Telerehabilitation advancements: A Computer-based speech/language treatment platform for Greek-speaking people with chronic aphasia and dysarthria

Nerantzini, M., Papakyritsis, I., Christou, Th., Zafeiri, A., Bagioka, D. & Terzi, A.
University of Patras

Abstract
The effective support of people with neurogenic communication disorders requires individualized intervention by speech/language pathologists (SLPs). Apart from traditional interventions, computer-based therapies have also been found to be effective. Speech/language Therapy Platform with Virtual Agent (PLan-V) is a novel, technologically assisted, personalized intervention platform, unavailable until now in Greek. It provides impairment-oriented treatment, offering patients the opportunity to work independently, via the assistance of a virtual SLP. The paper aims to describe the speech and language therapy protocols incorporated within the platform and the set of technological features utilized to facilitate patient performance and provide augmented feedback. Speech activities focus on respiration, phonation, and prosody, whereas language intervention programs include activities at the word and sentence level and target both production and comprehension skills.

Keywords: aphasia, dysarthria, telerehabilitation, speech, language, treatment

Introduction

Direct, face-to-face (FTF) clinical service has been traditionally considered the best intervention practice in the field of speech and language pathology. Effective service delivery requires from the clinician to be able to provide patients with real-time instructions and feedback directly responsive to their communication needs. Nevertheless, for some groups of patients, such as individuals with significant physical disabilities and people that reside in remote areas with limited access to health services, FTF sessions may often require high levels of physical, cognitive, and emotional effort from their part, as well as increased need for caregiver assistance and high financial cost. Remote clinical services are particularly relevant in the context of behavioral rehabilitation of neurogenic communication disorders, given the high levels of unmet needs and service demands.

Telerehabilitation practices involving computer-based intervention have been found to be effective for the treatment of speech and language deficits long before the COVID-19 pandemic made them an urgent necessity (Theodoros & Ramig, 2011; Jokel et al., 2009; Thompson et al., 2010). According to a survey of the American Speech-Language-Hearing Association, completed by 476 Speech-Language Pathologists (SLPs), 64% of clinicians endorse telepractice services and 96.4% of SLPs used it for intervention purposes (ASHA, 2016). Computer-based treatments are often delivered via a computer or a tablet either asynchronously (self-mediated, without the presence of an SLP) or synchronously (with the remote supervision of an SLP), following the same principles of traditional FTF behavioural interventions. Importantly, remote computerized intervention can be feasible and valid, with high levels of satisfaction gains reported from both health clinicians and patient users (Adrian et al., 2003; Fink et al., 2005; Jokel et al., 2009).

In this context, we present in this paper a novel, technologically assisted, personalized intervention platform, developed for the first time in Greek, the *Speech & Language Therapy Platform with Virtual Agent (PLan-V)*. The platform provides impairment-oriented treatment, offering patients the opportunity to work independently, without the physical presence of a clinician, via the assistance of a virtual SLP. The virtual assistant is used to provide oral instructions and feedback, in combination with visual cues (e.g., colour-coded written stimuli, pictures, animations, acoustic representations of speech) as required in each therapy task, via a natural and friendly interface facilitating accurate performance, and indicating incorrect and successful trials. The platform enables SLP users to design and administer remotely individualized intervention plans meant to complement the FTF clinical sessions. Additionally, PLan-V can aid with the time-consuming process of patient assessment and monitoring of the intervention outcome without the need for clinical visits. Among several other services, a system for aphasia detection (Chatzoudis et al., 2022) and automatic assessment of aphasia severity has been developed and integrated within the platform, using state-of-the art speech and language technologies and machine learning techniques to analyze spoken narratives (Stamouli et al., 2023). PLan-V contains original customized speech and language therapy materials, well-grounded on evidence-based intervention practices. In this paper we present the design of the therapy activities incorporated within the platform and the set of cues and technological features available. Intervention protocols are broadly categorized into speech and language therapy tasks, described in detail below.

1. Speech therapy protocols

Individuals with neurogenic communication disorders are often treated for both speech and language difficulties, given that aphasia and dysarthria often co-occur (Duffy, 2019). In many cases, if a patient cannot generate intelligible speech output, work on language cannot commence. For that reason, it was important to incorporate an intervention program targeting speech within the PLan-V platform. Traditionally, the focus of speech therapy has been placed on articulation, whereas suprasegmental aspects of speech production have often been ignored. To this effect, we developed a speech therapy protocol that targets speech breathing, phonation, and intonation. The tasks are based on evidence-based literature regarding the behavioral management of respiratory, phonatory and prosodic dysfunction (Yorkston et al., 2007). Improvements in these areas have a global effect on speech intelligibility and naturalness of patients with any type of dysarthria (Duffy, 2019). The protocol is impairment-oriented and has been developed for speakers with a specific speech profile, i.e., individuals that experience difficulties with respiratory support for speech, breathing control and coordination between phonation and breathing. These patients often run out of breath during speech and speak with weak voice and flat intonation, making the act of speaking laborious and the speech output unnatural and often reduced in intelligibility.

The speech intervention program we have developed is characterized by two key features in line with the basic principles of motor learning (Maas et al., 2008): (a) *task hierarchy*, i.e., tasks are organized in a sequential order of increasing difficulty, (b) *structured presentation of augmented feedback*, i.e., the feedback given in addition to the patient's own feedback. Regarding the latter, the computer interface of the platform provides visual information of both the speech target and the user's trials. Hence, it is possible to draw users' attention to two, distinct, but equally important, aspects of speech, the movement patterns of the vocal mechanism (knowledge of performance), and the movement outcome, i.e., the acoustic signal produced (knowledge of results). Both types of feedback are found to be beneficial. However, placing the focus on knowledge of performance is probably more effective in the early stages of treatment, when patients cannot yet distinguish correct from incorrect productions (Maas et al., 2008).

The intervention program consists of three, interrelated, activities: (a*) maximum vowel prolongation tasks using modal voice*, (b) *vowel prolongation tasks using non typical pitch, high, low, and varying*, and (c) *intonation tasks targeting sentence focus*.

In activity (a) the goal is to provide training on the use of behavioral strategies that aim to modify patterns of two types of speech movements: i) the speech breathing pattern using a variation of the inspiratory checking technique (Netsell, 1998), i.e., "slow, controlled expiration" focusing on the movement of the diaphragm, and ii) the resonance pattern by increasing the degree of mouth opening using the "open mouth" technique (Kummer, 2011). Both movements are introduced and illustrated via oral instructions, pictures, and animations, and trained in a simple, speech-like task, i.e., vowel prolongation, that requires the maintenance of a static posture of the vocal tract. The goal is to maintain the vowel for as long as possible, with steady and loud voice, without shouting, by increasing awareness of the speech movements, allowing for the active control of the diaphragm and the mouth. The protocol consists of five distinct sections:

1) Two baseline vowel prolongation trials from which the following acoustic measures are extracted: vowel duration, F_0 variation, intensity variation and mean intensity.
2) Training of the “slow, controlled expiration” technique during silent breathing
3) Training of the “slow, controlled expiration” technique during speech, using a vowel prolongation task.
4) Training of the “open mouth” technique using the vowel prolongation task.
5) Practice of the simultaneous use of both the “slow, controlled expiration” and the “open mouth” techniques during vowel prolongation.

The patient users can watch the target movements before each trial and receive real-time visual feedback of pitch, duration, and loudness, via an acoustic application. In sections 2-4 patients carry out a minimum of two trials and in section 5 the target is to obtain six consecutive vowel prolongation attempts. Again, similarly to baseline, (section 1) the acoustic measures of vowel duration, F_0 variation, intensity variation and mean intensity are collected. These measures and the actual vowel prolongation recordings can be accessed by the clinician user of Plan-V to monitor the patient’s progress.

The second activity (b) focuses on the flexibility of the vocal tract via the following exercises: i) vowel prolongation with stable high pitch, ii) vowel prolongation with stable low pitch, iii) production of a series of short vowels with progressively higher pitch levels, iv) vowel prolongation gliding from habitual to maximum pitch and back to habitual. Each of these exercises is illustrated via visual and acoustic modelling and the user also receives real time feedback during and following each trial. For each exercise, the goal is for the patient to carry out a minimum of two trials. The clinician user has access to both the patient’s attempts and acoustic measures of mean pitch and pitch variation.

The third activity (c) targets the linguistic use of prosody, in an attempt to establish natural intonation patterns in connected speech. It involves a production and a comprehension sentence focus task. At the production level, the user is required to produce sentences of varying complexity, following a specific focus pattern which is demonstrated via oral and written modes, both via direct imitation and via priming. The user receives real-time acoustic feedback and oral feedback following each attempt. At the comprehension level, the users again read and listen to sentences produced with a specific part highlighted and they must identify the word(s) in focus via a written multiple-choice paradigm. After a failed attempt, an additional cue is given and the patient carries out a second, final trial. Although, typically, target patients are not expected to demonstrate perceptual difficulties, therapy of disturbed production of prosody relies on adequate reception of prosody (Martens et al., 2016). For both exercises, target items are categorized in terms of complexity by means of sentence length. Four categories are included: i) adjective-noun phrases, ii) subject-verb sentences, iii) subject-verb-object (SVO) sentences and iv) sentences with a direct and an indirect object (SVO_1O_2). Within each category, items are further distinguished by mean of word length into sentences/phrases made up of disyllabic content words or polysyllabic words. 150 items were included in each of the 4 categories, making up a total of 600 sentence items. In the production task, the SLT has access to both the patients' actual trials, the generated F_0 track, as well as relevant quantitative data, i.e., the F_0 and intensity variation of the stressed syllable of the word in focus. In the comprehension task, the platform generates the percentage of items in which the focused word was accurately identified. The clinician has also access to the exact patient responses to identify possible item and cue effects.

Overall, clinicians combine the above activities and exercises to develop individualized treatment plans for their clients, adjusting and modifying the difficulty level, based on dysarthria severity and the patient's ongoing performance.

2. Language therapy protocols

Original materials were designed to address individuals with language difficulties of various types and severity degrees. Both language comprehension and production are targeted, covering different linguistic levels (phonology, lexicon, morphosyntax). Each task incorporates a ranking according to difficulty/complexity, depending on individual therapy needs. Language therapy protocols, thus, are designed to enhance lexical- and sentence-level processes, including semantic processing, word retrieval, and sentence construction tasks.

2.1. At the word level

2.1.1. Lexical-Semantic processing tasks

Semantic processing tasks are used in treatment to improve both comprehension and production of nouns, verbs, adjectives, and adverbs. Several studies have shown that such tasks facilitate word retrieval, via the amplification of meaning-form links and the strengthening of lexical access. Tasks such as the auditory word-to-picture matching, written word-to-picture matching, picture-to-spoken description matching, answering yes/no questions about the target stimuli, picture and spoken word categorization, and relatedness judgment tasks have been used to strengthen the associations in the semantic network (Boyle & Coehlo, 1995; Drew & Thompson, 1999; Wambaugh et al., 2001; Howard et al., 2006). Although these tasks are primarily (lexical) semantic in nature, they also target phonological representations, providing a combination of semantic-phonologic cueing. Our materials were developed in a way that could potentially discriminate these two independent processes. The semantic processing protocol of PLan-V includes 10 tasks covering different types of stimuli (oral, written, and visual) and are presented below.

The *auditory word-to-picture matching task* (a) consists of 6 noun and 4 verb categories. Noun categories consist of i) 25 target animals, ii) 30 food items, iii) 18 body parts, iv) 24 household objects, v) 20 animate professions and vi) 25 other objects. Verb categories include the following verb types: i) 42 verbs with one argument, ii) 27 verbs with two arguments, of which the internal one is not obligatory, iii) 46 verbs with two obligatory arguments, and iv) 14 verbs with three arguments. For every target item, two phonological, two semantic and two unrelated distractors were included. All items were matched for length and word frequency using two Greek corpora (GreekLex: Kyparissiadis et al., 2017; HelexKids: Terzopoulos et al., 2016). Color drawings were developed by a professional designer and picture norming took place collecting ratings of 45 neurotypical adults between the ages of 45 and 75 on picture name agreement and picture visual complexity. Patients are asked to select the target picture that matches the auditory word, selecting from three-to-seven pictures. In cases where a non-target item is selected, appropriate feedback is provided (semantic or phonological, depending on the error). The exact same items were used for the development of the (b) *written word-to-picture matching* and the (c) *picture-to-written word matching* tasks, where users are asked to either select the target picture that matches the visually presented written word, or the written word that matches the presented picture.

A few more (lexical) semantic tasks were developed involving processing of auditory information such as the *picture-to-spoken description matching task* (d), where patients are asked to select the target picture that matches the auditory description. This task consists of 5 noun categories, i) 50 target animals, ii) 53 food items, iii) 23 body parts, iv) 54 household objects and v) 55 other objects. Additionally, the *auditory synonym matching task* (e) and the *auditory antonym matching task*s (f) consist of 3 categories including i) 71 adjectives, ii) 65 adverbs and iii) 70 verbs. In these tasks individuals are asked to match the auditorily presented item with a synonym or an antonym target word respectively; an antonym or a synonym and two unrelated items of the same category are presented as distractors.

Within the semantic processing domain, we also included two tasks that rely upon category-based processes: (g) *the categorical sorting task* and (h) *the semantic odd-one-out judgment task*. Boyle (2017) suggested that concepts that share semantic features are more strongly related than those that do not. Semantic distance across categories is also a strong predictor of category organization and lexical access (Vigliocco et al., 2002). The categorical sorting task (superordinary categorization) was designed to strengthen the semantic network using pictures or written words. Patients are asked to sort a series of pictures or written words into categories based on conceptual distinctions according to their superordinate category (i.e., animals and vehicles). The task consists of 5 noun categories: i) 144 animals, ii) 156 food items, iii) 66 body parts, iv) 141 other objects and v) 153 household objects. Three levels of difficulty are included: the easy level with 3 stimuli (2 within the same category), the medium level with 4 stimuli (2 within the same category) and the advanced level with 5 stimuli. Each array consisted of two target words embedded in a set of semantically unrelated words. Within each difficulty level three types of associations were taken into consideration: high associations (e.g., related: *cat, dog,* unrelated: *glass*), low associations (e.g., related: *hen, pig,* unrelated: *fridge*) and within category associations (e.g., related: *fly, mosquito,* unrelated: *whale*). The semantic odd-one-out judgment task (Crutch et al., 2009) is organized in a similar manner. Typically, the time taken to identify the odd word out reflects not only the time needed to establish the semantic associations but also the time required to establish a mental model that links the remaining words. The odd-one-out task comprises of 4 noun categories and three levels of difficulty. Specifically, we included: i) 96 animals, ii) 104 food items, iii) 90 other objects and iv) 102 household objects. The easy level has 3 pictures (1 from a different category), the medium level has 4 pictures (2 from a different category) and the advanced level has 5 pictures (3 from a different category). Each array consisted of one target word embedded in a set of semantically related words. Finally, the semantic processing category included two pure phonological tasks i) *the initial syllable-to-picture matching task* and j) *the rhymes-to-picture matching task*. Both tasks consist of 6 noun categories with a total of 142 target items, same as those used in the auditory word-to-picture matching task. However, in the phonological tasks, participants are asked to select the target picture that either matches the word starting with the provided initial syllable, or the word that rhymes/ends with the provided syllable, respectively.

In these tasks, if a non-target item is selected, individuals can re-listen to the instructions and try again. A time limit for task completion can be established from the SLP. Nickels (2002) suggested that "repetition priming" enhanced target word elicitation as it strengthens the mapping from the semantic representation to the phonological form and thereby improves subsequent retrieval success. Therefore, the correct answer is always provided after two erroneous trials in our treatment tasks, based on the premise that when a correct word form is provided, both semantic and phonological forms are activated, increasing the likelihood of correct word retrieval on a subsequent attempt.

2.1.2. Word retrieval tasks

The language therapy protocol also includes word finding activities. To facilitate naming, different cue types are provided, such as phonological (Leonard et al., 2008), orthographical (Basso et al., 2001) and semantic (Boyle, 2017). According to Doesborgh et al. (2004) a combination of these cues can be used to activate the target word. Three-word retrieval therapy protocols are integrated into the platform: (a) the *multicue therapy* (Doesborgh et al., 2004), (b) the *semantic feature analysis* (Boyle, 2017), and (c) the *phonological feature analysis*.

The *multicue therapy protocol* consists of 6 noun categories with 386 target items and 3 verb categories with 136 target items in total. Nouns include: i) 50 target animals, ii) 53 food items, iii) 23 body parts, iv) 50 household objects, and v) 57 other objects and v). Verb categories include different verb types: i) 44 verbs with one argument, ii) 29 verbs two arguments, where the internal argument is not obligatory, iii) 48 verbs with 2 obligatory arguments, and iv) 15 verbs with 3 arguments. Patients are asked to name a visually presented picture and their response is recorded. In case of uncertainty or erroneous response, participants are instructed to replay the recording and make a second attempt. The interface includes four help buttons that provide different types of cues: a) written definition, b) phonological cueing – initial syllable, c) written word form and d) sentence completion. There is no fixed, pre-conceived cueing hierarchy, but instead, the user is free to select the most informative cue to complete his partial knowledge and facilitate word retrieval.

Semantic feature analysis (SFA**)** requires participants to generate semantic features related to object nouns and verbs. A picture is placed at the center of the screen and participants are asked to name it and their responses is recorded. Following their initial naming attempt, participants are asked to answer a set of questions about 6 semantic features associated with the target picture (e.g., superordinate category, use, physical properties, location, association, and sentence generation). Target's most distinguishing features were included, to ensure their strict association with the target item over other category members. Target responses for all features appear in the appropriate box. To elicit feature generation participants are prompted with questions such as "*What do you use it for*?" and a list of 4 possible answers is presented, from which participants need to select the target response. In case of an error, feedback is provided. Following feature prompting, participants are asked to review, repeat, and generate them independently. SFA was adapted to enhance verb retrieval (Wambaugh & Ferguson, 2007), and semantic features were modified to include category, subject of the associated verb phrase, purpose of action, associated body part, association with other action, and sentence generation. A picture of the target verb was provided. A total of 258 object nouns and 101 verbs are included as target items.

The Phonological Feature Analysis (following Leonard et al., 2008; Coelho et al., 2000) follows the same procedure as the SFA but participants are asked to identify phonological components related to the target item. We combined both phonological and orthographic cueing, since both types of features equally facilitate word retrieval in individuals with aphasia, and patients are expected to be literate. A set of 6 phonological and orthographic features associated with the target picture are provided, including number of syllables, first syllable cueing, final syllable cueing/rhyme, rhyming words, first grapheme cueing and final grapheme cueing. A total of 250 object nouns and 136 verbs were included as target items. Pure phonological and orthographic cues have effectively been used in some studies for the treatment of word finding difficulties in aphasia (Best et al., 2002; Hickin et al., 2002; Herbert et al., 2003; Leonard et al., 2008). Numbers of items will be rounded after piloting.

2.2. At the sentence level

2.2.1. Sentence comprehension tasks

Individuals with neurogenic communication disorders often face difficulties in understanding semantically reversible non-canonical sentence types (e.g., object relative clauses) as opposed to canonical sentence types (e.g., subject relative clauses) (c.f., Caramazza & Zurif, 1976), or passive sentences as opposed to active sentences (Grodzinsky, 1986). For the treatment of sentence comprehension deficits two tasks were developed: (a) *the sentence comprehension task without pictures* and (b) *the auditory sentence-to-picture matching task*. *In the sentence* *comprehension task without pictures*, 6 categories with a total of 130 items were targeted: sentences with non-active verbs surfacing in passive (40 items, with animate and inanimate subjects, e.g., *The man with long hair was arrested by a police officer in the evening. The storybook was translated very easily by the teacher.*), sentences with prototypically reflexive verbs (10 items, e.*g., The woman with the bandage was washed by the nurse in the evening.*), anticausative structures (20 items, e.g., *The brightly colored awning was torn by the wind in the afternoon.*), active sentences (20 items), and subject and object relative sentences (40 items). Participants are auditorily presented with the sentence and they are subsequently provided with a comprehension question related to *θ*-role interpretation (*who* did *what* to *whom*). In case of an error, feedback is provided.

In the *auditory sentence-to-picture matching task*, 5 categories comprising a total of 240 items are targeted: 40 sentences with non-active verbs in passives, 80 subject and object relative sentences, 40 simple active sentences, 40 sentences with focused objects and non-canonical word order (OVS), and 40 sentences with clitic doubling. A target picture depicting the actions in the sentence and a foil picture depicting the same items and actions with reversed thematic roles is presented at the same time. The position of the target picture on either side of the screen is counterbalanced across trials. Participants are asked to listen carefully and select the picture that matches the auditorily presented sentence. In case of an error, feedback is provided *"Please pay more attention to who did what to whom.*" After that, the foil picture is removed and participants are asked to identify the target verb, the doer and the receiver of the action. Then, the virtual assistant repeats the sentence once again and provides further explanation of the *θ*-roles in the picture (e.g., "*Be careful, in this sentence, the receiver of the action appears at the beginning of the sentence*"). In the final step, the target sentence is repeated while both the target and the foil picture are provided. Participants are asked again to choose which picture best matches the sentence they hear.

2.2.2. Sentence production tasks

Three tasks were designed: (a) *the sentence production priming task,* (b) *an adaptation of the treatment of underlying forms (TUF)* and (c) *a word sorting/anagram task*.

In the *sentence production priming* task (Cho & Thompson, 2012), a total of 120 sentences were targeted, 20 for each of the following structures: simple active, focus constructions with reversible order (OVS), sentences with clitic doubling, passives, subject and object relative sentences. A pair of semantically reversible action pictures is provided, while a prime sentence is auditorily presented for the picture on the left side. Participants are asked to produce a sentence of the same structure for the picture on the right side. Following that, participants are asked to identify the target verb, the doer and the receiver of the action in response to the target sentence. In the final step, participants are asked to make grammaticality judgments in sentences with reversible arguments or agreement errors.

The computerized version of the *Treatment of Underlying Forms* (TUF) (following Thompson et al., 2010; Ballard & Thompson, 1999; Thompson & Shapiro, 1995) was adapted for facilitating the production, but also comprehension, of complex sentences. Training focused on the production of object relative sentences: twenty such sentences were developed, using 20 transitive verbs controlled for frequency, imageability, and length. Two pictures were developed, one corresponding to the target sentence and one depicting its semantically reversed. Structural priming has been shown to facilitate syntactic production in individuals with neurogenic disorders and aphasia. Structural priming refers to the tendency of language users to produce syntactic structures that they have previously encountered in speech (Pickering & Ferreira, 2008). In the sentence-production priming task developed participants hear an object relative sentence (prime) while seeing the picture that corresponds to it and are asked to produce another object relative sentence (its semantically reversible) presented in another picture). Participants then go through a series of steps to derive the surface form of the object relative sentence, while, in the final step a sentence-picture matching task is used to assess comprehension.

The last sentence construction task is a *word sorting task,* also known as *Anagram*. It has been designed for aphasic patients with significant language production difficulties that are often further complicated by word retrieval issues and interfering working memory deficits (Thompson et al., 2012; Weintraub et al., 2009). These tasks require participants to sequentially sort cards with written words to generate meaningful sentences, i.e., “*Please build a sentence using these word cards*”. A total of 80 items are designed including simple reversible and irreversible canonical SVO sentences (with lexical items in singular and plural form). Numerous studies have suggested that inflectional morphology is deficient in aphasia, with access to temporal information identified as a source of difficulty in agrammatic production (Bastiaanse, 2008; Faroqi-Shah & Thompson, 2007). Recently, Faroqi-Shah, (2008) showed that agrammatic individuals benefited maximally from practice in mapping tense features (e.g.,

+PAST, +PRESENT, +FUTURE) onto semantically matched verb forms (*folded*, *folds*, *will fold*). Thus, trials addressing tense and aspectual marking are also included. In the tense marking condition participants are presented with a temporal adverb for the requested sentence (e.g., *Yesterday* ____, *Tomorrow* ____) to cue target sentence production. Participants choose the relevant written cards (with the appropriate verb inflection) to generate a grammatical sentence (e.g., *the nurse/washes/washed/will wash/the patient*). A total of 80 items were designed to map tense features (e.g., +PAST, +FUTURE) to reversible and irreversible SVO sentences (with lexical items in singular and plural form). Both regular and irregular verbs are included. Similarly, a total of 120 items addressing aspectual marking are included targeting perfective and imperfective distinctions in the marking of past or future events. Again, the initial words for the requested sentence were provided to facilitate target sentence generation (e.g., *Yesterday perpetually/ immediately*____, *Tomorrow perpetually/ immediately* ____).

3. Scoring Procedure

Clinicians can monitor patient progress through a comprehensive learning analytics system. Recorded tasks, quantitative scores, including response accuracy, reaction time and acoustic measures, enable clinicians to choose and sequence content and activities, depending on participants' performance, providing control, support planning and monitoring. Clinicians can develop individualized treatment plans for their clients (treatment phase used different word sets), can customize the difficulty level of the tasks while monitoring participants' progress both on the whole battery of tasks and on each task individually so that specific patterns of difficulties can be identified/emerge. Feasibility measures include evaluation of the remote technical setup and support, while acceptability is evaluated via a Likert rating scale, measuring participant engagement and satisfaction with the remote setting. The virtual clinician provides non-specific feedback at regular intervals in the form of verbal praise, additionally to specific feedback of response accuracy for the comprehension tasks. In the production tasks, voice recording begins and ends by selecting the "microphone" icon and it is submitted to the system by pressing the "next task" icon. All responses are automatically tracked by the computer for data analysis purposes. Data are automatically transcribed and processed with Natural Language Processing (NLP) tools available for Greek (Prokopidis & Piperidis, 2020), using state-of-the art speech and language technologies and machine learning techniques to perform accurate syntactic parsing. For reliability purposes, SLPs have access to the transcribed responses, which can be edited, and accuracy scores can be modified.

4. Discussion and Conclusions

People with neurogenic communication disorders require individualized intervention by SLPs for improvement of their communication needs. The quality and frequency of the intervention is crucial for optimal results. However, resources, including health insurance benefits, are usually scarce. Moreover, many beneficiaries reside in geographically remote locations, and it is not easy for them to commute to rehabilitation centers. The speech and language platform we have presented here aims to cover these needs, and is the first of its type for Greek.

The platform, designed by an expert team of speech pathologists and linguists, includes both speech and language tasks that are based on internationally well-grounded evidence-based intervention practices. The various types of tasks exceed by far the number of tasks available in conventional intervention protocols, and the number of items in each task is also much more extensive than in conventional intervention. The platform offers the individuals in need of speech and language intervention the opportunity to practice independently, with the guidance of an avatar, but the tasks are assigned to them by the responsible SLP, who monitors their progress remotely and modifies the intervention plan accordingly. The efficacy of the platform and the tasks remains to be evaluated via the forthcoming results from the individuals with neurogenic disorders who are currently acting as the pilot group.

Acknowledgments

This research has been co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship, and Innovation, under the call RESEARCH-CREATE-INNOVATE (project code: T2EDK-02159, "A Speech and Language Therapy Platform with Virtual Agent" - PLan-V, PI: Arhonto Terzi).

References

Adrian, J., Gonzalez, M., & Buiza, J. (2003). CASE STUDY. The use of computer-assisted therapy in anomia rehabilitation: A single-case report, *Aphasiology*, *17*: 981-1002, DOI: 10.1080/02687030344000256

American Speech-Language-Hearing Association (2016). *2016 SIG 18 Telepractice Survey Results*. Available online at: https://www.asha.org/siteassets/practice- portal/ telepractice/2016-telepractice-survey.pdf (accessed April 13, 2022).

Ballard, K. J., & Thompson, C. K. (1999). Treatment and generalization of complex sentence production in agrammatism. *Journal of Speech, Language, & Hearing Research*, *42*: 690–707. doi:10.1044/jslhr.4203.690

Basso, A., Marangolo, P., Piras, F., & Galluzzi, C. (2001). Acquisition of New “Words” in Normal Subjects: A Suggestion for the treatment of Anomia. *Brain & Language, 77*: 45-59.

Bastiaanse, R. (2008). Production of verbs in base position by Dutch agrammatic speakers: Inflection versus finiteness, *Journal of Neurolinguistics, 21*(2):104-119.

Best, W., Herbert, R., Hickin, J., Osborne, F., & Howard, D. (2002). Phonological and orthographic facilitation of word-retrieval in aphasia: Immediate and delayed effects, *Aphasiology*, *16*(1-2): 151-168, DOI: 10.1080/02687040143000483

Boyle, M. (2017). Semantic Treatments for Word and Sentence Production Deficits in Aphasia. *Seminars in Speech & Language*, *38*(01): 052-061 DOI: 10.1055/s-0036-1597256

Boyle, M., & Coelho, C. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology, 4:* 94-98.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain & language*, *3*(4): 572-582.

Chatzoudis, G., Plitsis, M., Stamouli, S., Dimou, A. L., Katsamanis, A., & Katsouros, V. (2022). Zero-Shot Cross-lingual Aphasia Detection using Automatic Speech Recognition. *arXiv preprint arXiv:2204.00448*.

Cho-Reyes S., & Thompson, C.K. (2012). Verb and sentence production and comprehension in aphasia: Northwestern Assessment of Verbs and Sentences. *Aphasiology, 26*(10): 1250–1277.

Coelho, C. A., McHugh, R. E., & Boyle, M. (2000). Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology*, *14*(2): 133-142.

Crutch, S., Connell, S., & Warrington, E. (2009) The different representational frameworks underpinning abstract and concrete knowledge: Evidence from odd-one-out judgements, *The Quarterly Journal of Experimental Psychology, 62* (7): 1377-1390, DOI: 10.1080/17470210802483834

Doesborgh, S., van de Sandt-Koenderman, M., Dippel, D., van Harskamp, F., Koudstaal, P., & Visch-Brink, E. (2004). Cues on request: The efficacy of Multicue, a computer program for wordfinding therapy. *Aphasiology*, *18*(3): 213-222.

Duffy, J. R. (2019). *Motor speech disorders: substrates, differential diagnosis, and management.* 4th ed. St. Louis, Mo.: Elsevier Mosby.

Drew, R. L., & Thompson, C. K. (1999). Model-based semantic treatment for naming deficits in aphasia. *Journal of Speech, Language, and Hearing Research, 42*(4): 972–989. https://doi.org/10.1044/jslhr.4204.972

Faroqi-Shah, Y., & Thompson, C. K. (2007). Verb inflections in agrammatic aphasia: Encoding of tense features. *Journal of Memory & Language*, *56*(1): 129-151.

Faroqi-Shah, Y. (2008). A comparison of two theoretically driven treatments for verb inflection deficits in aphasia. *Neuropsychologia*, *46*(13): 3088-3100.

Fink, R., Brecher, A., Sobel, P., & Schwartz, M. (2005). Computer-assisted treatment of word retrieval deficits in aphasia. *Aphasiology, 19*(10-11): 943-954.

Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language 27*: 135–159.

Herbert, R., Best, W., Hickin, J., Howard, D., & Osborne, F. (2003). Combining lexical and interactional approaches to therapy for word finding deficits in aphasia. *Aphasiology, 17*(12): 1163-1186.

Hickin, J., Best, W., Herbert, R., Howard, D., & Osborne, F. (2002). Phonological therapy for word-finding difficulties: A re-evaluation. *Aphasiology, 16*(10-11): 981-999.

Howard, D., Hickin, J., Redmond, T., Clark, P., & Best, W. (2006). Re-visiting "semantic facilitation" of word retrieval for people with aphasia: Facilitation yes but semantic no. *Cortex, 42*(6): 946-962.

Jokel R., Cupit, J., Rochon, E., & Leonard, C. (2009). Relearning lost vocabulary in non-fluent progressive aphasia with MossTalk Words. *Aphasiology, 23* (2):175-191.

Kyparissiadis, A., van Heuven, W.J.B., Pitchford, N.J., & Ledgeway, T. (2017). GreekLex 2: A comprehensive lexical database with part-of-speech, syllabic, phonological and stress information. *PLoS ONE, 12*(2), e0172493. doi: 10.1371/journal.pone.0172493

Kummer, A. W. (2011). Perceptual assessment of resonance and velopharyngeal function. *Seminars in Speech & Language*, *32*: 159–167.

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

Maas, E., Robin, D. A., Hula, S. N A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of Motor Learning in Treatment of Motor Speech Disorders. *American Journal of Speech - Language Pathology*, *17* (3): 277-298.

Martens, H., Van Nuffelen, G., Wouters, K., De Bodt, M. (2016). Reception of Communicative Functions of Prosody in Hypokinetic Dysarthria due to Parkinson's Disease. *Journal of Parkinson's Disease, 6*(1):219-29. doi: 10.3233/JPD-150678. PMID: 26889630.

Netsell, R. W. (1998). Speech Rehabilitation for Individuals with Unintelligible Speech and Dysarthria: The Respiratory and Velopharyngeal System. *Journal of Medical Speech-Language Pathology*, *6*(2): 107-110.

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

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological bulletin*, *134*(3): 427.

Prokopidis, P., & Piperidis, S. (2020). A Neural NLP toolkit for Greek. In *11th Hellenic Conference on Artificial Intelligence* (pp. 125-128).

Stamouli, S., Nerantzini, M., Papakyritsis, I., Katsamanis, A., Chatzoudis, G., Dimou, A.-L., Plitsis, M., Katsouros, V., Varlokosta, S. & Terzi, A. (2023). A web-based application for eliciting narrative discourse from Greek-speaking people with and without language impairments. *Frontiers in Communication, 8*: 919617. doi: 10.3389/fcomm.2023.919617

Terzopoulos, A., R., Duncan, L.G., Wilson, M.A.J., Niolaki, G.Z., & Masterson, J. (2016) ‘HelexKids: A word frequency database for Greek and Cypriot primary school children’. *Behavior Research Methods*. Available from http://dx.doi.org/10.3758/s13428-015-0698-5

Theodoros, D., & Ramig, L. (2011). Telepractice supported delivery of LSVT LOUD. *Perspectives on Neurophysiology and Neurogenic Speech & Language Disorders, 21* (3):107-119.

Thompson, C. K., & Shapiro, L. P. (1995). Training sentence production in agrammatism: Implications for normal and disordered language. *Brain & Language*, *50*(2): 201-224.

Thompson, C. K., Choy, J., Holland, A. and Cole, R. (2010). Sentactics®: Computer-automated treatment of underlying forms. *Aphasiology, 24* (10): 1242 - 1266.

Thompson, C. K., Weintraub, S., & Mesulam, M. M. (2012). Northwestern anagram test (NAT).

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3): B61-B69.

Wambaugh, J. L., & Ferguson, M. (2007). Application of semantic feature analysis to retrieval of action names in aphasia. *Journal of Rehabilitation Research & Development, 44:* 381–394.

Wambaugh, J., Linebaugh, C., Doyle, P., Martinez, A., Kalinyak-Fliszar, M. & Spencer, K. (2001). Effects of two cueing treatments on lexical retrieval in aphasic speakers with different levels of deficit. *Aphasiology, 15* (10-11): 933-950, DOI: 10.1080/02687040143000302

Weintraub, S., Mesulam, M. M., Wieneke, C., Rademaker, A., Rogalski, E. J., & Thompson, C. K. (2009). The northwestern anagram test: measuring sentence production in primary progressive aphasia. *American Journal of Alzheimer's Disease & Other Dementias*, *24*(5), 408-416.

Yorkston, K. M., Hakel, M., Beukelman, D. R., & Fager, S. (2007). Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review. *Journal of Medical Speech-Language Pathology*, *15*(2): xi-xxxvi.