I think it's a [incomprehensible], I don't know but I would love to find one that I do not have to pay for that is on my level.	I would love to find one that I do not have to pay for that is on my level.	
	I think I spend a half an hour a day...but I would like to spend more if the games were educational or for aphasia, I would love to release games, a list of games, that I could use. Because I find the games...I feel guilty playing games because it's games. I would feel better if it was [incomprehensible] or if I thought it was good for me.	Desire to engage with games that provide experiences that are enriching, other than engaging	Desire to engage with games that provide experiences that are enriching, other than engaging
	[Interviewer: When you said you'd like educational games, would you like something that teaches you about not only how to do exercise but also tells you something about aphasia, or you learn something more about aphasia while playing?] I think not aphasia, because I think I know enough about aphasia…	Desire to engage with games that provide experiences that are enriching, other than engaging	
	[Interviewer: Something that is more to your interest like something that you would like to know more about, could be sports, movies, art, anything related to your interests?] Yes.		

Subject	Meaning Unit	Condensed Meaning Unit	Theme
S5	I do a lot of games like memory games, I [played] bridge before I had my stroke. It requires a lot of memory. I'm working on that and sudoku, I didn't ever play sudoku before the stroke but I find it helps with my numbers...	Passion for games that engage memory	
	I mainly use the tablet for games because it has a bigger screen [than the iPhone] but I just download sudoku and card games and word games and word scramble, word games.	Passion for games that engage memory	
S2	[Interviewer: Is there something that you would like a video game for helping with aphasia to have?] Yes. [Partner: I am looking at the more recent ones, when I was searching for something I could read to him and he could follow, and the one thing that I found is their accent, it is difficult to understand. It is either too fast, we just can't find the right.. something suitable, I go to British English accent but it's strong English accent. The American is different. I think because they don't have a Canadian One, I don't know. That's what I found….None was quite suitable so that he can.. later he would probably be OK but starting off, it is difficult.]	Desirable features of video game targeting aphasia - option to use Canadian accent	Miscellanea
S3	The only thing I need, and that applies to everything, is gotta be based on one hand.	Desirable feature of video game targeting aphasia - user-friendly to people suffering from hemiparesis	
S9	We have a game called Fortune, I play it with my son and my wife, it works for me because I have to think about the words, and sound them out. I'm getting better at it, and because it's something you can do with family members, but you can also do it on your own.	Positive aspects of video game – play in group with family; engages with a task that has right level of challenging for a person with aphasia	

Table A.16. Summary of User Needs identified from literature review and from interviews

ID	User Need
1	The program should work on any device (Swales et al., 2016)
2	The program should be easy to use with the non-dominant hand (Greig et al., 2008; Palmer et al., 2013; Swales et al., 2016)
3	The program should be easy to use with mouse or via touchscreen (Greig et al., 2008; Palmer et al., 2013; Swales et al., 2016)
4	The program should not require users to execute fine movements (Greig et al., 2008; Palmer et al., 2013; Swales et al., 2016)
5	Users should be able to navigate the program in order to conduct rehabilitation sessions (Swales et al., 2016)
6	Tasks should be simple and require few steps to be completed (Greig et al., 2008; Swales et al., 2016)
7	The program should be easy to learn and play (Swales et al., 2016)
8	The User Interface should have the least possible number of buttons (Greig et al., 2008)
9	The User Interface should make minimal use of text (Greig et al., 2008)
10	Maximize consistency in the User Interface when releasing new versions (Brandeburg et al., 2017)
11	Avoid popup notifications (Brandeburg et al., 2017)
12	Users should easily get access to support on how to use the program if needed (Brandeburg et al., 2017; Greig et al., 2008; Hill & Breslin, 2016; Palmer et al., 2013)
13	The program should provide clear, simple and concise instructions to the user (Brandeburg et al., 2017; Greig et al., 2008; Palmer et al., 2013; Swales et al., 2016)
14	Users should be able to track their progress across sessions (Swales et al., 2016)
15	Users should receive immediate feedback on the accuracy of their naming responses (Finch & Hill, 2014a; Galliers et al., 2011; Hill & Breslin, 2016)
16	All information that is recorded (i.e. user performance, usage statistics, audio recordings etc.) should be stored in a personal account whose access is restricted to the associated user (Swales et al., 2016)
17	The program should provide a variety of stimuli and tasks (Finch & Hill, 2014a; Hill & Breslin, 2016; Palmer et al., 2013)
18	The program should adapt the level of difficult to user's skill level (Finch & Hill, 2014a; Hill & Breslin, 2016; Palmer et al., 2013; Swales et al., 2016)
19	Stimuli should be salient (Swales et al., 2016)
20	Tasks should be challenging (Finch & Hill, 2014a; Hill & Breslin, 2016; Palmer et al., 2013; Swales et al., 2016)

APPENDIX B

Table B. 1 List of software requirements and associated user needs

SWR ID	Software System/Domain	Software Requirement	Associated User Need ID
1		Webapp must run on any browser and device — Build Unity project using the WebGL (html5) option	1, 2
2	Backend	Personal accounts should store information about number of training sessions completed, and item-level information on accuracy, number of attempts, whether/which cue was used to support correct naming	15
3	Backend	Every user should have a personal folder to store the pictures that will be presented during the confrontational naming task, and associated audio recordings. The folder structure should be organised by '*subject_ID/session_number/type_of_stimuli*'	15
4	Frontend/Stage 2	Users must access their personal account via personal credentials	
5	Frontend/Stage 3	Before starting the naming task, users are provided with an overview of the items that they will name	
6	Frontend/Stage 4	The confrontational naming task should be supported by three types of cues (semantic, orthographic, printed word).	
7	Frontend/Stage 4	Users' naming attempts should be recorded and stored in their personal folder ('*subject_ID/session_number/audio_recordings'*).	
8	Frontend/Stage 4	During the confrontational naming task, users should have the option to listen to their last attempt at naming a word.	

SWR ID	Software System/Domain	Software Requirement	Associated User Need ID
9	Frontend/Stage 4	During the naming task, users should have the opportunity to move back and forth the list of items and retry specific ones	
10	Frontend/Stage 4	Connect ASR to webapp. Every time a naming attempt is recorded, it should be sent to the ASR for processing, and the result sent back to the webapp for display.	13
11	Frontend/Stage 4	A progress bar informs users about how many words they have named correctly, based on the judgement of the ASR	13
12	Frontend/Stage 5	Once users have completed the first round of naming for all picture stimuli, they should be presented an overview of item-level performance. Users should have the opportunity to try again on any of the items.	13
13	Frontend/Stage 6	At the end of the training session, participants should be presented an overview of the progress they made throughout all session they engaged in (i.e., number of correct words per session)	13
14	Frontend/All Stages	Every stage of the program should have an associated video tutorial	11, 12
15	Security	Passwords should be hashed and salted	
16	Security	Implement https protocol to handle communication between web browser and webserver	
17	Security	Remote access to webserver requires SSH keys	

Background Questionnaire

Section 1: Background Information

1. How old are you?

2. Is English your first language? If not, what is your first language?

   Yes No

3. When did aphasia occur?

a. Use Year Scale, if less than 1 year, then use Month Scale.

Section 2: Experience with Technology

4. Do you use a tablet or computer on a daily basis?

   Yes No

5. What do you use a computer/tablet for?

a. Browsing the Web

b. News

c. Email

d. Skype/video chat

e. Games

f. YouTube or other videos

g. Social Media

h. Entertainment (movies/shows/music)

Others:

6. What type of device do you prefer to use? Why?

7. Is there anything that makes it difficult for you to use any of these devices? If yes, what?

Section 3: computer-based therapies

8. Have you ever used online therapies to help with aphasia?

Yes No

9. Are you currently using any computer-based therapy for aphasia?

a) If yes, which ones?

Yes No

b) Do you think it is helping?

Yes No

c) What do you like most about using computer-based therapy?

d) Are there any aspects that you would like to improve?

Usability Questionnaire

1. How would you **judge** the **instructions** in the video?

Not clear **Neutral** **Clear**

2. How would you **judge** the **instructions** in the video?

Not helpful **Neutral** **Helpful**

3. What would you improve in the instructor video?

4. How would you describe your **experience** using the program?

Boring **Neutral** **Engaging**

5. Would you be **interested** in adopting improved versions of this program to do therapy at home?

Yes No

6. How did you find **using** the program?

Too Easy **My level** **Too Hard**

7. How did you find the **quality** of the **audio**?

Not good **Neutral** **Good**

8. How did you find the **quality** of the **images**?

Not good **Neutral** **Good**

9. How did you find to **log** into the **program**?

Easy **My level** **Too Hard**

10. In case you found it **hard**, **why**?

Too many pictures

Pictures were too small

Other:

---

11. How did you find the scene "Naming Scene" (shown below)?

Easy to navigate **Neutral** **Difficult to navigate**

12. If you found it **hard**, **why**?

Too many icons on the screen

Unclear what icons refer to

Other:

---

13. Did you **feel insecure** or **discouraged** at any point? If yes, why?

Yes No

---

14. How did you find the **feedback** on your vocal responses?

Not Accurate **Neutral** **Accurate**

15. How did you find the **feedback** on your vocal responses?

NotUseful **Neutral** **Useful**

System Usability Scale

For each of the following statements, please **mark one box** that **best describes your experience** using the program today.

1. I think I would use the program frequently

Strongly disagree		Neutral		Strongly agree

2. I found the program **unnecessarily complex**

Strongly disagree		Neutral		Strongly agree

3. I thought the program was **easy to use**

Strongly disagree		Neutral		Strongly agree

4. I think that I would **need** the **support of** a **technical person** to be able to use the program

Strongly disagree		Neutral		Strongly agree

5. I found the various **functions** in the program were **well integrated**

Strongly disagree **Neutral** **Strongly agree**

6. I thought **there was** too much **inconsistency** in the program

Strongly disagree **Neutral** **Strongly agree**

7. I would imagine that most **people** would **learn to use** the program very **quickly**

Strongly disagree **Neutral** **Strongly agree**

8. I found the **program** very **cumbersome** to **use**

Strongly disagree **Neutral** **Strongly agree**

9. **I felt** very **confident** using the program

Strongly disagree **Neutral** **Strongly agree**

10. I needed to **learn** a **lot** of **things** before I could **get going** with the program

Strongly disagree **Neutral** **Strongly agree**

List of Tasks of Prototype v1.0

1. Access web and type the website address
2. Login into appropriate account
3. Get overview of training material
4. In Naming Task scene:

a. Start trial

b. Use cues for help

c. Use “Try Again” button

d. Use “Replay Audio” button

e. Move to next item

f. Exit task at completion

5. Review item-level performance
6. Get overview of scores throughout study sessions
7. Quit program

Appendix C – ASSESSMENT OF THE RELIABILITY OF THE AUTOMATIC SPEECH RECOGNITION ENGINE

On Kaldi, an Open-Source Framework to Train and Deploy ASR

For this project, we employed an ASR built using the Kaldi (Povey et al., 2011) open source framework. Briefly, Kaldi provides the entire pipeline required to train and evaluate an ASR, and in principle it requires users to provide only the data required to do so. Nonetheless, the framework is also versatile as it allows users to tweak some key parameters based on their needs (for example, during the process of feature extraction, the width of the temporal windows used to compute the Mel Frequency Cepstral Coefficients ("Mel-Frequency Cepstrum," n.d.) (MFCC) can be changed; different algorithms are available to build an acoustic model).

The Kaldi ASR framework is composed of three essential processing blocks (see Figure C.1 below for a graphical overview): one performing feature extraction, one embedding an acoustic model, one embedding a language model. Since it is out of the scope of this study to describe in detail all the steps that are necessary to successfully train an ASR, this section provides only a high-level description of how the model is trained and how it operates during classification, with the intent to give a general idea about how it works.

Nowadays, in the field of speech recognition, the most successful algorithms can easily classify the audio files associated with different phonemes by looking at a specific feature, the MFCC. To this end, the feature extraction process consists of taking the raw audio as input, breaking it down into a series of adjacent, and partially overlapping, time-windows, and returning the corresponding series of MFCC as output.

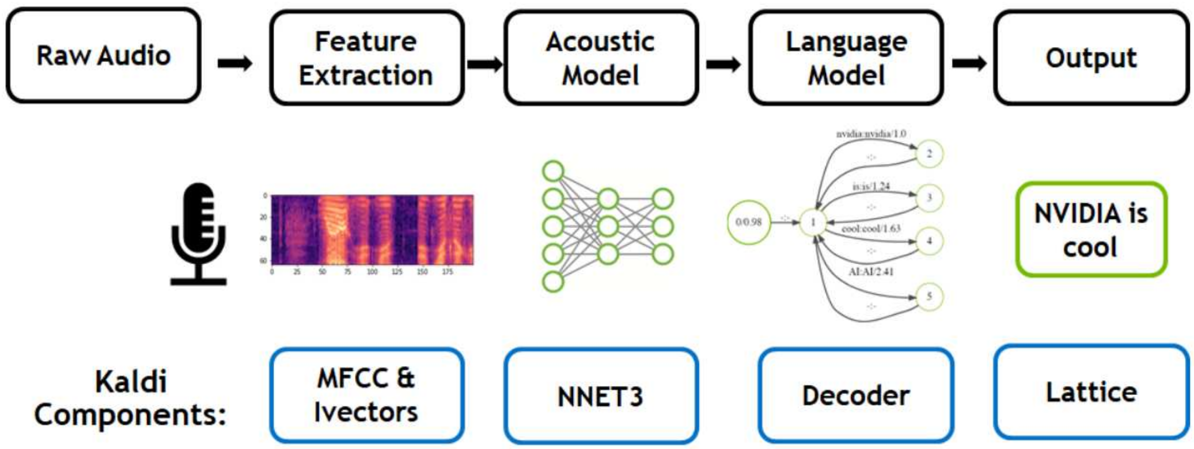

Figure C.1: Schematic overview of the main processing blocks of the Kaldi ASR framework (*NVIDIA Accelerates Real Time Speech to Text Transcription 3500x with Kaldi*, n.d.)*: above, in sequence, are listed the general steps; below are the corresponding tools used to implement each step. NNET3 is a particular Deep Learning architecture.*

The next task is to convert the series of MFCC into the corresponding series of phonemes. This can be done by building an acoustic model for the target language - a model that encodes information about how the sound properties of each phoneme of a language (i.e. American English in our case) map into their corresponding MFCC. In our case, the acoustic model was built by training a Deep Neural Network employing a Long Short-Term Memory (LSTM) architecture (Graves, 2012; Hochreiter & Schmidhuber, 1997). This has proven to be one of the most successful algorithms to model data, such as speech, that have a sequential structure (i.e. where the state of a signal at time *t* is influenced by the sequence of preceding states). During training, this model takes as input the series of MFCC obtained during the feature extraction process and attempts to output the associated string of phonemes. Naturally, at the beginning of the training phase, when the model has no knowledge about the relationship between MFCC and phonemes, the output will show little resemblance with the correct transcription. To allow the model to learn, in the training phase it is provided a detailed transcription of the audio file (the labels), along with information about when the speaker started and finished speaking. Based on the comparison between the desired output and the actual transcription, an error signal is generated and used to drive the update of the model parameters in order to attain a better transcription during subsequent attempts. This process is repeated, during the training phase, until the model does achieve satisfactory performance.

During the validation phase, where the model is given out-of-sample data – audio files and associated transcriptions that were not processed during the training phase – the input series of MFCC is converted into a series of phonemes. This is then further converted into a string of words by identifying the timing of pauses and by relying on knowledge about the lexicon.

Whenever the goal is to transcribe not just single words, but sentences, it proves useful to not only rely on an acoustic model to return strings of words, but also on language models, which encode information about the transition probabilities in chain of words (i.e., which word is likely to come next given knowledge of the preceding words.) Naturally, performing speech recognition in a context where the speakers engage with an open-ended conversation is very difficult, as the space of the possible correct mappings between an audio source and its transcription is infinite (although some combinations of words are more likely to occur than others). Nonetheless, in a naming task, where the goal is to determine whether a person is saying a given word or not, the ASR is not required to search the entire space of possibilities to determine whether the target word was pronounced. It can, instead, simply look at a model that is limited to the target word and a few other words that are likely to be said.

In this study, we exploited this scenario and restricted the space of possible output to three words, including the target one. The output of the ASR is considered to be correct if it matches with the target word, and incorrect otherwise.

Preliminary Study on the Accuracy of the Voice Recognition System

Since we have no information regarding how accurately the ASR can recognize vocal input from people with communication disorders, we conducted a preliminary analysis to test this aspect. To this end we first collected voice samples from people with aphasia, and asked two native English speakers to transcribe them. We then used these transcriptions as a benchmark for assessing the reliability of the ASR. Below are details about the methods that we employed to conduct this investigation.

METHODS

Participants

Six people with aphasia took part in this study. Four of them visited our lab, while the other two, who lived outside the Halifax Metropolitan Area and could not drive to town, did the task from their house. This was made possible by the fact that this study only required the availability of a computer and an internet connection.

Stimuli

We initially selected 500 images depicting animate or inanimate objects. Pictures were taken from the Bank of Standardized Stimuli (BOSS) dataset (Brodeur, n.d.; Brodeur et al., 2014) according to the following selection criteria: words had to be no longer than 4 syllables, with frequency (expressed as 1+ base 10 logarithm of occurrence per million words in the HAL database (Balota et al., 2007) not lower than 1.5. We further filtered out words related to fictional characters (i.e. wizard, pegasus, minotaur, werewolf etc.). This resulted in a reduced set of 434 words. To minimize the length of the voice collection process we created two sets of words, one with 230 and one with 204 words. Four participants were assigned to the first block, 2 to the second.

Task

Every trial (one per picture) started with the appearance of a picture at the center of the screen. Participants were asked to name it and to press a button when they were ready to move to the next one. A one second delay with a blank screen intervened between the keystroke and the onset of the next picture.

Vocal responses were recorded through a microphone, which was automatically activated upon the onset of the picture and deactivated at the end of the trial (i.e. either upon subjects signaling the intention to move on by pressing the appropriate button, or automatically after 10 seconds if no response was detected). The corresponding audio file was saved and tagged with the name of the target word.

Every participant was associated with a unique user account and anonymized ID created prior to the start of data collection. This was necessary in order to track which utterances were produced by each participant.

Data analysis

Two native English speakers transcribed the audio files, while blinded to the target item. A member of the research team subsequently labelled each transcription according to whether it matched the target word, or not. Therefore, even if raters provided two different transcriptions for a given audio file, they would be associated with the same label if they both did not match with the target. We made this decision because, for the sake of the preliminary studies proposed for this thesis, the aim is to provide binary feedback (correct/incorrect) about response accuracy, and not also about what specific (if any) part of the utterance was pronounced incorrectly.

After we created the binary accuracy responses, we quantified the extent to which raters agreed by employing the Inter-Rater Reliability (IRR) metric, defined as the percentage of responses for which the accuracy scores derived from the transcriptions of both raters matched. We used this subset of responses as the ground truth to compare the performance of the ASR against.

The performance of the ASRs was measured by employing signal detection theory (Donaldson, 1992) using the *psycho* package in R (Makowski, 2018). Briefly, this method measures both the extent to which the ASR is capable of discriminating between correct and incorrect responses, as measured by the parameter a' (a-prime), and the extent to which it might be biased towards a specific category of responses, as measured by the parameter β'' (beta prime prime d; Makowski, 2018). Essentially, the β'' parameter measures whether the likelihood (not the absolute number) of false positives and false negatives is the same. It is, therefore, particularly informative in cases where the two classes of responses that have to be evaluated by the ASR (correct and incorrect) are unbalanced (a scenario that is likely to occur throughout training sessions, as participants' performance improve and the proportion of correct responses will progressively outweigh that of incorrect ones).

Following Makowski (2018), values of a' around .5 indicate that the system performs at chance level. Conversely, the more a' approaches 1 (maximum possible value), the higher the ability of the system to discriminate between correct and incorrect responses. Values of β'' around 0 indicate no bias in response. Positive values indicate a bias towards scoring a response as incorrect, with a value of 1 indicating that all responses were scored as such. Vice versa for negative values.

RESULTS

The top two rows of Table C.1 present the percentage of raters' transcriptions (H1 and H2, respectively) that matched the target response. The third row shows the IRR between raters.

The last row indicates the number of responses for which both raters agreed; within parentheses is the number of these responses for which the ASR could return an output. As it turned out, some of the words that participants had to name were not included in the language model that the ASR was trained on. Therefore, these words could not be recognized, and we excluded them from further analysis.

In some cases, vocal responses could not be provided due to glitches that occurred during the experiment, where the pictures corresponding to the target words were not presented. For P5 we were able to collect only about a third of her responses: although he was able to complete the entire experiment, there were issues saving most of the audio files recorded during her session.

Broadly, the IRR, which is generally considered to reflect good agreement between raters when it is equal or greater than .75, was close or above this threshold for all participants except P2.

The ASR showed high levels of agreement with both raters, being above the .90 threshold for all six participants when comparing its performance with the ground truth.

Table C.2 presents subject-level confusion matrices. Overall, across subjects, the results of the analysis based on signal detection theory revealed an a' score of .96, indicating good ability to discriminate between correct and incorrect responses, and a β'' of -.57, indicating a moderate propensity to classify responses as correct.

DISCUSSION

This preliminary analysis shows promising results, as there was a high level of agreement between the output of the ASR and that of human raters. These results are in line with those of recent studies investigating the accuracy of novel ASR systems when tasked to classify pre-recorded datasets of single-words or short utterance recordings produced by people with aphasia

(Abad et al., 2013; Barbera et al., 2021; Le et al., 2018), or when tasked to provide real-time feedback on single word naming accuracy to people with aphasia using computer-based therapies to improve their ability to name words (Ballard et al., 2019).

Table C.1: Comparison between the performance of human raters with that of the ASR

	P1	**P2**	**P3**	**P4**	**P5**	**P6**
H1	.78	.50	.85	.88	.46	.79
H2	.92	.73	.94	.95	.55	.86
IRR	.81	.66	.84	.87	.78	.73
#items$_{H1=H2}$	181 (181)	150 (137)	190 (174)	196 (179)	59 (50)	150 (113)
ASR	.90	.93	.99	.99	.96	.98

Note. The first two rows indicate how many correct responses each participant provided, according to each human rater (H1 and H2, respectively). The third row indicates the Inter-Rater Reliability score, for each participant. The fourth row indicates the number of items that were scored, and within parentheses the number of items for which the ASR could return an output. The fifth row indicates the percentage agreement between the subset of responses for which human raters agreed and the corresponding output of the ASR.

All in all, the results of this investigation justify the integration of our ASR of choice in a prototype of computer-based therapy, with the aim to test the feasibility of enabling the provision of automatic feedback on vocal responses to assist people with aphasia with the recovery of their word finding and naming abilities.

Table C.2: Confusion matrices for the performance of the ASR relative to human raters

		Predict T	Predict F
S1	True T	154	16
	True F	1	10
S2	True T	89	3
	True F	7	38
S3	True T	168	1
	True F	1	4
S4	True T	174	1
	True F	1	3
S5	True T	27	1
	True F	1	21
S6	True T	105	1
	True F	1	6

Note. Green and red cells indicate, respectively, responses that were correctly and incorrectly labeled by the ASR.

Appendix D

Prototype v1.0 to v2.0: Summary of Changes

The first version of the prototype underwent four main kinds of changes: two scenes were removed; buttons were added, removed, or moved in different positions; video-tutorials were edited or fixed; the user database features additional fields to collect a wider spectrum of information on the user experience.

The two scenes that were removed were the ‘overview’ and the ‘review item-level performance’ scenes. Participants engaged with the overview scene immediately after the successful login into their account and prior to engaging with the naming task proper. The original purpose of this scene was to prime participants to the material that they would have engaged with. However, it was removed for two reasons: first, in the intended use case scenario participants are supposed to be already familiar with the training material; second, analysis of the effectiveness scores (see Table 3.3 in the previous chapter) revealed that participants with lower AQ scores did not find it immediately obvious what to do in that scene, nor why it was useful. The ‘review item-level performance’ scene occurred immediately after the end of the naming task proper, and was intended to give participants a chance to review which items they could never name correctly, and to focus on them before concluding the session. Nevertheless, since in the intended use case scenario participants are expected to cycle three times over the list of items associated with a particular training session (see Methods), we thought that an additional exposure was not necessary, especially in the interest of preventing training sessions from getting too long.

One participant recommended to align the design choices to those of most popular apps, in order to make the experience more consistent, and easier. Specifically, he asked to move the login button below the text fields associated with participants’ credentials. Another one requested to introduce a function that would allow users to play an audio recording of the target word made by a person without language disorders, for which a dedicated button was added to the user interface. The button used to move from the current trial to a previous one was removed, as it was redundant. The reason behind this choice was, again, that in the intended use case scenario participants are afforded multiple opportunities to name a given item. Furthermore, to minimize the number of buttons in the user interface, the list of tutorials was, by default, hidden behind a single button named 'video-tutorials'. In case a user would feel the need to watch a video-tutorial, he/she could reveal the entire list by simply clicking on this button.

A major change that was introduced was the complete redesign of the video tutorials. Where possible, the scripts were pruned and made more aphasia-friendly. Also, the videos were recorded to align with the new scripts.

Finally, as part of assessing what strategies users adopt when training with and without external guidance, it is valuable to collect information on the number and type of cues used during each trial. To do so, we expanded the range of information collected in the user database, adding a total of eight fields. Five fields were dedicated to record whether any of the five types of support afforded by the program: orthographic hint and semantic hint, along with how far in the hierarchy the user reached out; printed target word; listen to audio file of target word; and replay last recording of naming attempt. Three additional fields were added to keep track of the study session, session round (there were three rounds per session – see Methods), and how many times a given trial was repeated in a given round.

List of Tasks of Prototype v2.0

1. Access web and type the website address
2. Login into appropriate account
3. In Naming Task scene:
    a. Start trial
    b. Use cues for help
    c. Use “Try Again” button
    d. Use “Replay Audio” button
    e. Use “List to Target Word” button
    f. Move to next item
    g. Exit task at completion
4. Get overview of scores throughout study sessions
5. Quit program

Chapter 4 – Supplementary Figures

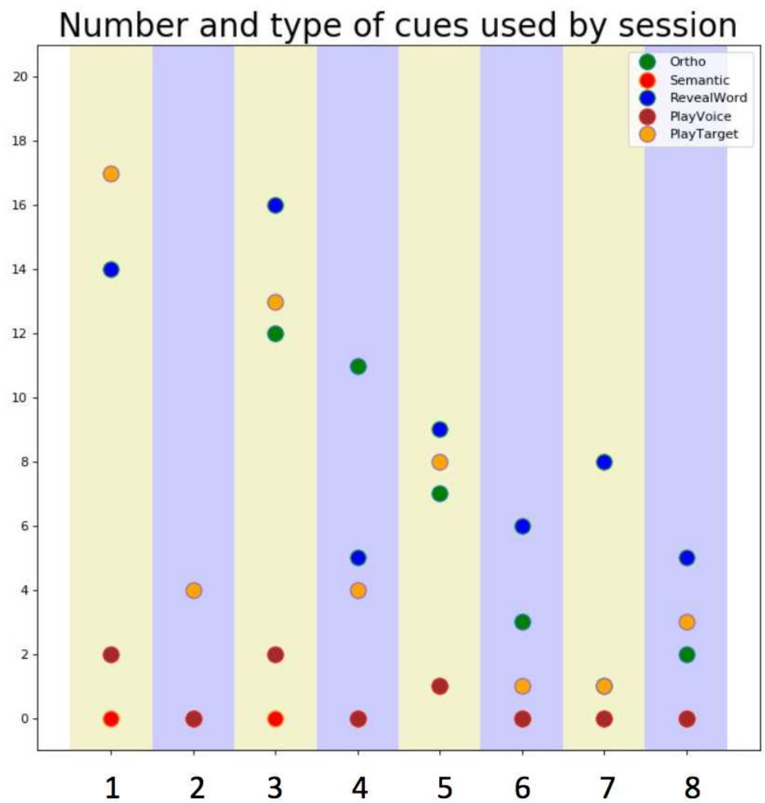

Figure D. 1 Cue usage of S2 throughout training sessions. Yellow background relates to the condition in which no feedback was provided, vice versa for the purple background.

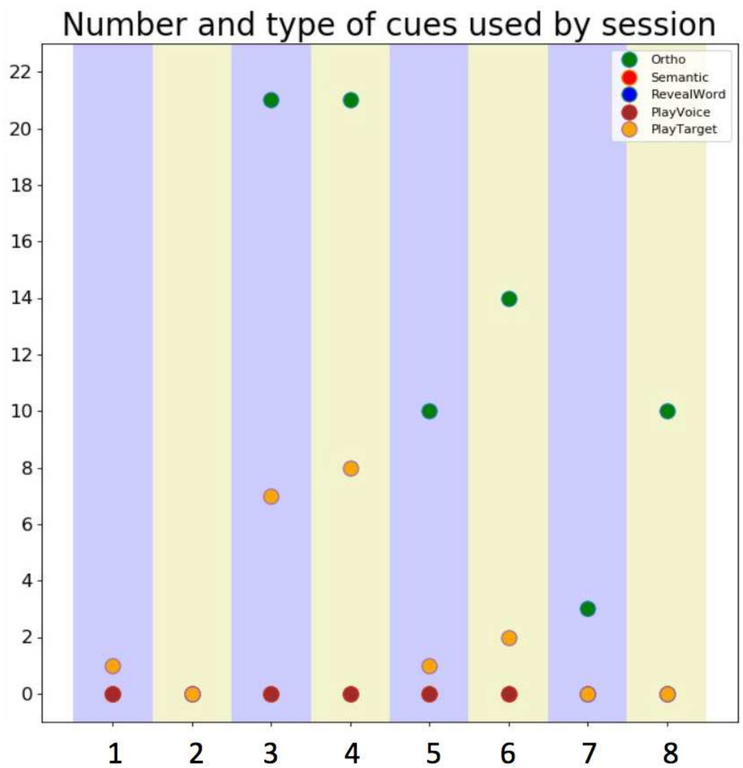

Figure D. 2 Cue usage of S5 throughout training sessions. Yellow background relates to the condition in which no feedback was provided, vice versa for the purple background.