frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 01 December 2023
DOI 10.3389/fcomm.2023.1210193

OPEN ACCESS

EDITED BY
Kristina Jonas,
University of Cologne, Germany

REVIEWED BY
Jean K. Gordon,
The University of Iowa, United States
Anna Rosenkranz,
University of Marburg, Germany

*CORRESPONDENCE
Judith Heide
✉ jheide@uni-potsdam.de

†These authors have contributed equally to this work and share first authorship

RECEIVED 21 April 2023
ACCEPTED 01 November 2023
PUBLISHED 01 December 2023

CITATION
Heide J, Netzebandt J, Ahrens S, Brüsch J, Saalfrank T and Schmitz-Antonischki D (2023) Improving lexical retrieval with LingoTalk: an app-based, self-administered treatment for clients with aphasia.
Front. Commun. 8:1210193.
doi: 10.3389/fcomm.2023.1210193

Improving lexical retrieval with LingoTalk: an app-based, self-administered treatment for clients with aphasia

Judith Heide[1*†], Jonka Netzebandt[2†], Stine Ahrens[1], Julia Brüsch[1], Teresa Saalfrank[1] and Dorit Schmitz-Antonischki[1]

[1]Department Linguistics, University of Potsdam, Potsdam, Germany, [2]Lingo Lab, Berlin, Germany

Introduction: LingoTalk is a German speech-language app designed to enhance lexical retrieval in individuals with aphasia. It incorporates automatic speech recognition (ASR) to provide therapist-independent feedback. The execution and effectiveness of a self-administered intervention with LingoTalk was explored in a case series study.

Methods: Three individuals with chronic aphasia participated in a highly individualized, supervised self-administered intervention lasting 3 weeks. The LingoTalk app closely monitored the frequency, intensity and progress of the intervention. Treatment efficacy was assessed using a multiple baseline design, examining both item-specific treatment effects and generalization to untreated items, an untreated task, and spontaneous speech.

Results: All participants successfully completed the intervention with LingoTalk, although one participant was not able to use the ASR feature. None of the participants fully adhered to the treatment protocol. All participants demonstrated significant and sustained improvement in the naming of practiced items, although there was limited evidence of generalization. Additionally, there was a slight reduction in word-finding difficulties during spontaneous speech.

Discussion: This small-scale study indicates that self-administered intervention with LingoTalk can improve oral naming of treated items. Thus, it has the potential to complement face-to-face speech-language therapy, such as within in a "flipped speech room" approach. The choice of feedback mode is discussed. Transparent progress monitoring of the intervention appears to positively influence patients' motivation.

KEYWORDS

aphasia, anomia, lexical retrieval, oral naming, app-based intervention, self-training, automatic speech recognition (ASR), LingoTalk

1 Word-finding disorders in aphasia

Aphasia is an acquired disorder of language processing that occurs after language acquisition has been completed. The most common cause of aphasia, responsible for more than for 80% of cases, is stroke (Engelter et al., 2006). While aphasia can affect all four modalities of language – reading, writing, oral production, and auditory language comprehension – the most prevalent symptoms involve difficulties in word retrieval (Goodglass and Wingfield, 1997; Nickels, 2002). The severity of the impairment can vary greatly, ranging from difficulties to formulate even single words to mild uncertainties in selecting the appropriate word. Word-finding disorders often lead to an enormously high level of despair: they severely limit everyday communication (Blom Johansson et al., 2012) and have a significant impact on the quality of life (Hilari et al., 2016). Word-finding disorders are observable in everyday conversation as well as in tasks requiring oral word production, such as picture naming, storytelling or word fluency.

Typical symptoms of word-finding difficulties include hesitations, the use of empty phrases, rephrasing or paraphrasing, as well as zero responses, phonological or semantic paraphasias or neologisms. Sometimes individuals with aphasia can retrieve partial information of a word, such as the number of syllables or the initial sound of a word (Goodglass et al., 1976; Anusuya and Shyamala, 2021). They might also have access to grammatical information, such as gender, without being able to retrieve the word form (Badecker et al., 1995).

Both neuro- and psycholinguistic research have developed models of language production that can explain word retrieval disorders in aphasia [see Nickels (2001) for a comprehensive discussion]. A fundamental distinction can be made between serial-modular and connectionist models. Connectionist models assume a continuous flow of information, leading to multiple processing steps being active in parallel and mutually influencing each other. This information flow can be either strictly feedforward (Plaut and Shallice, 1993a) or interactive, allowing feedback from later processing steps to earlier ones (Dell, 1986). On the other hand, serial-modular models propose autonomous modules that process incoming information independently and sequentially. For the neurolinguistic diagnosis of word retrieval disorders, serial-modular models are highly suitable as they allow for a very precise localization of the underlying impairment (Lorenz, 2004). Examples of serial-modular models include the Logogen model (Patterson et al., 1987) and Levelt's two-stage model (Levelt, 1993). These models differ in their stance on whether lexical access and phonological retrieval constitute a one- or two-step process. The Logogen model (Patterson et al., 1987) posits a direct link between a semantic concept and the corresponding phonological word form, rendering lexical access a one-step process. In contrast, Levelt's model (Levelt, 1993; Levelt et al., 1999), assumes two distinct steps within lexical access, distinguishing between accessing abstract lexical representations (lemmas) and the retrieval of the phonological word forms (lexemes). In the context of this study, we rely on the Logogen model (Patterson et al., 1987), which describes the oral and written production as well as auditory and visual comprehension of monomorphemic words and neologisms. Figure 1 illustrates the modules that are relevant to oral picture naming.

The recognition of an object leads to the activation of the item's semantic features in the semantic system. This semantic activation is forwarded to the corresponding entries in the phonological output lexicon (POL). The target word, which receives the highest activation, is then retrieved, while competing word forms are inhibited. In the next step, the target word is briefly stored in the phonological output buffer (POB) before it is articulated. The modules involved in lexical retrieval may be disrupted selectively or in combination (Hillis and Caramazza, 1994). Thus, impairments in oral word production can be due to semantic and/or post-semantic disorders (Ellis et al., 1992), namely

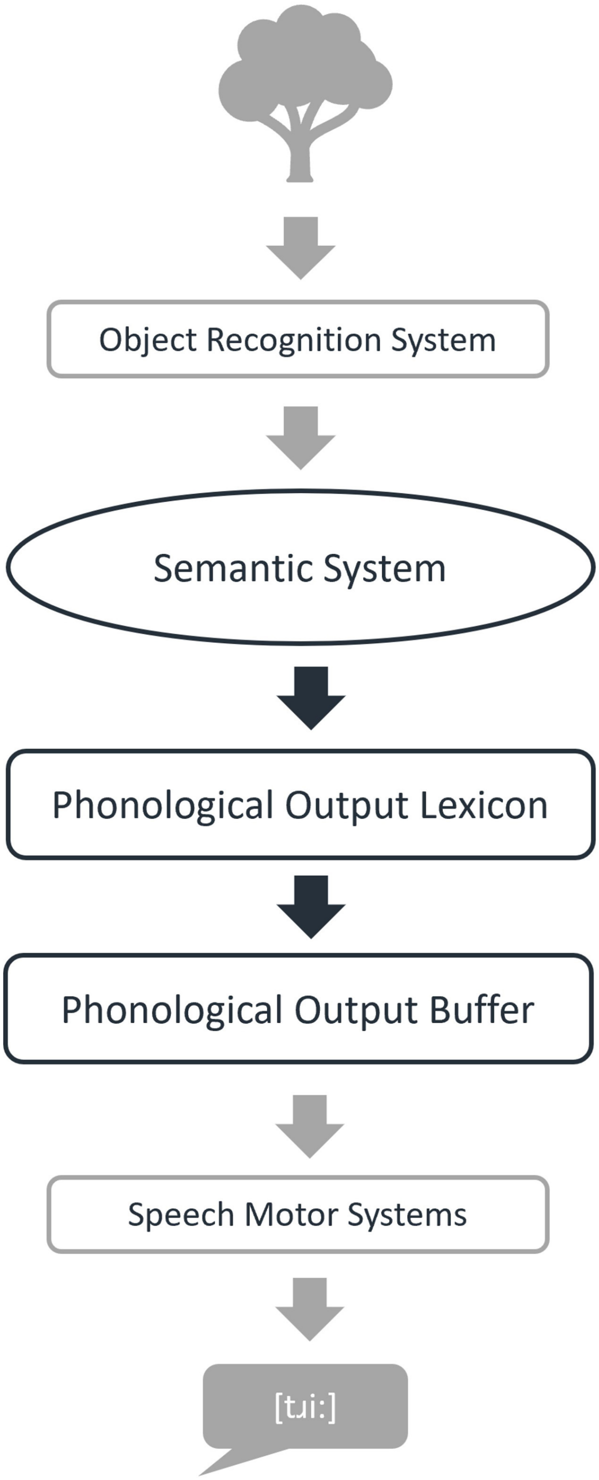

FIGURE 1
Lexical retrieval in oral picture naming according to the Logogen model (based on Patterson et al., 1987).

1. Faulty or insufficient activation of semantic knowledge.
2. Impaired access from the semantic system to the corresponding lexical entry in the POL.
3. Missing or underspecified lexical entries in the POL; i.e., deficient knowledge about the phonological word form.

4. A limited capacity of the POB which leads to deficits in maintenance and assembly of phonemes.

Various (psycho-)linguistic parameters are known to influence word production in aphasia, as well as in neurotypical individuals. For instance, concrete words and typical members of a semantic category are retrieved more easily than abstract words and atypical members. Both the concreteness effect (Plaut and Shallice, 1993b) and the typicality effect (McRae et al., 1997) are attributed to the word's semantic features and their representation in the semantic system. Frequency effects – where high-frequency words are easier to retrieve than low-frequency words – are associated with lexical disorders. In the framework of two-step models, Kittredge et al. (2008) argue that word frequency affects both stages of lexical retrieval, i.e., access to both lemmas and lexemes. In the Logogen model, word frequency effects arise from the POL, as representations of high-frequency words require less activation to be retrieved than representations of low frequency words (Morton, 1969). Also, the influence of part-of-speech and lexical neighborhood density is attributed to the lexicon (Harley and Bown, 1998; Laiacona and Caramazza, 2004). Whether effects of age of acquisition originate at the lexical or at the semantic level is still under discussion (Morrison and Gibbons, 2006). Word length has an impact on phoneme maintenance and sequencing and is associated with post-lexical processes, i.e., the phonological output buffer (Baddeley et al., 1975; Haluts et al., 2020). Contrasting performance for morphologically complex vs. simple words is attributed to morpho-lexical processing of complex words (Lüttmann et al., 2011).

These parameters and their specific impact on cognitive components should be considered when compiling materials for an intervention aimed at facilitating lexical retrieval. A substantial body of research has demonstrated the effectiveness of repetitive word retrieval training (Hickin et al., 2002; Boyle, 2004; Renvall et al., 2013a). Patients with post-semantic word-finding disorders appear to benefit from tasks that require both phonological and semantic processing (Lorenz and Ziegler, 2009). Effective cueing techniques include phonological, semantic, auditory, and visual hints (Sze et al., 2021).

When addressing word retrieval deficits, the aim is typically twofold: On the one hand, the goal is to enhance naming performance for treated items. On the other hand, there is an aspiration for generalization to untreated items and/or untreated tasks, as this would result in a broader improvement beyond therapy. Generalization to untreated items requires that they share semantic or phonological features with treated items, and that these features are targeted during intervention. In this case, spreading activation within the semantic system or the POL may lead to improvement of untreated semantic concepts or untreated word forms (Webster et al., 2015). However, if the word retrieval deficit is due to an impaired connection between semantics and the POL, generalization to untreated items is not expected: as the association of a semantic concept with a particular word form is, in most cases, purely incidental, each connection must be rebuilt in its own right (Miceli et al., 1996; Howard, 2000). Consequently, the selection of the items to be practiced is of great importance. The items should be meaningful for the client and have relevance in everyday contexts (Renvall et al., 2013a). Corpus analyses conducted by Renvall et al. (2013a) showed that the 100 most frequent English words include verbs, pronouns, adverbs, and prepositions. Similarly, in the German corpus "Wortschatz Leipzig" (Universität Leipzig, Institut für Informatik, Projekt Deutscher Wortschatz, 1998-2023), the 50 most frequent German words consist mainly of function words (determiners, prepositions, pronouns) and various forms of the light verbs "to have" and "to be." This underscores the importance of incorporating words other than nouns in an intervention aimed at facilitating lexical retrieval (Renvall et al., 2013a). Generalization to untreated tasks may occur within the same linguistic level, e.g., if naming by definition improves after the treatment of oral picture naming. Even more meaningful are across-level generalizations, where improved lexical retrieval extends to sentence production, connected speech or everyday communication (Webster et al., 2015).

2 Apps in speech-language therapy

The dosage of speech-language therapy plays a crucial role in its effectiveness (Bhogal et al., 2003). A recent review conducted by the RELEASE collaborators [The REhabilitation recovery of peopLE with Aphasia after StrokE (RELEASE) Collaborators, 2022] found that the most significant improvements in language and functional communication occurred when the intervention was administered 5 days a week. Breitenstein et al. (2017) demonstrated that receiving ten or more hours of speech-language therapy per week led to sustained improvements in aphasic communication disorders after a stroke. German guidelines for the rehabilitation of aphasic disorders after a stroke (Deutsche Gesellschaft für Neurologie, 2011) recommend daily speech therapy as the minimum dosage. However, the reality in outpatient speech therapy facilities differs (Bürkle et al., 2022). In Germany, the standard practice typically involves one to two therapy sessions per week (Asmussen et al., 2013). Therefore, therapeutic homework is employed to increase the frequency of the intervention through complementary self-training (Wendlandt, 2002). In this context, patients bear a high level of responsibility, as they need to complete their assignments regularly, comprehensively, and in the desired manner. Digital applications, such as apps and computer programs, can provide the necessary guidance and support that individuals with aphasia, in particular, may require (Braley et al., 2021). Participation in digital technologies – in the case of aphasia rehabilitation, the competent use of a high-quality speech language app – can therefore contribute to the self-determined pursuit of individual health goals.

2.1 Advantages of app-based approaches

While therapeutic homework can increase the frequency of interventions, it lacks the interaction between the client and the speech-language therapist (SLT). Most notably, traditional "paper and pencil" tasks do not provide any feedback. Consequently, clients remain unaware of their performance until their next session with their SLT. In contrast, digital applications, can offer immediate feedback that is independent of the therapist, objective, and reliable. Outcome-oriented feedback which visualizes the learning progress is motivating (Kurland et al., 2014). Clients are encouraged to practice more frequently, thereby achieving the intended intensive treatment (Stark and Warburton, 2018; Leinweber, 2021). At the same time, clients assume greater (shared) responsibility for their therapy (Palmer et al., 2019), as app-based learning fosters personal responsibility and self-determination (Kurland et al., 2014).

The availability of high-quality speech-language applications is still limited in German-speaking countries. Only a few apps are specifically designed for adults with aphasia. These apps usually focus on training reading comprehension, writing, or auditory comprehension – tasks where the app can easily provide feedback. In contrast, oral naming tasks require the clients to self-assess their own production, such as by comparing their answer to the target word that is presented auditorily and/or visually by the app. Obviously, this self-evaluation can be error-prone when self-monitoring is impaired. Therefore, there is a demand for external feedback in word production tasks as well. Achieving therapist-independent training for verbal speech production necessitates the use of automatic speech recognition (ASR) technology to recognize and assess spoken words. There is initial evidence suggesting that digital speech recognition technologies utilizing ASR can improve verbal word production in individuals with aphasia and apraxia of speech (Ballard et al., 2019).

2.2 LingoTalk – a speech language app with automatic speech recognition

LingoTalk (© LingoLab 2021–2023) is a German speech-language app designed to enhance word retrieval in individuals with aphasia. The app has been available in the Google and Apple app stores for tablet computers and large smartphone displays since spring 2021. LingoTalk focuses on training lexical retrieval through cued verbal picture naming. The design and objectives of LingoTalk are based on the ICF framework (World Health Organization, 2001). Training with the app is intended to facilitate lexical retrieval from the POL, ultimately enhancing communication in everyday life.

LingoTalk's linguistic database covers words that are highly relevant to everyday life and topics of general interest, allowing for a patient-oriented selection of word materials. As of April 2023, this database contained over 3200 words, categorized into 28 topics. Each word is represented by a color photo. The items are classified based on 17 linguistic parameters and divided into four levels of increasing difficulty (easy, medium, demanding, hard). The difficulty level of an item is determined by various linguistic criteria, including word frequency (high vs. medium vs. low; derived from Digitales Wörterbuch der Deutschen Sprache Universität Leipzig, Institut für Informatik, Projekt Deutscher Wortschatz, 1998-2023, word length in syllables (ranging from 1 to more than 4 syllables), morphological complexity (simple vs. complex), syllable complexity (absence vs. presence of consonant clusters), and stress pattern (trochaic vs. non-trochaic). If norm data are available (see Schröder et al., 2012), age of acquisition (early vs. late) and familiarity (high, medium, low) are taken into consideration. Shifts in place of articulation (none, few, many) and phoneme-grapheme regularity (regular vs. irregular) are considered if the other criteria result in an inconclusive classification. Easy words, for instance, are of high or medium frequency, acquired early, morphologically simple, have one or two syllables, do not contain consonant clusters, and two-syllable words have a trochaic stress pattern. In contrast, words that are morphologically complex (e.g., compounds and or reflexive, prefix, or particle verbs) or consist of four or more syllables are classified as demanding or hard. Further details on how these different criteria correspond to the four levels of difficulty can be found in the Supplementary material.

The words to be practiced can be selected either based on a specific topic (e. g., "city life and traffic") or according to psycholinguistic variables (e.g., "two-syllable high-frequency nouns with /n/ or /m/ in the initial sound"). In both cases, one can choose not only content words but also items from the core vocabulary. The core vocabulary includes high-frequency function words (e. g., "more," "not") that are not tied to a particular topic but are applicable in various contexts (Boenisch and Sachse, 2020). LingoTalk's items encompass 11 part-of-speech (nouns, different verb forms, adjectives, adverbs, perfect participles, pronouns, numerals, prepositions and interjections), including both concrete and abstract concepts.

LingoTalk is the first German speech-language app to incorporate ASR, enabling app-based evaluation and feedback in a verbal picture naming task. When ASR is employed, the app assesses the response and indicates whether the item was named correctly or not. LingoTalk's ASR relies on the speech recognition software provided by Apple (SIRI) and Android-based devices (Google Speech). Data protection regulations are strictly adhered to and the use of ASR requires explicit consent from the clients. To determine the accuracy of the app's ASR functionality, we analyzed 1801 utterances from ten neurotypical native German speakers (six men, four women, aged 20–70). Each speaker named 50 pictures depicting nouns, verbs, and adjectives up to six times on two different days. We tested the quality of Google Speech's ASR with 15 Samsung tablet computers (Galaxy Tab S2 or Galaxy Tab S6; Android 7.0) and Apple's SIRI ASR with five devices (iPad Pro and iPad Air; PadOs 14.1.1). The correctness of the utterances was assessed by both the ASR and two experienced SLTs. There was a high level of agreement in the ratings between ASR and the SLTs, reaching 98.05% for Google Speech (1259/1284 utterances) and 99.26% for SIRI (508/517 utterances). Google Speech rejected 24 responses that the SLTs rated as correct and accepted one response that the SLTs rated as incorrect. SIRI rejected 9 responses that the SLTs rated as correct. With both Google Speech and SIRI operating with an accuracy rate of over 95%, we consider the feedback provided by the app to be reliable. Moreover, ASR offers the opportunity for feedback-driven learning outside of face-to-face sessions. This enhances clients' independence from their SLT and strengthens their sense of competence, responsibility, and self-efficacy in rehabilitating their language skills. As an alternative to ASR, the correct/incorrect rating can also be performed by either the SLT or by the clients themselves. For self-evaluation, clients can play an audio file containing the target word and then reveal the written word form to compare their reaction with the target item.

The app is designed in a tandem version, allowing the client's and therapist's applications to be linked via an encrypted code. The SLT can compile patient-oriented and linguistically tailored materials, define the type of feedback, and choose from ten evidence-based cues (Sze et al., 2021). Linked accounts also allow the therapist to monitor the patient's treatment routine and their progress. LingoTalk records the date, time and duration of each practice session, as well as the number of correct responses (with and without cueing), naming latencies, and the cues that were most frequently employed.

During the training session (Figure 2), the patient is presented with the picture to be named and hears the instruction "Please name the picture!" or "What can be seen here?" On the right side (or optionally on the left side), cues are displayed, offering phonological, semantic, and graphemic support in a hierarchical sequence. Additionally, an audiovisual articulation video is provided. When ASR is enabled, the patient presses the microphone button while uttering the word and the app responds with outcome-oriented feedback. Successful naming is indicated by a green flag along with a confirming sound, while incorrect responses are marked with a gray tag and an error sound. The patient is granted three attempts to name the item correctly and can make use of cues if necessary. After three unsuccessful attempts, the client receives corrective feedback, and the target word is presented auditorily and in written form beneath the picture. Then the next picture is presented. The progress of the exercise, including the number of remaining items, is displayed in a progress bar at the top.

When the exercise is completed, the evaluation screen (Figure 3) shows a pie chart that summarizes the current session (naming accuracy with and without cues) and a bar chart that shows the therapy progress (naming accuracy with and without cues, average naming latencies). A table displays date, time and duration of a practice session. It also includes average naming latencies, the most frequently used cues, and naming accuracy with and without cues.

3 Research questions

The aim of our case series study was twofold. Firstly, we monitored if and how participants with aphasia were able to carry out the self-administered training with LingoTalk and asked:

1. Can participants with aphasia manage their app-based treatment independently and adhere the treatment protocol as instructed?

Secondly, we investigated whether intensive training with the LingoTalk app leads to improved oral naming in aphasia. We aimed to answer the following questions:

2a. Does treatment with LingoTalk result in improved oral naming of treated items?
2b. Does this improvement generalize to (a) untreated but similar items and/or (b) treated items in an untreated task that is similar to the treated task (naming by definition)?
2c. Is there a transfer from the highly structured practice sessions to spontaneous speech and everyday communication?

4 Materials and methods

The research presented in this paper was conducted in the context of three Bachelor's projects (Schmitz-Antonischki, 2021; Ahrens, 2022; Saalfrank, 2023) carried out at the University of Potsdam and the P.A.N. Center for Post-Acute Neurorehabilitation in Berlin. The research was conducted in accordance with the relevant institutional guidelines, including the EU General Data Protection Regulation (GDPR) and the Brandenburg State Data Protection Law (BbgDSG) as well as the German Research Foundation's Guidelines for Safeguarding Good Research Practice. The project was approved by the research coordinator of the P.A.N Center.

4.1 Participants

An intervention with LingoTalk is appropriate for native speakers of German who have impaired lexical retrieval due to aphasia. For the present study, the participants had to meet the following criteria:

Inclusion criteria

- Native speaker of German
- Chronic aphasia, at least 6 months post-onset
- Word-finding difficulties and impaired oral naming due to a post-semantic deficit, i.e. preservation of basic semantic knowledge
- Written informed consent
- Interest in working with a speech-language therapy app and informal commitment to practice on a daily basis
- Ability to use a tablet computer and the LingoTalk app without assistance
- Completion of a test trial
- Access to a stable Wi-Fi connection at home

Exclusion criteria

- Moderate or severe speech motor disorder
- Severely impaired auditory comprehension that might compromise the understanding of instructions

The diagnosis of aphasia was established using either the ACL (Kalbe et al., 2010) or the AAT (Huber et al., 1983). In the ACL, participants needed to score below the cutoff, i.e., <135 points. The outcome of the AAT had to indicate "aphasia." Additionally, participants' performance in the Wortproduktionsprüfung [WPP/subtest 3 (Blanken et al., 1999)], where they were required to orally name 60 nouns, had to be < 90% correct. As LingoTalk does not allow for semantic treatment, participants had to score above cutoff in subtests 1–3 of the Bogenhausener Semantik Untersuchung (BOSU, Glindemann, 2002), which require thematic and taxonomical semantic knowledge to judge on situations and features. Participants were also required to demonstrate the ability to initiate, turn off, and recharge their tablet computers. After receiving an extensive demonstration of how to use the LingoTalk app and the ASR, participants had to successfully complete a test trial with five to ten items without assistance.

Three participants (P1, P2, P3), who met all inclusion criteria and none of the exclusion criteria were recruited in the P.A.N. Center for Post-Acute Neurorehabilitation (Berlin) and through University of Potsdam's Patholinguistics patient database. The participants were provided with information

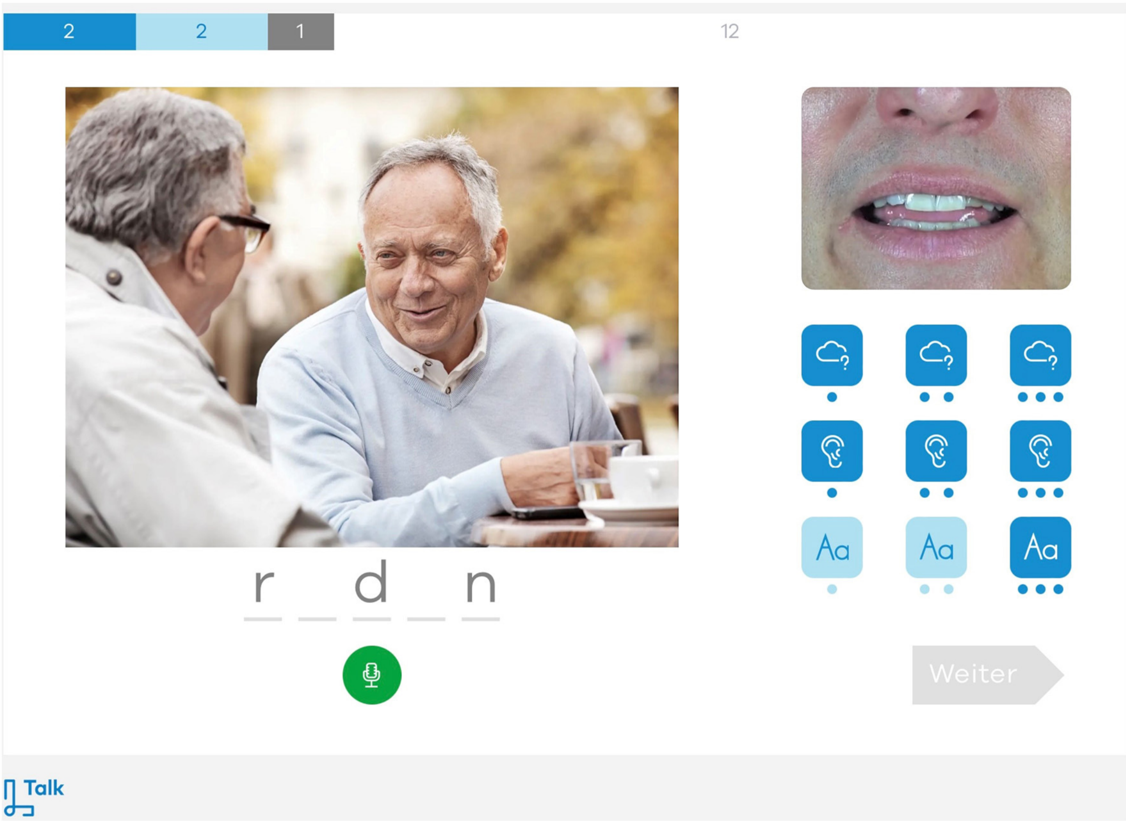

FIGURE 2
Target item "reden" (to talk) with semantic, phonological, orthographic cues and the articulation video. The green button allows ASR-based feedback on the spoken response.

about the study's purpose, procedures, collection and use of data and the potential risk of an unsuccessful therapy outcome. They provided written consent to participate in the study and they could withdraw from the study at any time and without giving any reason. All participants had received speech-language therapy before their involvement in the study, but no additional speech-language treatment was administered during the intervention with LingoTalk. P1 and P2 already owned tablet computers, while P3 was provided with a loan device for the duration of the study. Table 1 shows the demographic information for each patient and their language profiles.

The purpose of the neurolinguistic assessment was twofold: to confirm that the participants met the inclusion criteria and to examine the severity and the nature of their word-finding difficulties. Using the WPP (Blanken et al., 1999), we investigated the participants' overall naming performance and examined the influence of word frequency, word length, and articulatory complexity on lexical retrieval. The BOSU assessment (Glindemann, 2002) ruled out severe semantic impairment. Comparing oral naming to oral reading performance [either WPP subtest 3 vs. 6 (Blanken et al., 1999) or LEMO 2.0 T13 vs. T8 (Stadie et al., 2013)] provided insight whether impaired oral naming was caused by damaged access from semantics to POL and/or impairment of representations in the POL.

P1 was a 23-year-old German-speaking woman with 16 years of education. She had suffered a traumatic dissection of the left carotid artery and an occlusion of the internal carotid artery (ICA) due to an accident 1.5 years prior to this study, resulting in damage in the entire left ICA territory. At the time of the study, P1 resided in a center for post-acute neurorehabilitation and was highly motivated to maximize her rehabilitation potential. P1 used a smartphone on a daily basis and had recently acquired a tablet for Augmentative and Alternative Communication (AAC). Assessment with the AAT (Huber et al., 1983) confirmed a global aphasia, although clinical observation was more indicative of Broca's aphasia. Her spontaneous speech was non-fluent and agrammatic, characterized by frequent word-finding difficulties, and displayed mild symptoms of apraxia of speech. Neurolinguistic assessment revealed impaired oral naming [50% correct in WPP (Blanken et al., 1999)] with no effects of word frequency, word length, or articulatory complexity. Errors included semantic, phonemic, and formal paraphasias. According to BOSU (Glindemann, 2002), basic semantic knowledge was preserved, whereas oral reading of regular and irregular words was impaired [23% correct in LEMO 2.0/subtest T8 (Stadie et al., 2013)]. Therefore, the naming disorder was attributed to the access from semantics to the POL and/or the POL itself.

P2 was a 39-year-old woman with 9 years of education. 1:6 years prior to this study, she suffered a left carotid ischemic stroke. P2

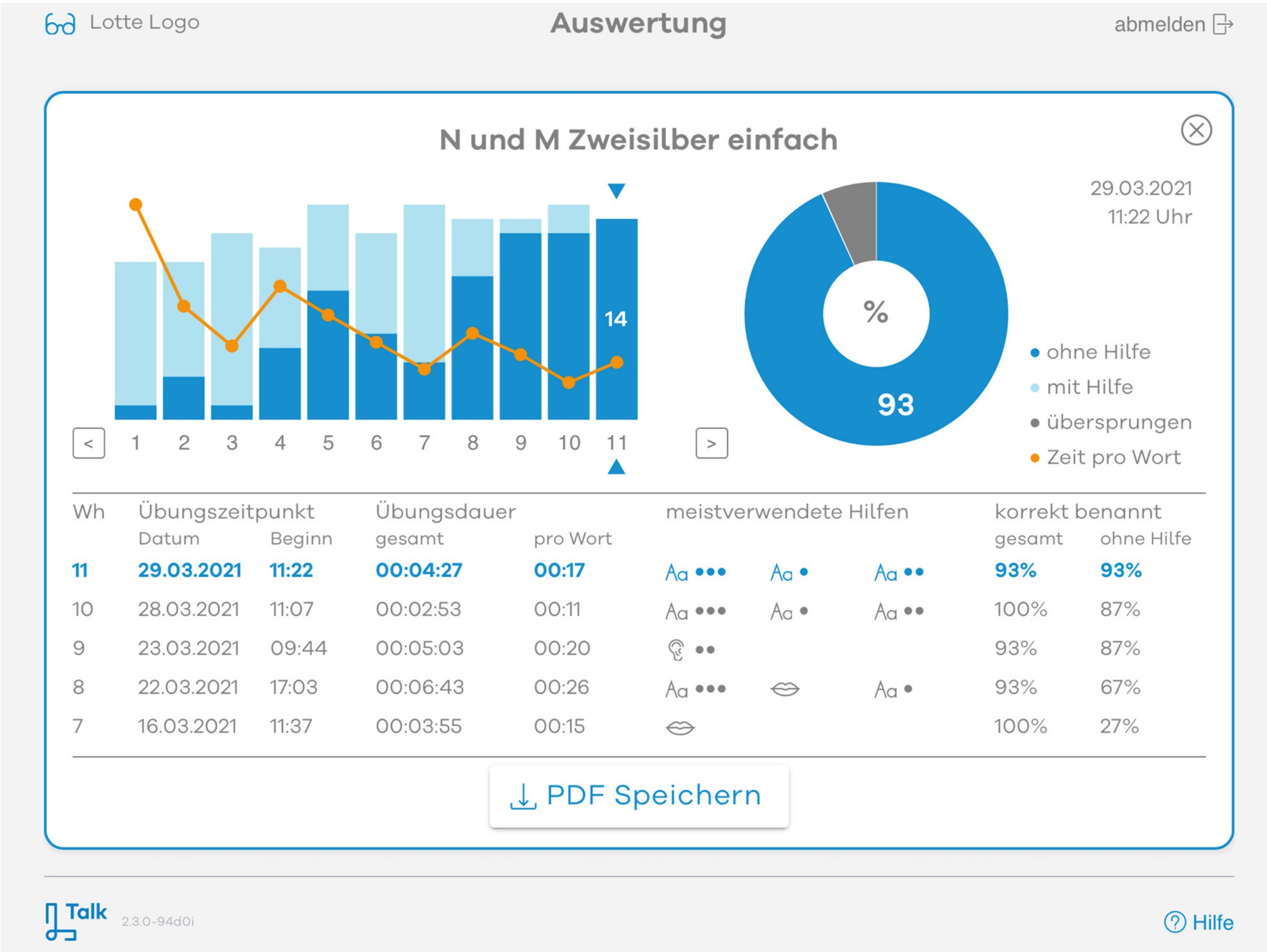

FIGURE 3
Evaluation screen.

resided in a center for post-acute neurorehabilitation and received treatment there. She was an enthusiastic computer gamer and proficiently used a laptop and a smartphone. She also owned a tablet for AAC but did not actively use it. Assessment with the ACL (Kalbe et al., 2010) confirmed the diagnosis of aphasia. Her spontaneous speech was non-fluent, characterized by incomplete sentences due to word-finding difficulties. Oral naming was impaired [65% correct in WPP (Blanken et al., 1999)], and performance for high-frequency words was significantly better than for low-frequency words (24/30 vs. 15/30, $p = 0.029$; Fisher's exact test). Word length and articulatory complexity did not influence naming performance. There were few phonological paraphasias, but most errors were semantically related to the target. As basic semantic knowledge was preserved [cf. BOSU (Glindemann, 2002)], the latter errors were attributed to insufficient activation of the lexical entry in the POL, resulting in the retrieval of a semantically similar response. Since reading and oral naming were equally affected [cf. LEMO 2.0 (Stadie et al., 2013), T13 vs. T8, 17/20 vs. 41/60, $p = 0.25$], the impairment was localized within the POL itself, rather than in lexical access.

P3 was a 69-year-old German-speaking man with 15 years of education. He had suffered a left carotid ischemic stroke more than 20 years prior to this study. After a pause of several years, P3 requested the resumption of speech language therapy and was included in the study in an outpatient setting. P3 showed great interest in working with an app but had never used a tablet computer or a smartphone before. The ACL (Kalbe et al., 2010) confirmed a persistent mild to moderate aphasia. His spontaneous speech was fluent but marked by various symptoms of word-finding difficulties, including hesitations, rewording, empty phrases, and repetition of words and phrases.

Oral naming was impaired [70% correct in WPP (Blanken et al., 1999)] and affected by word frequency (high: 25/30 correct vs. low: 17/30; $p = 0.047$) and word length (1 syllable: 17/20 vs. 3 syllables: 10/20, $p = 0.041$). Both semantic and phonologic errors occurred. Articulatory complexity did not influence the naming performance. Basic semantic knowledge was preserved [cf. BOSU (Glindemann, 2002)], and the WPP (Blanken et al., 1999) showed that oral reading was significantly better than naming of the very same words (59/60 vs. 42/60, $p < 0.001$). Therefore, the naming disorder was attributed

TABLE 1 Patient demographic information and language profiles.

	P1	P2	P3
Demographic information			
Gender	f	f	m
Age (years)	23	39	69
Educational level/former profession	High school/student	Secondary school/shop assistant	Bachelor's/retired CEO
Etiology	Traumatic dissection LCA and occlusion ICA	Ischemia left MCA	Ischemia left MCA
Time post injury (months)	17	18	248
Neurolinguistic assessment			
AAT Overall result	Aphasia (global)	—	—
ACL Overall result	—	Aphasia (79/148)	Aphasia (114/148)
Severity of aphasia (clinical observation)	Severe	Moderate	Mild
BOSU (subtests 1–3)	26/30 all subtests above cut-off	27/30 all subtests above cut-off	29/30 all subtests above cut-off
WPP (subtest 3) oral naming	30/60	39/60 frequency effect	42/60 frequency effect length effect
WPP (subtest 6) oral reading	45/60	—	59/60
LEMO 2.0 (T13) oral naming	16/20 impaired	17/20 impaired	18/20 impaired
LEMO 2.0 (T8) oral reading regular/irregular nouns	14/60 impaired	41/60 impaired	51/60 impaired

LCA, left carotid artery; ICA, internal carotid artery; MCA, middle cerebral artery; AAT, Aachen Aphasia Test (Huber et al., 1983); ACL, Aphasie Check Liste (Kalbe et al., 2010); WPP, Wortproduktionsprüfung (Blanken et al., 1999); LEMO 2.0, Lexikon modellorientiert (Stadie et al., 2013); BOSU, Bogenhausener Semantik Untersuchung (Glindemann, 2002).

to an impaired access from semantics to the POL while semantic and lexical representations were intact.

4.2 Materials

LingoTalk allows for an individual selection of items that takes into account both the patient's needs and interests as well as the degree of language impairment. PlanBe (Pfeiffer and Leisner, 2016) was used to identify the patients' interests and hobbies, their interlocutors and the communicative topics and situations they engage in. Involving patients in the item selection process and developing individual item sets makes the materials relevant to everyday life and usually increases motivation for the intervention (Renvall et al., 2013b). Based on the information from PlanBe (Pfeiffer and Leisner, 2016), topics of interest were selected in LingoTalk individually for each participant. In addition, the difficulty of the items (easy, medium, demanding, hard) was adjusted to the severity of the oral naming impairment. To investigate item-specific effects as well as generalization, Brüsch (2022) suggests to use three item sets for each participant: (1) treated items, (2) untreated items from treated topic, (3) untreated items from untreated topic. Each set should contain 20–30 items and preferably different part of speech. Table 2 shows the item selection for each participant. A full list of items can be found in the Supplementary material.

Using PlanBe, P1 identified six topics relevant to her daily life, with four of them being treated and two remaining untreated. The intervention comprised 120 items that were divided into three sets: Treated items ($n = 50$), untreated items from the treated topic ($n = 30$), and untreated items from an untreated topic ($n = 40$). Each set contained the same proportion of nouns (60%), verbs (20%), and adjectives (20%) and the same number of low, medium, and high frequency items. Treated and untreated sets were matched for item difficulty. As P2 was about to move into an assisted living facility, the treated items were chosen from topics related to living at home. Untreated items were chosen from other topics that were relevant at that time. All item sets (treated items, untreated items from treated topic, untreated items from untreated topic) comprised 20 nouns, 10 verbs and 10 adjectives. Each set contained the same number of low, medium, and high-frequency items. Treated and untreated sets were matched for item difficulty. P3 is an avid sailor and therefore identified summer holidays and weather as relevant topics. As previous interventions had revealed a specific deficit for morphologically complex words (Wegener et al., 2010), only low frequent compound nouns were treated. Compounds of low and medium frequency, belonging to the same topic as the treated items, served as untreated control items.

4.3 Planned treatment and procedure

The overall aim of the intervention was to facilitate lexical retrieval for items that were chosen according to the participants' needs and interests. The treatment was planned within a multiple

TABLE 2 Items individually selected for each participant.

	Item difficulty (LingoTalk classification)	Treated topics	Untreated topics	Part of speech	Treated items	Untreated items treated topic	Untreated items untreated topic
P1	Medium Demanding Hard	• Family • Grocery shopping • In the kitchen • Climate and weather	• City life and traffic • Nature and environment	Nouns Verbs Adjectives	$N = 50$	$N = 30$	$N = 40$
P2	Demanding Hard	• Laundry • Living at home • In the morning • Cleaning	• Family • COVID-19 pandemic • Easter	Nouns Verbs Adjectives	$N = 40$	$N = 40$	$N = 40$
P3	Hard	• Holidays • Weather	—	Compound nouns	$N = 22$ (low frequency)	$N = 44$ (low and medium frequency)	—

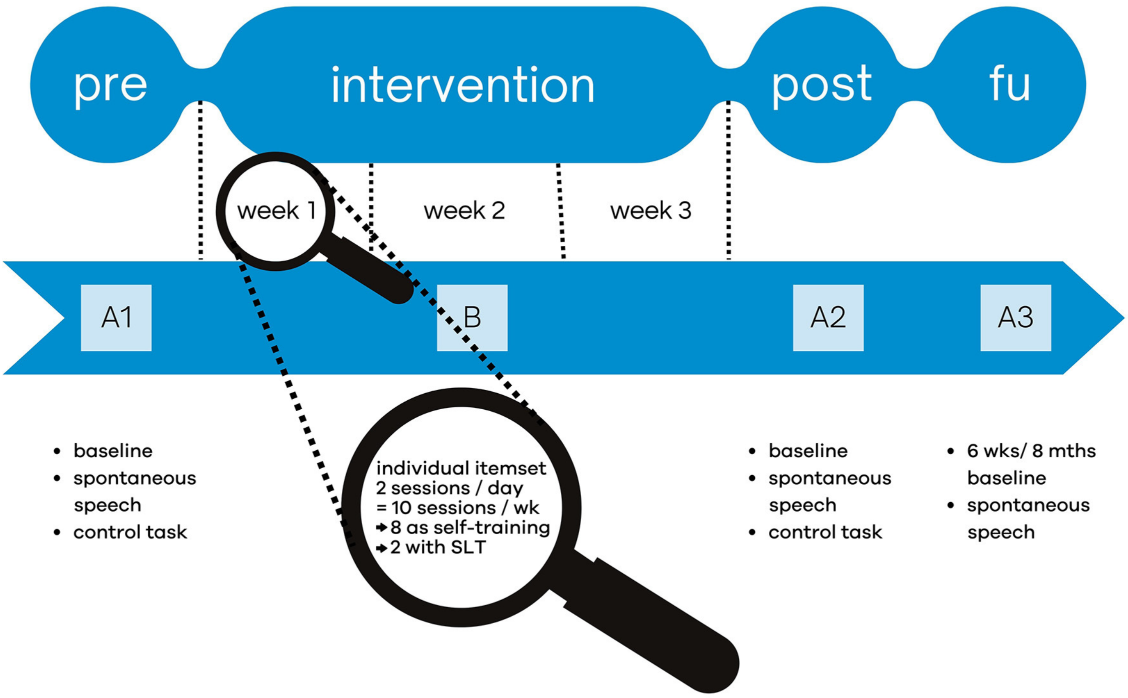

FIGURE 4
Timeline of baseline measures and intervention.

baseline design (A1-B-A2-A3). The timeline is illustrated in Figure 4.

In the intervention phase (B), participants were asked to complete 10 training sessions per week within a period of 3 weeks, resulting in 30 sessions in total. We expected them to practice twice a day on 5 days a week, leading to 15 (out of 21) training days. The very first session took place in a face-to-face setting to ensure that the participants knew how to use the tablet computer and the app. Familiarization with LingoTalk included instructions on how to start/end a training session, how to use ASR, and how to systematically choose hierarchical cues in case of incorrect answers. The SLT observed the training session and assisted if necessary until the participant felt comfortable using the app. Afterwards, the participants started the self-administered intervention with LingoTalk (session 2–6). The SLT was able to monitor the patients' progress in the professional version of LingoTalk, allowing the experimenter to check if the patient completed the therapy sessions as planned. Session 7, again, took place in a face-to-face setting. The SLT and patient reflected on the already completed sessions and addressed any queries or technical issues. In weeks 2 and 3, two sessions were also supervised by a therapist. Thus, the 30 therapy sessions consisted of six supervised and 24 self-administered sessions. To help the participants track their progress, they were provided with a schedule they could check off when they had completed a therapy session.

Participants were asked to practice all treated items in each therapy session. Treatment was always administered with LingoTalk, and the task was oral picture naming. Participants were shown a picture and heard the instructions “What can be seen here?” and “Please name the picture!” Afterwards, they named the picture. The pictures were presented either in one block (P3) or split up into several blocks (P1, P2) to keep the blocks shorter and more homogeneous. For example, for P2, we created six blocks according to topic and part of speech. For patients to receive direct feedback and to monitor their performance throughout the intervention, ASR should be used. A response would be considered correct if the ASR could identify the response. To make use of the ASR, the patients had to press the “record” button (symbolized by a green vibrating microphone), hold the button, give their answer, and then release the button (cf. Figure 2). P1 and P2 could use ASR immediately. P3, however, had major difficulties in keeping the button pressed while giving the answer. Despite intensive training, he would press the button, release it, and only then name the picture. Therefore, P3 was asked to use the self-assessment mode to classify his reactions as correct or incorrect. If an item was named incorrectly or could not be named at all, the participants could make use of gradually increasing phonological, semantic, and/or graphemic cues. An audiovisual mouth image could be used as a maximal cue. Although it was recommended to start with the weakest cue, participants were free to select whichever cue they found helpful. After three incorrect naming attempts, the target word was presented auditorily and in written form below the picture.

To answer research question 1 – can participants with aphasia manage their app-based treatment independently and do they follow the treatment protocol as instructed? – we used LingoTalk’s data documentation to closely monitor the intervention phase. We evaluated how often and how regularly the participants named the treated items and compared these data to the treatment protocol. Naming accuracy for treated items was recorded for every single session to monitor each participant’s progress throughout the intervention. To account for treatment effects, and to answer research questions 2a–c, baseline measures were conducted before treatment (A1), directly after treatment (A2) and in a follow-up at least 5 weeks after the intervention (A3). Baselines measures examined item-specific effects, i.e., whether treated items improved in a treated task (oral picture naming). Furthermore, different types of generalization were investigated: (1) generalization to untreated items a treated task (oral picture naming of untreated items), (2) generalization to treated items in an untreated task (oral naming by definition of treated items), (3) generalization to a comparable task [oral picture naming, WPP/subtest 3 (Blanken et al., 1999)]. Transfer to communication was measured with the Amsterdam Nijmegen Everyday Language Test (ANELT) (Blomert and Buslach, 1994). The analysis of spontaneous speech focused on the number of word-finding difficulties and phrases, as their ratio (one word-finding difficulty every n = x phrases) is a sensitive marker for the frequency of word-finding difficulties (Bayer, 1986). The higher the ratio, the fewer word-finding difficulties occur per phrase. Spontaneous speech was collected in semi-structured interviews that covered both treated and untreated topics (cf. Table 2). In addition, the Communicative Activity Log (CAL) (Pulvermüller and Berthier, 2008) informed about each participant’s everyday communicative practice as perceived by the participants (CAL self-assessment) and/or a conversation partner (CAL external assessment). To control for unspecific, general improvement, a task unrelated to oral naming and not practiced during intervention was administered before and after treatment. The unrelated control task was writing non-words to dictation for P1, written picture naming for P2, and oral non-word repetition for P3. Statistical analyses with either the McNemar test or the Fisher’s exact test were carried out for all baseline measures except spontaneous speech.

TABLE 3 Execution of the treatment protocol.

	Training days (within 3 weeks)	Training sessions	Sessions per day	Average number of sessions per training day
Planned	15	30	2	2
P1	17	30	1–2	1.82
P2	12	29	2–4	2.42
P3	16	46	1–5	2.88

5 Results

5.1 Execution of the treatment protocol

Participants had been instructed to complete 30 training sessions, evenly distributed over 15 days (i.e. twice a day), within a total period of 21 days. However, it was observed that none of the participants fully adhered to this protocol (cf. Table 3). While P1 deviated only marginally by completing 30 sessions within 17 days, P2 and P3 showed greater variation. P2 found it challenging to practice on a regular basis but still aimed to complete the 30 training sessions. As a result, she increased the number of sessions per day, resulting in 29 sessions within 12 days. In contrast, P3 enjoyed working with the tablet computer and completed 46 training sessions within 16 days.

All participants became considerably faster throughout the therapy process. Initially, sessions lasted ∼40–45 minutes, but by the end of the treatment period, they were completed in just 5–10 min. Visual inspection of naming accuracy (Figure 5) showed unexpected data for P3 in sessions 28–37. During this period, his naming accuracy suddenly dropped from 70% to zero. Upon investigation, P3 explained that he had attempted to use the ASR once more. It seems that he failed to do so without noticing. Consequently, none of his responses were identified as correct during this period. When P3 reverted to self-assessment in session 38, his naming accuracy returned to 100%.

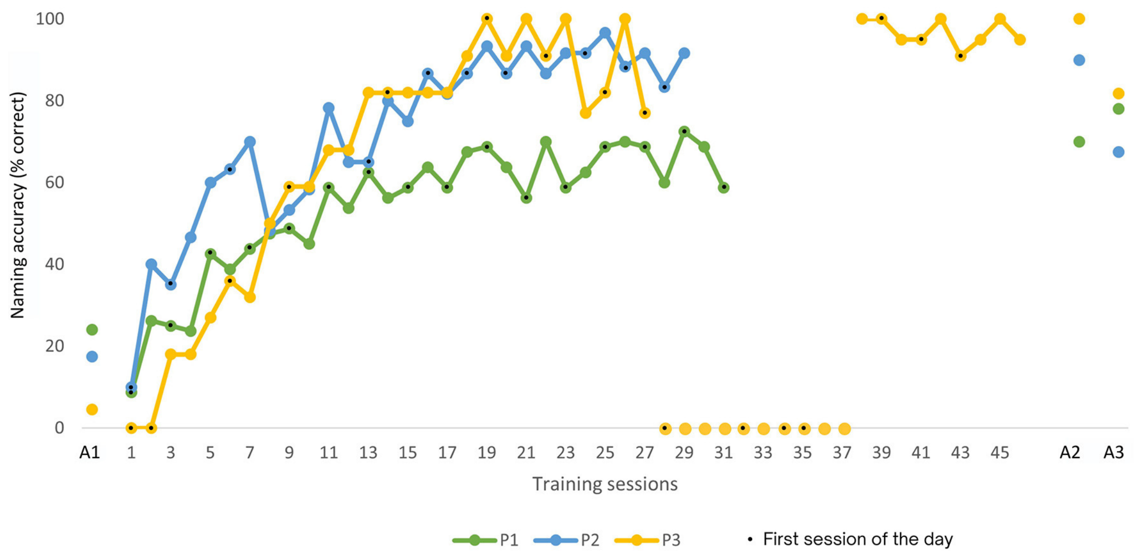

FIGURE 5
Improvement of naming accuracy during intervention, with A1/A2/A3 being baseline measures before/after/follow-up. Number of training sessions differed for P1 ($n = 30$), P2 ($n = 29$), and P3 ($n = 46$). Missing data for P3 in sessions 28–37 due to problems using the ASR.

5.2 Outcome measures

Naming accuracy for treated items continuously increased for all participants as depicted in Figure 5.

The percentage of non-overlapping data (PND) (Scruggs et al., 1987), indicating the number of training sessions where performance was better than in the initial baseline, exceeded 90% for all participants. Therefore, the improvement in treated items is considered highly reliable. Table 4 shows the results of the baseline measures before (A1), after (A2) and in a follow up (A3). The main result is that all participants showed significant improvement for treated items that sustained at least 5 weeks after treatment had been withdrawn. Generalization effects occurred to different extents.

P1 showed a significant improvement of naming accuracy for treated items (before: 12/50 correct vs. after: 35/50, $p < 0.001$, McNemar Test). This training effect was sustainable and naming accuracy 8 months after treatment was still significantly better than before (12/50 vs. 39/50, $p < 0.001$). Immediately after treatment, there was no generalization to untreated items that belonged to a treated category (4/30 vs. 9/30, $p = 0.074$), but the improvement became significant in the follow up-assessment (4/30 vs. 22/30, $p < 0.001$). There was no generalization to untreated items of an untreated category, but naming accuracy in the WPP (Blanken et al., 1999) improved significantly (30/60 vs. 48/60, $p < 0.001$). There are still very many word-finding difficulties in spontaneous speech although their amount decreased a little bit (from one in 1.54 phrases to one in 2.18 phrases and one in 5.3 phrases in the follow up). Assessment with the CAL (Pulvermüller and Berthier, 2008) could not detect any changes in P1's communication in daily life. As the performance in an unrelated control task [LEMO 2.0 T9, writing non-words by dictation (Stadie et al., 2013)] remained stable, the item-specific training effect and the generalization to untreated items and the WPP (Blanken et al., 1999) can be attributed to the intervention with LingoTalk and are not caused by some general or unspecific improvement.

P2 showed a significant improvement of naming accuracy for treated items (before: 7/40 correct vs. after: 36/40, $p < 0.001$). That improvement remained stable in a follow up test 5 weeks after treatment (7/40 vs. 27/40, $p < 0.001$). There was no generalization to any of the untreated items, including WPP (Blanken et al., 1999). Spontaneous speech analysis was not very informative as P2's reactions in baseline A2 were rather taciturn and brusque as she knew that she had answered the very same questions already before. While both the performance in the ANELT (Blomert and Buslach, 1994) and the external assessment of P2's communicative abilities (CAL) (Pulvermüller and Berthier, 2008) did not change, P2 herself reported that "she speaks much better" resulting in a significantly better self-assessment with the CAL (before: 35/65 points vs. after 51/65, $p = 0.001$, Fisher's exact test). Performance in an unrelated control task [WPP written naming (Blanken et al., 1999)] did not change, tracing back the item-specific training effect to our intervention.

P3 showed a significant improvement of naming accuracy for treated items (before: 1/22 correct vs. after: 22/22, $p < 0.001$). That improvement remained stable in a follow up test 6 weeks after treatment (1/22 vs. 18/22, $p < 0.001$). Improved oral picture naming generalized to naming by definition (1/18 vs. 12/18, $p = 0.003$), i.e. treated items improved sustainably in an untreated task. There was no significant generalization to any of the untreated items, including WPP (Blanken et al., 1999). The amount of word findings difficulties dropped from one in 5.05 phrases ("very many") to one in 6.85 phrases ("many"). The CAL self-assessment (Pulvermüller and Berthier, 2008) and the ANELT (Blomert and Buslach, 1994) could not detect any changes in P3's communicative behavior. Again, the performance in an unrelated control task [Lemo 2.0 T5, repeating non-words (Stadie et al., 2013)] remained

TABLE 4 Results of baseline measures.

		A1 (before)	A2 (after)	A3 (follow up)
P1	Treated items ($N = 50$)	12	35***	39***
	Untreated items/treated topic ($N = 30$)	4	9	22***
	Untreated items/untreated topic ($N = 40$)	5	11	11
	WPP oral naming ($N = 60$)	30	48***	50***
	CAL self assessment (65 points)	49	51	Not tested
	CAL external assessment (45 points)	26	27	Not tested
	Control task: LEMO 2.0 T9 writing non-words to dictation ($N = 40$)	0	0	Not tested
	Spontaneous speech			
	No. of phrases	77	72	90
	No. of word-finding difficulties	50	33	17
	Ratio phrases/WFD (i.e. 1 WFD every $N = x$ phrases)	1.54	2.18	5.3
	Frequency of WFD according to Bayer (1986)	Very many	Very many	Very many
P2	Treated items ($N = 40$)	7	36***	27***
	Untreated items/treated topic ($N = 40$)	11	11	12
	Untreated items/untreated topic ($N = 40$)	5	3	4
	WPP oral naming ($N = 60$)	39	40	Not tested
	CAL self assessment (65 points)	35	51*	Not tested
	CAL external assessment (45 points)	25	24	Not tested
	ANELT comprehensibility (40 points)	27	25	27
	ANELT intelligibility (40 points)	35	34	29
	Control task: WPP written naming ($N = 60$)	31	31	34
	Spontaneous speech			
	No. of phrases	55	40	69
	No. of word-finding difficulties	4	0	3
	Ratio phrases/WFD (i.e. 1 WFD every $N = x$ phrases)	13.75	Incalculable	23
	Frequency of WFD according to Bayer (1986)	Some	None	Hardly any
P3	Treated items (low frequency) ($N = 22$)	1	22***	18***
	Untreated items (low and medium frequency) ($N = 44$)	2	3	Not tested
	Treated items/untreated task (naming by definition) ($N = 18$)	1	12**	10*
	WPP oral naming ($N = 60$)	42	50	Not tested
	CAL self assessment (65 points)	41	43.5	Not tested
	ANELT comprehensibility (40 points)	33	30	31
	ANELT intelligibility (40 points)	38	40	29
	Control task: LEMO 2.0 T5 repeating non-words ($N = 40$)	33	31	29
	Spontaneous speech			
	No. of phrases	86	89	Not tested
	No. of word-finding difficulties	17	13	Not tested
	Ratio phrases/wfd (i.e. 1 WFD every $N = x$ phrases)	5.05	6.85	Not tested
	Frequency of WFD according to Bayer (1986)	Very many	Many	Not tested

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$ (McNemar); #$p < 0.01$ (Fisher's exact). WPP, Wortproduktionsprüfung (Blanken et al., 1999); CAL, Communication Activity Log (Pulvermüller and Berthier, 2008); LEMO 2.0, Lexikon modellorientiert (Stadie et al., 2013); ANELT, Amsterdam Nijmegen Everyday Language Test (Blomert and Buslach, 1994); WFD, word-finding difficulties.

stable and all improvement can be ascribed to the intervention with LingoTalk.

6 Discussion

Three participants with aphasia took part in an intervention using the speech-language app LingoTalk to improve lexical retrieval. Using PlanBe (Pfeiffer and Leisner, 2016), all participants were able to identify topics that were either of general interest or important in their daily life. Although PlanBe (Pfeiffer and Leisner, 2016) was originally developed for individuals who use AAC, it proved to be suitable and beneficial for clients with aphasia as well. LingoTalk's extensive database allowed for item selection tailored to the patients' individual needs. Floor or ceiling effects were avoided by adjusting the level of difficulty based on LingoTalk's item difficulty rating. Importantly, the training covered various part of speech as recommended by Renvall et al. (2013a).

The intervention with LingoTalk was supervised, but mostly self-administered by the participants. Therefore, we aimed to determine whether participants with aphasia could manage their app-based treatment independently. LingoTalk recorded how often and how regularly the participants named the treated items and tracked the naming accuracy for treated items in each session. The data demonstrated that all participants were able to use LingoTalk on their own, as each of them completed at least 29 training sessions within 3 weeks. If they could not name an item, they were able to make use of gradually increasing cues and decided themselves if they wanted to use phonological, semantic, and/or graphemic cueing. Digital technology enabled individuals with aphasia to actively and autonomously pursue their therapy goals. However, none of the participants followed the intervention protocol completely, resulting in a different number of training sessions in total and per day. This divergence might be attributed to motivational factors as indicated by participants' comments. P1 liked working with LingoTalk but found practicing on her own somewhat monotonous. She favored using the app in a face-to-face setting alongside her SLT. In the case of P2, practicing independently on a daily basis was demanding, leading to a decrease in motivation and occasional complaints. Nevertheless, she did not withdraw from the study. P3, on the other hand, enjoyed training on his own as he was "less nervous when there is no therapist present." After the treatment study concluded, he acquired a tablet computer of his own. He not only continued using LingoTalk but also started to use the internet, for example, for searching about his hobbies. The intervention with a digital application eventually sparked interest and instilled self-confidence for participating in digital services such as the use of search engines and websites.

The self-administered treatment using digital technology significantly increased the therapy frequency. Instead of six face-to-face sessions within 3 weeks, the participants completed 12 to 17 training days, which aligns much closer with the recommended minimum daily speech therapy dosage (Deutsche Gesellschaft für Neurologie, 2011). After a comprehensive introduction, digital technologies facilitate self-administered treatments that are mostly independent of therapists. This is especially beneficial in cases of a shortage of SLT services, for instance, in rural areas or when there is a lack of therapists. In such circumstances, digital technologies can ensure ongoing care.

Our second objective was to assess the effectiveness of the self-administered intervention using LingoTalk. Following a three-week intervention, all three participants showed improved lexical retrieval of practiced words. Stable performance in unrelated control tasks confirmed that the improvement could be attributed to the intervention. Consequently, we conclude that LingoTalk is a suitable app for improving word retrieval in aphasia for practiced material. The results are less clear when it comes to generalizations effects. There was no generalization to untreated items, except for P1 in the follow-up assessment. However, the lack of generalization to untreated materials, is in line with the literature [see Sze et al. (2021) for a review]. Training with LingoTalk re-established the connection between a semantic concept and its corresponding word form in the POL, as the participants repeatedly named the same set of pictures. While the LingoTalk intervention employed both semantic and phonological cues to facilitate word retrieval, there were no tasks that explicitly targeted semantic or phonological features and which could have triggered spreading activation within the semantic system or the POL. In this case, item-specific improvement without generalization to untrained materials is expected (Miceli et al., 1996; Howard, 2000). For one participant (P3), there is evidence of within-level generalization (Webster et al., 2015) for treated items in an untreated task: naming by definition improved after oral picture naming had been trained. Unfortunately, we did not collect such data for P1 and P2. Two participants (P1 and P3) showed fewer word-finding difficulties in spontaneous speech, as indicated by the ratio of word-finding difficulties to phrases, suggesting the possibility of across-level generalization (Webster et al., 2015).

All of the results should be interpreted with caution, as our study has several limitations. Firstly, the study is quite small in scale, involving only three participants. A larger sample size would provide more robust insights into whether individuals with aphasia can independently manage the LingoTalk intervention and its effectiveness. Future studies on LingoTalk should aim for greater methodological consistency. Our research originated from three separate Bachelor's theses, each investigating a single case, and as such, there were slight variations in methodology during both neurolinguistic assessment and intervention. For the purposes of this paper, we combined these three single cases *post-hoc* into one case series, resulting in some lack of coherence. For instance, different tests [AAT (Huber et al., 1983) and ACL (Kalbe et al., 2010)] were used to diagnose aphasia, making the patient profiles not entirely comparable. Treatment frequency and intensity differed among participants as none of them fully followed the treatment protocol. The participants also used different feedback modes (ASR vs. self-assessment) as one of them was not able to handle ASR. When this participant nevertheless attempted to use the ASR, he failed to do so without noticing. This led to missing data on naming accuracy in some of the training sessions. The efficacy of the intervention was demonstrated for treated items, but evidence for generalization is very limited. Generalization to treated items in an untreated task was only addressed for one participant, and generalization to spontaneous speech was based on a rather general, though established, indicator, i.e. the ratio of word-finding difficulties to phrases (Bayer, 1986). While our study investigated the effectiveness of an intervention with LingoTalk, it did not compare LingoTalk to other app-based interventions [e.g., neolexon (Jakob and Späth, 2023)] or to traditional face-to-face approaches.

Despite the study's limitations, it has yielded some interesting findings. LingoTalk is the first German speech-language app that incorporates ASR, enabling app-based evaluation and feedback in an oral picture naming task. ASR was successfully utilized by two out of three participants and offered the advantage of an immediate and objective feedback. The third patient relied on self-assessment and demonstrated a high level of reliability in evaluating his own responses. Both feedback modes provide an opportunity for patients to gain independence from their SLT. While there appeared to be an initial effect of having a supervisor present – the performance of all participants dropped in the first self-training session compared to their baseline performance – this effect did not persist over the long term.

Interestingly, the mode of feedback did not seem to influence treatment efficacy. However, the decision regarding the feedback mode should be made thoughtfully. Self-assessment requires a sufficient level of self-monitoring and honesty in evaluating incorrect responses. On the other hand, utilizing ASR demands coordination between button press and speech output which proved to be challenging for one patient. Therefore, both feedback modes should be individually tested with each patient to ensure reliable feedback.

The study revealed improved naming accuracy for treated items across all three participants. The most significant increase in accuracy occurred during the first week, followed by continued improvement in the second week, and finally, a consolidation in the third week. This pattern suggests that a two-week intervention might be sufficient to achieve ∼85–90% of the overall improvement. In the cases of two participants, P2 and P3, their accuracy rates exceeded 75% and remained stable after already 15 sessions, which could indicate a ceiling effect. This might have contributed to a sense of monotony in their training and potentially affected P2's motivation negatively. Implementing a dynamic item set (Conroy et al., 2009), where additional training items are introduced once others can be named correctly, might help maintain interest and engagement over a longer training period.

LingoTalk's evaluation screen provided patients with transparency regarding their progress, motivating them to persist with the treatment. Even when P1 and, to a greater extent, P2 faced challenges, they maintained their commitment to the scheduled training sessions. They realized that they became much faster in naming items over time, resulting in shorter training sessions. Although the number of practiced items varied significantly among participants, the initial training sessions took ∼45 min for all of them. This corresponds to the typical duration of a therapy session in standard outpatient care and should not be exceeded to ensure patient engagement.

The treatment with LingoTalk demonstrated robust and long-lasting practice effects but there was limited generalization to spontaneous speech. This outcome is probably not surprising since the treatment did not encompass any functional communication tasks. In future studies, it may be beneficial to combine the self-administered LingoTalk treatment with functional-pragmatic tasks conducted in a face-to-face setting with a SLT. One might think of this as an SLT variant of the flipped classroom model (Bergmann and Sams, 2012) where monotonous and learning-intensive content is made the responsibility of the learners. We are not aware of any scientific studies on "flipped therapies" in the context of SLT but Wu (2023) has introduced the idea from an SLT perspective. In a "flipped speech room," clients would practice specific sets of items intensively at home, while in-person sessions could focus on transferring these items into meaningful communication contexts. For instance, an SLT might create an item set for a simulated visit to the market in spring, which the client practices independently with LingoTalk. During face-to-face sessions, the SLT can then integrate these learned items into sentence structures or interactive communication tasks. Eventually, in an *in vivo* intervention, the SLT could accompany the client on an actual trip to the market to make planned purchases. Improving lexical retrieval through a self-administered, app-based treatment then aligns with a participation-oriented speech-language therapy approach, enhancing everyday communication as advocated by the International Classification of Functioning, Disability, and Health (ICF, World Health Organization, 2001).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The study was conducted in accordance with the relevant institutional guidelines, including the EU General Data Protection Regulation (GDPR) and the Brandenburg Data Protection Act (BbgDSG), as well as the guidelines of the German Research Foundation to ensure good research practice. The participants have given their written informed consent to participate in this study and to the publication of the collected data in anonymized form.

Author contributions

JN developed the app LingoTalk. DS-A, JB, JH, and JN designed the treatment study. SA, TS, and DS-A compiled the treatment materials, guided and monitored the treatment of P1, P2 and P3 respectively, and collected the data. SA, TS, DS-A, JH, and JN performed statistical analyses of the outcome measures. JH and JN wrote the manuscript. All authors read and approved the submitted version.

Funding

This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project Number 491466077.

Acknowledgments

We gratefully acknowledge the efforts and reliability of P1, P2, and P3 who made this study possible. We thank the reviewers for their very helpful comments on an earlier version of this paper.

Conflict of interest

JN is the developer of the speech therapy app LingoTalk and managing director of the e-health start-up Lingo Lab UG (haftungsbeschränkt).

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2023.1210193/full#supplementary-material

References

Ahrens, S. (2022). *Anwendung eines operationalisierten Protokolls für eine Wortabruftherapie mit der App LingoTalk*. Bachelor's Thesis. Potsdam: University of Potsdam.

Anusuya, M., and Shyamala, K. C. (2021). "Tip of the tongue phenomenon" in normal and aphasic adults: an exploratory study. *Int. J. Sci. Res. Pub. IJSRP* 11, 1–8. doi: 10.29322/IJSRP.11.04.2021.p11202

Asmussen, L., Bremer, W., Heldt, C., and Krüger, S. (2013). Therapiefrequenz in der ambulanten logopädischen Praxis. *Forum Logopäedie* 27, 12–19. doi: 10.2443/skv-s-2013-53020130202

Baddeley, A. D., Thomson, N., and Buchanan, M. (1975). Word length and the structure of short-term memory. *J. Verb. Learn. Verb. Behav.* 14, 575–589. doi: 10.1016/S0022-5371(75)80045-4

Badecker, W., Miozzo, M., and Zanuttini, R. (1995). The two-stage model of lexical retrieval: evidence from a case of anomia with selective preservation of grammatical gender. *Cognition* 57, 193–216. doi: 10.1016/0010-0277(95)00663-J

Ballard, K. J., Etter, N. M., Shen, S., Monroe, P., and Tien Tan, C. (2019). Feasibility of automatic speech recognition for providing feedback during tablet-based treatment for apraxia of speech plus aphasia. *Am. J. Speech Lang. Pathol.* 28(2S), 818-834. doi: 10.1044/2018_AJSLP-MSC18-18-0109

Bayer, J. (1986). *Die linguistische Bewertung aphasischer Spontansprache: Eine Anleitung für die Praxis.* Aphasie: München, TUDUV, 9–46.

Bergmann, J., and Sams, A. (2012). *Flip Your Classroom. Reaching Every Student in Every Class Every Day*. Michigan: International Society for Technology in Education; ProQuest.

Bhogal, S. K., Teasell, R., and Speechley, M. (2003). Intensity of aphasia therapy, impact on recovery. *Stroke* 34, 987–993. doi: 10.1161/01.STR.0000062343.64383.D0

Blanken, G., Döppler, R., Schlenck, K. J., and Bautz, M. (1999). *Wortproduktionsprüfung*. Hofheim: NAT-Verlag.

Blom Johansson, M., Carlsson, M., and Sonnander, K. (2012). Communication difficulties and the use of communication strategies: from the perspective of individuals with aphasia. *Int. J. Lang. Commun. Disord.* 47, 144–155. doi: 10.1111/j.1460-6984.2011.00089.x

Blomert, L., and Buslach, D. C. (1994). *Amsterdam-Nijmegen Everyday Language Test (ANELT)-Deutsche Fassung*. Lisse: Swets and Zeitlinger.

Boenisch, J., and Sachse, S. K. (2020). "Core vocabulary–importance for daily communication," in *Kompendium Unterstützte Kommunikation*, eds J. Boenisch and S. K. Sachse (Stuttgart: Kohlhammer), 108–116.

Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *Am. J. Speech Lang. Pathol.* 13, 236–249. doi: 10.1044/1058-0360(2004/025)

Braley, M., Pierce, J. S., Saxena, S., Oliveira, D., Taraboanta, E., Anantha, L., et al. (2021). A virtual, randomized, control trial of a digital therapeutic for speech, language, and cognitive intervention in post-stroke persons with aphasia. *Front. Neurol.* 12, 626780. doi: 10.3389/fneur.2021.626780

Breitenstein, C., Grewe, T., Flöel, A., Ziegler, W., Springer, L., Martus, P., et al. (2017). Intensive speech and language therapy in patients with chronic aphasia after stroke: a randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. *The Lancet* 389, 1528–1538. doi: 10.1016/S0140-6736(17)30067-3

Brüsch, J. (2022). *Wie kann die Wirksamkeit einer Wortabruf-Therapie mit der App "Lingo Talk" überprüft werden? Entwicklung eines Studiendesigns*. Bachelor's Thesis. Potsdam: University of Potsdam.

Bürkle, L., Collasius, V., Djuric, A., Heinemann, S., Beushausen, U., Grötzbach, H., et al. (2022). Update Aphasietherapie: Aktuelle evidenzbasierte Empfehlungen für die Sprachtherapie bei Aphasie. *Neurol. Rehab.* 28, 7–15. doi: 10.14624/NR2201001

Conroy, P., Sage, K., and Lambon Ralph, M. A. (2009). Errorless and errorful therapy for verb and noun naming in aphasia. *Aphasiology* 23, 1311–1337. doi: 10.1080/02687030902756439

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93, 283–321. doi: 10.1037/0033-295X.93.3.283

Deutsche Gesellschaft für Neurologie (2011). *Leitlinie "Rehabilitation aphasischer Störungen nach Schlaganfall"*. Berlin: Deutsche Gesellschaft für Neurologie.

Ellis, A. W., Kay, J. A. N. I. C. E., and Franklin, S. (1992). *Anomia: Differentiating Between Semantic and Phonological Deficits. Cognitive Neuropsychology in Clinical Practice.* New York, NY: Oxford University Press, 207–228.

Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., et al. (2006). Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. *Stroke* 37, 1379–1384. doi: 10.1161/01.STR.0000221815.64093.8c

Glindemann, R. (2002). *Bogenhausener Semantik-Untersuchung: BOSU*. Urban & Fischer: Elsevier.

Goodglass, H., Kaplan, E., Weintraub, S., and Ackerman, N. (1976). The "tip-of-the-tongue" phenomenon in aphasia. *Cortex* 12, 145–153. doi: 10.1016/S0010-9452(76)80018-4

Goodglass, H., and Wingfield, A. (1997). "Word-finding deficits in Aphasia," in *Cognitive Neuropsychology in Clinical Practice*, eds D. I. Margolin (New York, NY: Oxford University Press), 3–27.

Haluts, N., Trippa, M., Friedmann, N., and Treves, A. (2020). Professional or amateur? The phonological output buffer as a working memory operator. *Entropy* 22, 662. doi: 10.3390/e22060662

Harley, T. A., and Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *Br. J. Psychol.* 89, 151–174. doi: 10.1111/j.2044-8295.1998.tb02677.x

Hickin, J., Best, W., Herbert, R., Howard, D., and Osborne, F. (2002). Phonological therapy for word-finding difficulties: A re-evaluation. *Aphasiology* 16, 981–999. doi: 10.1080/02687030244000509

Hilari, K., Cruice, M., Sorin-Peters, R., and Worrall, L. (2016). Quality of life in aphasia: State of the art. *Folia Phoniatrica et Logopaedica*,67, 114–118. doi: 10.1159/000440997

Hillis, A. E., and Caramazza, A. (1994). "Theories of lexical processing in the rehabilitation of lexical deficits," in *Cognitive Neuropsychology and Cognitive Rehabilitation*, eds M. J. Riddoch, and G. Humphreys (Hove: Lawrence Erlbaum).

Howard, D. (2000). "Cognitive neuropsychology and aphasia therapy: the case of word retrieval," in *Acquired Neurogenic Communication Disorders: A Clinical Perspective,* ed I. Papathanasiou (London: Whurr Publishers Ltd).

Huber, W., Poeck, K., Weniger, D., and Willmes, K. (1983). *Aachener Aphasie Test (AAT).* Zürich: Hogrefe.

Jakob, H., and Späth, M. (2023). *Neolexon*. München: Limedix GmbH.

Kalbe, E., Reinhold, N., Ender, U., and Kessler, J. (2010). *Aphasie-Check-Liste (ACL)*. Istein: Schulz-Kirchner Verlag GmbH. Available online at: https://www.testzentrale.de/shop/aphasie-check-liste.html

Kittredge, A. K., Dell, G. S., Verkuilen, J., and Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognit. Neuropsychol.* 25, 463–492. doi: 10.1080/02643290701674851

Kurland, J., Wilkins, A. R., and Stokes, P. (2014). iPractice: piloting the effectiveness of a tablet-based home practice program in aphasia treatment. *Seminars Speech Lang.* 35 35, 51–63. doi: 10.1055/s-0033-1362991

Laiacona, M., and Caramazza, A. (2004). The noun/verb dissociation in language production: varieties of causes. *Cognit. Neuropsychol.* 21, 103–123. doi: 10.1080/02643290342000311

Leinweber, J. (2021). App-Einsatz in der Logopädie/Sprachtherapie: Strategien und Kriterien. *Spektrum Patholinguistik* 14:69. doi: 10.25932/publishup-50016

Levelt, W. J. (1993). *Speaking. From Intention to Articulation*. Cambridge: MIT Press.

Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–38. doi: 10.1017/S0140525X99001776

Lorenz, A. (2004). *Die Behandlung von Wortabrufstörungen bei Aphasie: Eine methodenvergleichende Studie zum Bildbenennen (Doctoral dissertation*). Potsdam: Universität Potsdam.

Lorenz, A., and Ziegler, W. (2009). Semantic vs. word-form specific techniques in anomia treatment: a multiple single-case study. *J. Neuroling.* 22, 515–537. doi: 10.1016/j.jneuroling.2009.05.003

Lüttmann, H., Zwitserlood, P., Böhl, A., and Bölte, J. (2011). Evidence for morphological composition at the form level in speech production. *J. Cognit. Psychol.* 23, 818–836. doi: 10.1080/20445911.2011.575774

McRae, K., De Sa, D., and Seidenberg, V. R. (1997). On the nature and scope of featural representations of word meaning. *J. Exp. Psychol. Gen.* 126, 99. doi: 10.1037/0096-3445.126.2.99

Miceli, G., Amitrano, A., Capasso, R., and Caramazza, A. (1996). The treatment of anomia resulting from output lexical damage: analysis of two cases. *Brain Lang.* 52, 150–174. doi: 10.1006/brln.1996.0008

Morrison, C. M., and Gibbons, Z. C. (2006). Lexical determinants of semantic processing speed. *Visual Cognit.* 13, 949–967. doi: 10.1080/13506280544000129

Morton, J. (1969). Interaction of information in word recognition. *Psychol. Rev.* 76, 165–178. doi: 10.1037/h0027366

Nickels, L. (2001). *Spoken Word Production. The Handbook of Cognitive Neuropsychology: What Deficits Reveal About the Human Mind*. London: Psychology Press, 291–320.

Nickels, L. (2002). Therapy for naming disorders: revisiting, revising, and reviewing. *Aphasiology* 16, 935–979. doi: 10.1080/02687030244000563

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M., Bowen, A., et al. (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. *The Lancet Neurol.* 18, 821–833. doi: 10.1016/S1474-4422(19)30192-9

Patterson, K., Shewell, C., Coltheart, M., Sartori, G., and Job, R. (1987). "Speak and spell: dissociations and word-class effects," in *The Cognitive Neuropsychology of Language*, eds M. Coltheart, G. Sartori, and R. Job (London: Lawrence Erlbaum).

Pfeiffer, N., and Leisner, S. (2016). *PlanBe. Planen und Bewerten von Kommunikationssituationen.* Bremen: Rehavista.

Plaut, D. C., and Shallice, T. (1993a). Perseverative and semantic influences on visual object naming errors in optic aphasia: a connectionist account. *J. Cognit. Neurosci.* 5, 89–117. doi: 10.1162/jocn.1993.5.1.89

Plaut, D. C., and Shallice, T. (1993b). Deep dyslexia: A case study of connectionist neuropsychology. *Cognit. Neuropsychol.* 10, 377–500. doi: 10.1080/02643299308253469

Pulvermüller, F., and Berthier, M. L. (2008). Aphasia therapy on a neuroscience basis. *Aphasiology* 22, 563–599. doi: 10.1080/02687030701612213

Renvall, K., Nickels, L., and Davidson, B. (2013a). Functionally relevant items in the treatment of aphasia (part I): challenges for current practice. *Aphasiology* 27, 636–650. doi: 10.1080/02687038.2013.786804

Renvall, K., Nickels, L., and Davidson, B. (2013b). Functionally relevant items in the treatment of aphasia (part II): further perspectives and specific tools. *Aphasiology* 27, 651–677. doi: 10.1080/02687038.2013.796507

Saalfrank, T. (2023). *Evaluation einer app-gestützten Wortabruftherapie im ambulnten Setting - Eine Einzelfallstudie mit LingoTalk* (Bachelor's Thesis). Potsdam: University of Potsdam.

Schmitz-Antonischki, D. (2021). *Therapie von Wortabrufstörungen mit der App LingoTalk bei einer Patientin mit Aphasie* (Bachelor's Thesis). Potsdam: University of Potsdam.

Schröder, A., Gemballa, T., Ruppin, S., and Wartenburger, I. (2012). German norms for semantic typicality, age of acquisition, and concept familiarity. *Behav. Res. Methods*, 44, 380–394. doi: 10.3758/s13428-011-0164-y

Scruggs, T. E., Mastropieri, M. A., and Casto, G. (1987). The quantitative synthesis of single-subject research: methodology and validation. *Remedial Special Educ.* 8, 24–33. doi: 10.1177/074193258700800206

Stadie, N., Cholewa, J., and Bleser, D. e. R. (2013). *LEMO 2, 0. Lexikon modellorientiert: Diagnostik für Aphasie, Dyslexie und Dysgraphie.* Hofheim: NAT-Verlag.

Stark, B. C., and Warburton, E. A. (2018). Improved language in chronic aphasia after self-delivered iPad speech therapy. *Neuropsychol. Rehab.* 28, 818–831. doi: 10.1080/09602011.2016.1146150

Sze, W. P., Hameau, S., Warren, J., and Best, W. (2021). Identifying the components of a successful spoken naming therapy: a meta-analysis of word-finding interventions for adults with aphasia. *Aphasiology* 35, 33–72. doi: 10.1080/02687038.2020.1781419

The REhabilitation and recovery of peopLE with Aphasia after StrokE (RELEASE) Collaborators (2022). Dosage, intensity, and frequency of language therapy for aphasia: a systematic review–based, individual participant data network meta-analysis. *Stroke* 53, 956–967. doi: 10.1161/STROKEAHA.121.035216

Universität Leipzig, Institut für Informatik, Projekt Deutscher Wortschatz. (1998-2023). *Wortschatz Leipzig.* Available online at: https://cls.corpora.uni-leipzig.de/de/deu_newscrawl_2011/3,2.1_The%20Most%20Frequent%2050%20Words.html (accessed July 4, 2023).

Webster, J., Whitworth, A., and Morris, J. (2015). Is it time to stop "fishing"? A review of generalisation following aphasia intervention. *Aphasiology* 29, 1240–1264. doi: 10.1080/02687038.2015.1027169

Wegener, V., Meinhold, M., and Heide, J. (2010). "The whole is more than the sum of its parts: Improving compound naming through full forms and constituents," in *Poster Presented at the Science of Aphasia Conference* (Potsdam).

Wendlandt, W. (2002). *Therapeutische Hausaufgaben. Materialien für die Eigenarbeit und das Selbsttraining. Eine Anleitung für Therapeuten, Betroffene, Eltern und Erzieher*. Stuttgart: Thieme.

World Health Organization (2001). *ICF. International Classification of Functioning, disabilIty and Health.* Geneva: WHO.

Wu, S. (2023). *The Flipped Speech Room. The Flipped Classroom Model for Speech Pathologists*. Available online at https://speechisbeautiful.com/2016/09/the-flipped-speech-room/ (accessed July 4, 2023).