

PLan-V: A computer-based speech/language therapy platform for people with chronic neurogenic communication disorders.

Nerantzini Michaela, Dimou Nasia, Papakyritsis Ioannis,
Stamouli Spyridoula, Varlokosta Spyridoula & Terzi Arhonto



Co-financed by Greece and the European Union

Background

- The effective support of people with neurogenic communication disorders requires individualized, systematic and regular intervention by speech/language therapists (SLTs).
- Apart from traditional interventions, **computer-based therapies** have been found to be effective in the treatment of **voice and speech disorders** (Halpern et al., 2012), as well as in **word- and sentence-level deficits** (Jokel et al., 2009; Thompson et al., 2010).
- We present the design of a comprehensive clinician-guided computer therapy program addressing the needs of Greek patients.



The present study



The project "A Speech and Language Therapy Platform with Virtual Agent" (**PLan-V**), aims at developing a novel, integrated, technologically assisted, speech/language intervention platform that can support the self-management of people with chronic neurogenic communication disorders.

The PLan-V platform provides for the first time in Greek,

- impairment-oriented treatment,
 - clients can practice independently in their own environment, without the physical presence of a clinician,
 - with the assistance of a virtual agent (Avatar).
-



Original and rich speech and language therapy material



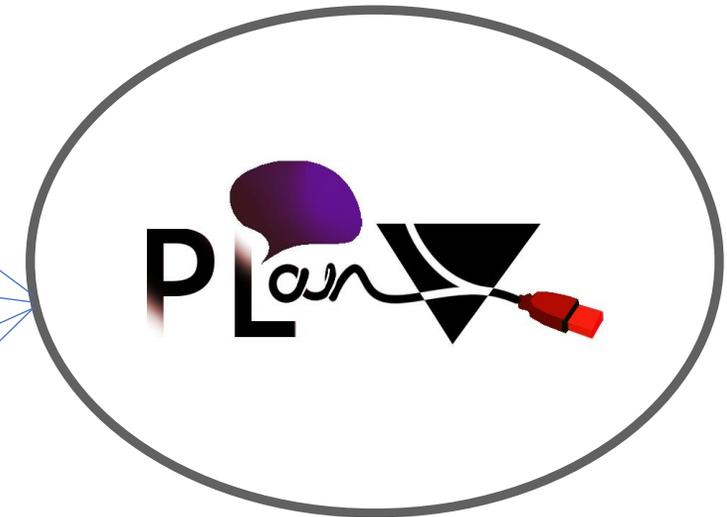
Innovative system of technologically-assisted intervention



Management of content and user data



Automatic assessment of aphasia severity



These products will be integrated in a comprehensive platform, available as a web and as a smartphone application



- The material addresses individuals with **speech** and **language difficulties** of various types and severity degrees.
 - It covers language **comprehension** and **production** at different linguistic levels (phonology, lexicon, morphosyntax).
 - There is a **ranking of tasks according to degrees of difficulty**, depending on individual therapy needs.
-



Original and rich speech and language therapy material

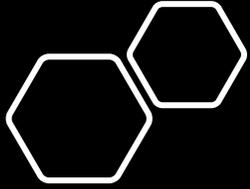


Speech therapy protocols

- Target population: individuals with progressive and non-progressive dysarthria
 - Many patients with neurogenic communication disorders are treated for both speech and language difficulties (e.g., Duffy, 2019)
 - Speech tasks are divided into three main categories:
 - (a) **Warm-up** exercises
 - (b) **Production** exercises that focus on a specific speech system
 - (c) **Perception** exercises that focus on a specific speech system
-

Warm-up exercises

- They are carried out in the beginning of every session
 - Two types of exercises based on level of difficulty:
 - A. Exercises that require the **maintenance** of a **static** posture of the vocal tract
 - i. Maximum prolongation of a voiceless sound (voiceless fricative /s/)
 - ii. Maximum prolongation of a voiced sound (low vowel /a/)
 - iii. Maximum prolongation of a voiced sound (high vowel /i/)
 - B. Exercises that require **flexibility** of the vocal tract
 - i. Production of a series of /a/ with progressively higher **pitch levels**
 - ii. Prolongation of /a/ with rising pitch: gliding from habitual to maximum **pitch**
 - iii. Production of a series of /a/ with progressively higher **loudness levels**
 - iv. Prolongation of /a/ with rising pitch: gliding from habitual to maximum **loudness**
-



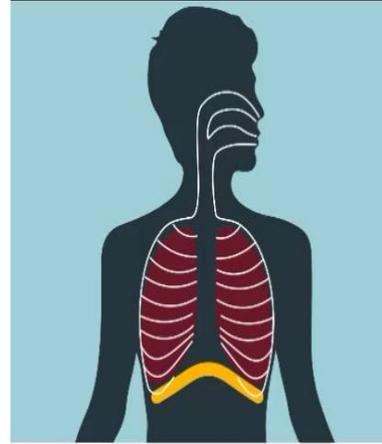
Warm-up exercise example: vowel prolongation

The task is divided into five parts:

1. **Baseline:** task is performed without facilitation or feedback
2. **Preparatory / breathing** exercises: mechanics of silent breathing
3. Speech technique A: **slow, controlled expiration during speech** (“inspiratory checking” Netsell, 1995)
4. Speech technique B: **increased mouth opening** (Kummer, 2011)
5. Main part: **technique application** with **real-time feedback** of **pitch, duration** and **loudness**

Ασκηση Ομιλίας

Έναρξη



Goals:

- Maximum vowel prolongation with loud (**not shouting**) and **steady** voice
- Change of the speech **breathing pattern** by increasing awareness of the “movements of respiration” focusing on the **diaphragm** and the active control of **air flow** via visualizations
- Change of **resonance pattern** by increasing awareness and by actively controlling the **degree of mouth opening**

Production exercises that focus on a specific speech system

The following areas are targeted:

- i. **Lexical stress:** ability to convey stress location
 - ii. Intonation (**focus**): ability to highlight different parts of a message
 - iii. Intonation (**sentence type**): ability to convert a statement into a question
 - iv. **Emotional prosody:** ability to signal emotional state via speech
- Manipulation of **pitch, duration** and **loudness** at the **word** and **sentence** levels
 - Each area will include a maximum of three levels of difficulty
 - e.g., for lexical stress: a. single words, b. minimal word pairs, c. phrases
 - Exercises are based on **direct imitation, visualizations** and **real-time feedback**
-

Perception exercises that focus on a specific speech system

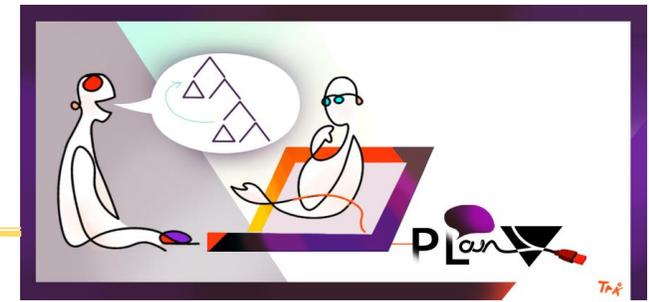
- The same four areas are targeted
- Clients are asked to **indentify** the focused information, sentence type, emotional state, stress position
- Although target clients do not typically have perceptual difficulties, these exercises help individuals:
 - develop **internal feedback**
 - successfully **judge** online their **own** productions
 - depend less on **external** (clinician's) feedback and **generalize** accurate speech production

OVERALL

- Improvement on these tasks has a global effect on speech intelligibility and naturalness of patients with any type of dysarthria
 - The tasks used are based on evidence-based literature regarding the behavioral management of respiratory, phonatory and prosodic dysfunction (Yorkston et al., 2003; 2007)
-



Original and rich speech and language therapy material



Language therapy protocols are designed to target both comprehension and production and enhance lexical and sentence-level processes.

The software includes:

- (a) single-word production/naming tasks
 - (b) word/sentence comprehension tasks
 - (c) sentence generation activities, for various types and aphasia severity levels.
-

For the treatment of semantic deficits

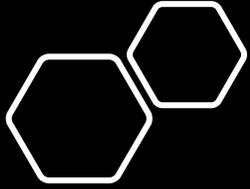
The application includes **comprehension tasks that target semantic processing:**

- Auditory word to picture matching task
- Spoken description to picture matching task
- Auditory synonym / antonym matching task
- Categorical sorting task (superordinary categorization) (pictures)
- Odd one out (pictures)

For the treatment of phonological deficits

The application includes **comprehension tasks that target phonological processing.**

- Initial syllable to picture matching task
- Rhyme to picture matching task



Auditory word to picture matching task

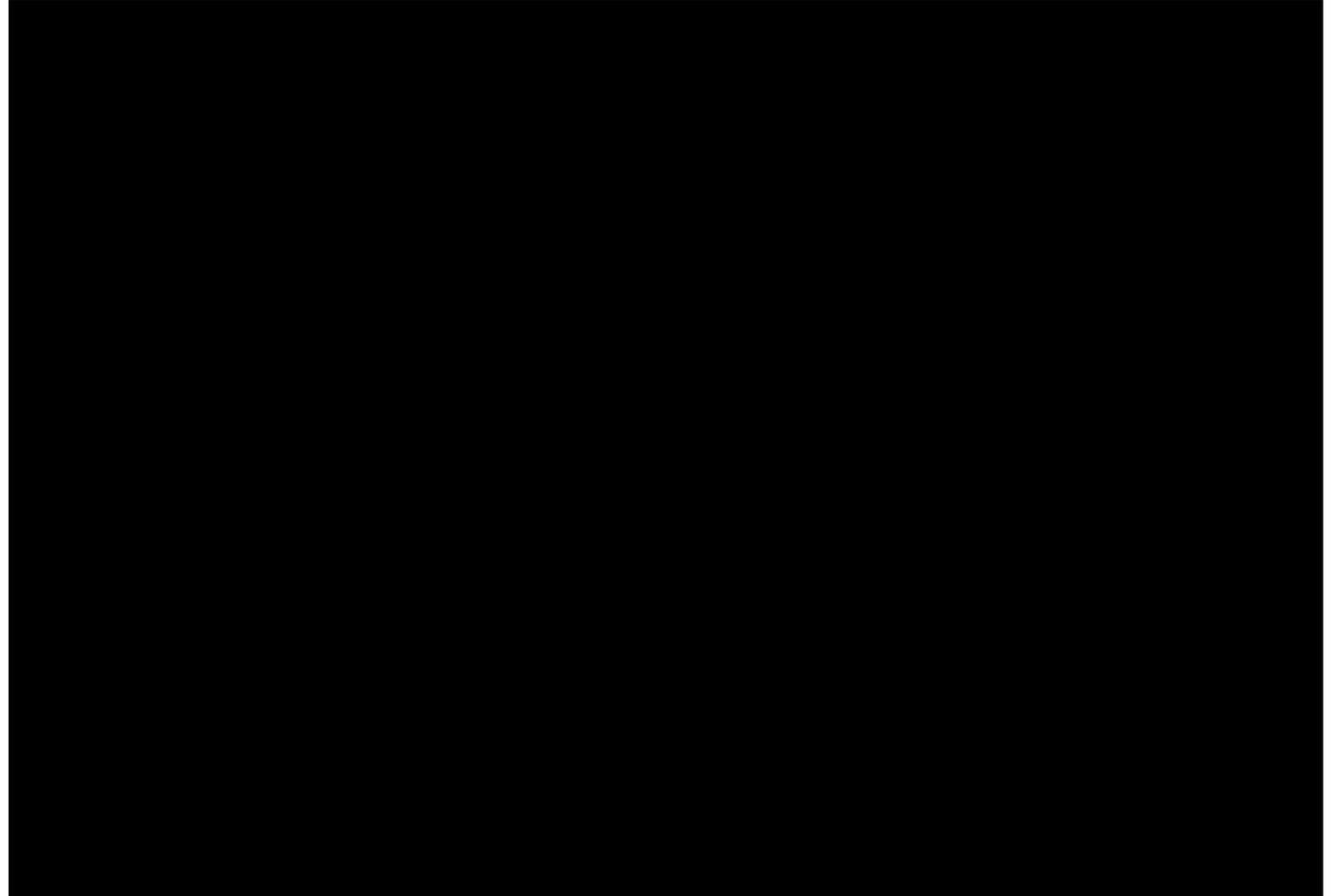
Settings:

- Complexity adaptivity (number of distractors can be adjusted)
- Adjustable picture size
- As accuracy increases difficulty level changes (delayed presentation of pictures)

Available tools:

- Repetition of instructions / target word
- Elaborative feedback for correct/incorrect responses (verbal+visual cue)
- RTs are collected

Regarding stimuli: Picture norming / Words have been matched for number of syllables/word frequencies



For the treatment of word retrieval deficits

the application provides **Feedback** after **correct** and **incorrect** responses.

- **repetition** (pronouncing of the target word for the patient to repeat)
- To aid word retrieval: **multi-modal cues**
 - **phonemic cueing** (providing the initial syllable of the target word)
 - **semantic cueing** (providing semantic information about the target word)
 - **orthographic cueing** (providing the initial grapheme of the target word)
 - and **sentence cueing** (providing a sentence which must or can be completed with the target word).

(Doesborgh et al., 2004)

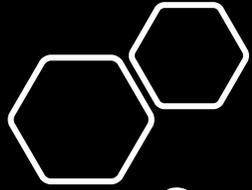
Cues can be customized allowing the clinician to choose which type of cues to assign to specific clients.

For the treatment of word retrieval deficits

specific treatment protocols are included:

- Semantic Feature Analysis (SFA: Boyle, 2004) (for noun & verb retrieval)
- Phonological Feature Analysis (Leonard et al., 2008)
- Orthographic Feature Analysis (Basso et al., 2001)

where participants are required to generate the semantic/phonological/orthographic characteristics of each noun/verb by answering specific questions per category.



Word elicitation
(noun retrieval)

Semantic Feature Analysis

Settings:

- Complexity adaptivity (number of features can be adjusted)

Available tools:

- Repetition of instructions

- Elaborative feedback for correct/incorrect responses (verbal+visual cue)

- RTs are collected

Παραγωγή Λέξης

Εναρξη



For the treatment of sentence level deficits

linguistically based treatment protocols are included:

- Treatment of Underlying Forms (TUF) (Thompson et al., 2010)
- Verb-centered Mapping therapies (Marshall, 1995)
- Activities for verb inflection/tense marking (Faroqi-Shah, 2013)

In addition to:

- Word sorting and
 - Scene description tasks that require sentence generation
-



Innovative system of technologically-assisted intervention



- Tasks are assigned remotely by a speech and language therapist and are carried out by the language impaired individual with the aid of a virtual assistant.
- The virtual assistant is used to **provide instructions and feedback** required by each therapy task via a natural and friendly interface.
- Type of feedback
 - visual cue for correctness** (e.g., red/green colour to indicate incorrect/successful trial)
 - verbal elaborative feedback** (e.g. semantic/phonological cueing)



Management of content and user data

The clinician can:

- **develop individualized treatment plans** for their clients
 - **customize/modify the difficulty level of the tasks**
 - **monitor participants' progress** both on the whole battery of tasks and on each task individually so that specific difficulties can be identified
-
- A comprehensive **learning analytics** system will provide control, support planning and monitoring for clinicians and patients.
 - The system collects and processes data regarding the users and the usage of PJan-V to fulfill two objectives:
 1. To support clients in their everyday engagement with the Plan-V system (performance tracking, visualization primitives).
 2. To support data processing and facilitate explorative analysis by clinicians (data reporting that will enable them to choose and sequence content and activities).
-



Management of content and user data

(A) Task selection / assignment to participants
Creation of individual treatment plan

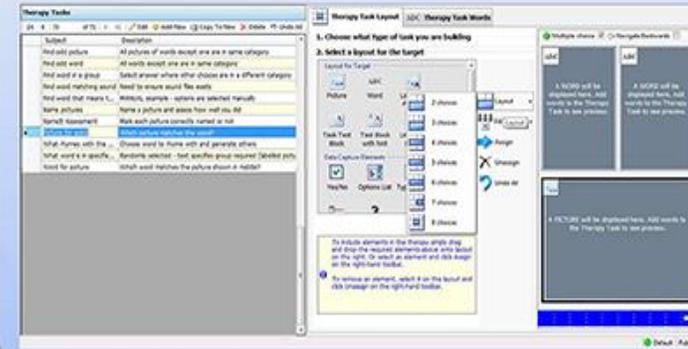
(B) Customization of difficulty level based on participants' progress

(C) Monitor participants' progress per task/
Learning analytics

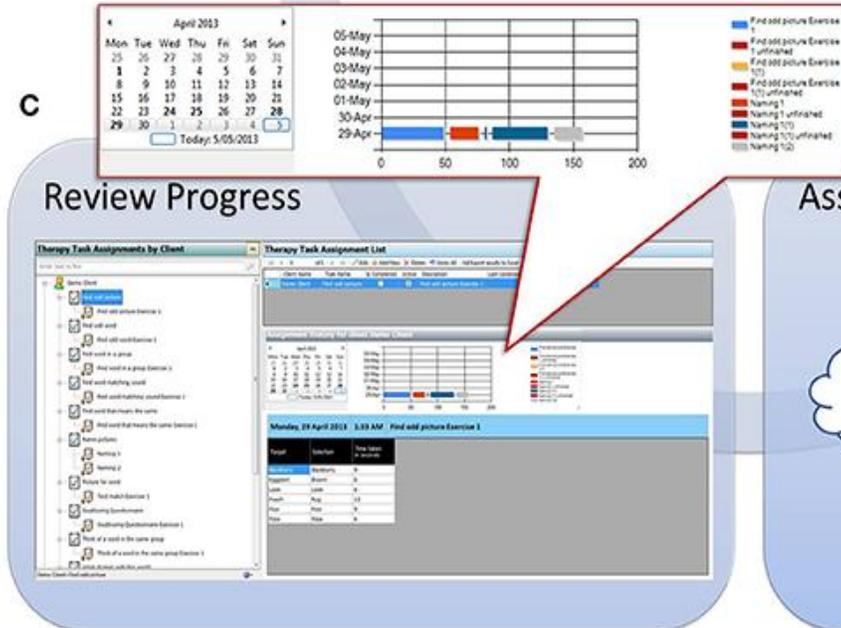
Usage Data

- Application usage
 - Users' Info
 - Start/finish time
- Performance Report
 - task actions/choices
 - voice recordings of verbal responses
 - reaction times
 - info regarding provided feedback
 - final outcome

A Modify Tasks



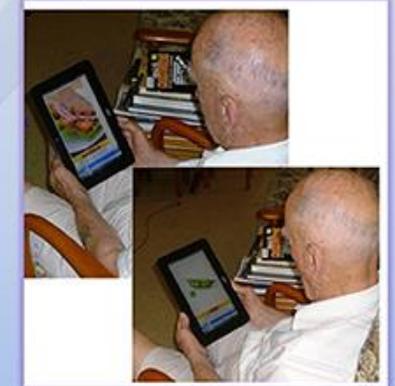
C

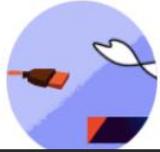


Review Progress

Assign Tasks

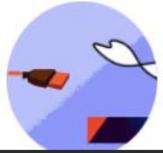
B





Automatic assessment of aphasia severity

- The platform includes an automated system for the **evaluation of the severity of the speech and language deficits**, before and after the end of the intervention, through speech samples.
 - Patients' oral discourse (narratives) will be evaluated through an automated speech recognition algorithm, using novel speech technologies and artificial intelligence/machine learning techniques.
 - This system will allow for the comprehensive assessment and monitoring of the intervention progress.
-

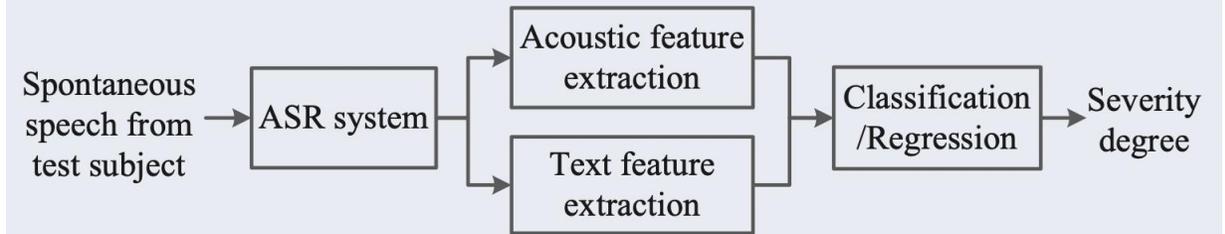


Automatic assessment of aphasia severity

– Dataset:

- **Spontaneous speech:** 80 aphasic and 30 unimpaired subjects.
 - **Narrative tasks:**
 - **Task 1:** Unaided production of a **personal narrative** (“stroke story”).
 - **Task 2:** Production of an **unknown story based on a 6-picture series** (“the party”).
 - **Task 3:** **Retelling of an unknown story**, aided by a 5-picture series (“the ring”).
 - **Task 4:** **Retelling of a familiar story** (“hare and tortoise” Aesop’s fable).
 - **Task 5:** Production of an **unknown story based on a 6-picture series** “Refused umbrella” (adapted from AphasiaBank)
 - **Task 6:** Production of an **unknown story based on complex picture** “Cat Rescue” (adapted from AphasiaBank)
 - **Task 7:** Production of a **familiar story** “Cinderella story” (adapted from AphasiaBank)
 - Both **acoustic** and **linguistic features will be extracted** from **narratives**.
- Aphasia Quotient (**AQ**: 0-100). Low → severe (WAB-R; Kertesz, 2007)

Framework for Fully Automatic Assessment



Acoustic features:

- Non-speech-to-speech duration ratio.
- Average duration of silence segments (>0.5s).
- Average duration of speech segments.
- Ratio of silence segment count to syllable count.
- Ratio of pauses to word count.

Linguistic features:

- Words per minute
- N/V ratio
- Total number of words
- Average word length
- Type/token ratio

Thank you!

- This research has been co-financed by 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).

