© Journal of Accessibility and Design for All (JACCES), Volume 16, Issue 2, 2026, ISSN: 2013-7087
DOI: https://doi.org/10.17411/jacces.v16i2.727

A rule-based multimodal computer-access system for upper-limb accessibility

Design and preliminary usability evaluation of the Turkish voice-command module

Keleş M.S., Izmir Konak Vocational School, Türkiye,
ORCID 0000-0001-8923-8805, muhammet.keles@konak.edu.tr

Alaybeyoğlu A., İzmir Kâtip Çelebi University, Türkiye
ORCID 0000-0001-9204-7356, aysegul.alaybeyoglu@ikc.edu.tr

Received: 2026-02-16 | Accepted: 2026-09-06 | Publication: 2026-11-11

Abstract: Computer access can be difficult for people who cannot use conventional mouse-and-keyboard input effectively. This study documents the design of a rule-based multimodal computer-access prototype comprising a Turkish voice-command module and a face/head-motion cursor-control module, and reports a preliminary usability evaluation limited to the voice-command component. Before development, formative consultations were conducted with two individuals with upper-limb impairments, one individual with a visual impairment, and six non-disabled stakeholders. After development, 25 non-disabled proxy participants completed a predefined scenario of 19 common computer tasks using the Turkish voice-command module. Task completion, completion time, facilitator assistance, and descriptive post-scenario feedback were recorded. All participants reached the predefined outcome for every task within an assisted protocol. Facilitator assistance was required in 28 task-participant instances across six tasks, most frequently when copying selected text, selecting a search-result link, and creating a new file. Mean task-completion times ranged from 9 to 240 seconds. The observations point to command operations and workflow transitions that may warrant further refinement, including clearer feedback, more visible error recovery, simplified multi-step interactions, and greater personalisation. However, the evaluation did not include participants with upper-limb impairments, did not test the face/head-motion module, and did not use standardised usability, workload, or cognitive-load measures. The results therefore offer preliminary proxy-user observations about the Turkish voice-command module rather than validation of the complete multimodal system or its accessibility for the intended user population.

Keywords: Assistive technology, computer accessibility, Turkish voice-command interface, multimodal interaction, upper-limb impairment, usability evaluation, Design for All

1. Introduction

Digital participation has become integral to education, employment, health services, communication, and civic life. However, access to digital systems is shaped by design decisions, and people with disabilities may be excluded when interfaces presume a narrow range of physical, sensory, or cognitive abilities. The World Health Organization estimates that approximately 1.3 billion people, corresponding to about 16% of the global population, experience significant disability, with prevalence increasing partly because of population ageing

and a rise in chronic health conditions (World Health Organization [WHO], 2023). Article 9 of the Convention on the Rights of Persons with Disabilities further recognises access to information and communication technologies as a condition for independent living and participation on an equal basis with others (United Nations, 2006). Digital accessibility should therefore be treated not as a specialist addition introduced after development, but as a fundamental requirement of inclusive design.

For people with upper-limb impairments, conventional desktop interaction may be difficult when keyboard and mouse use requires sustained fine motor control, reach, grip, or repetitive movement. Alternative computer-access methods include switch-based input, eye tracking, head-based pointing, and speech input. These methods impose different motor, sensory, cognitive, environmental, calibration, and equipment requirements; therefore, no single access method can be assumed to meet every user's abilities and preferences.

Research on head-mounted pointing devices indicates that technical feasibility alone does not determine whether an assistive interface will be adopted; perceived usefulness and ease of use are also important (Kabir et al., 2023). Research on multimodal alternative access has likewise demonstrated the potential value of offering more than one input pathway, while also identifying challenges involving modality coordination, personalisation, and task-specific interaction demands (Mandak et al., 2022). In the present study, this literature provides the design context for developing voice-command and face/head-motion modules. It should not be interpreted as evidence that both modules were evaluated with participants.

Speech recognition has long been considered an alternative input method within electronic assistive technology, particularly for users who experience difficulty operating conventional physical controls (Hawley, 2002). However, speech-based computer access depends on more than the conversion of speech into text. Users also require clear feedback about what the system has recognised, predictable relationships between spoken expressions and resulting actions, and recoverable responses when a command is misunderstood or unsupported.

These requirements also have a language-specific dimension because command wording, alternative expressions, and recognition errors may differ across languages. The prototype examined in this study addressed this design context through a Turkish command vocabulary and explicit rule-based mappings between recognised utterances and predefined computer actions (Keleş, 2024). The present evaluation therefore focuses on the usability of these Turkish voice-command interactions rather than on the performance of speech-recognition technology in general.

Accessibility standards provide useful design guidance, but their scope should be distinguished from formal conformance. WCAG 2.2 organises web-content accessibility around four principles: perceivable, operable, understandable, and robust. Because the present prototype is desktop software rather than web content, WCAG 2.2 is not treated as a direct conformance framework. The W3C's WCAG2ICT guidance explains how WCAG principles and success criteria can inform non-web documents and software, while noting that non-web accessibility may require considerations beyond WCAG alone (World Wide Web Consortium [W3C], 2023, 2025).

ISO 9241-171:2025 is more directly relevant because it specifies requirements and provides guidance for the accessibility of interactive software across a wide range of physical, sensory, and cognitive abilities (International Organization for Standardization [ISO], 2025). Applied to the present prototype, this perspective directs attention to predictable command–action relationships, perceivable system status, error prevention and recovery, user control, and

adaptability. These standards are used in this article to frame design requirements and future evaluation; the prototype was not subjected to a formal WCAG or ISO conformance assessment.

Against this background, the present article documents an early-stage rule-based computer-access prototype developed as part of a master's thesis in Türkiye. The prototype comprised a Turkish voice-command module for discrete computer operations and a face/head-motion module intended to support cursor movement. The development process was preceded by formative stakeholder consultations, whereas the subsequent participant-based evaluation was limited to the Turkish voice-command module (Keleş, 2024).

The empirical component had three objectives: first, to report task completion and facilitator-assistance patterns within a predefined 19-task voice-command scenario; second, to describe task-completion-time profiles and identify operations requiring further investigation; and third, to use the preliminary observations to outline considerations for subsequent inclusive design and target-user evaluation. The study did not compare the two input modules, evaluate multimodal conflict resolution, or assess the face/head-motion module with participants.

Within these limits, the article documents the prototype architecture, reports preliminary task-level observations from non-disabled proxy participants, and outlines design and methodological considerations for future work. It does not validate the complete multimodal system or establish that the prototype is accessible, effective, or independently usable by individuals with upper-limb impairments.

2. Methodology

2.1. Study design

The study followed a sequential development-and-evaluation design comprising three stages. First, formative consultations were conducted to explore perceived barriers to computer access and expectations regarding alternative interaction methods.

Second, a rule-based multimodal computer-access system was developed. The system comprised a Turkish voice-command module for discrete computer operations and a face/head-motion module intended to support cursor movement.

Third, a preliminary scenario-based usability evaluation was conducted with 25 non-disabled proxy participants. This evaluation focused on the Turkish voice-command module and included 19 everyday computer tasks. Task completion, completion time, facilitator assistance, and post-scenario feedback were recorded.

The face/head-motion cursor-control module was technically examined by the developer during implementation but was not included in the participant-based usability evaluation. Accordingly, the empirical findings reported in this article represent preliminary proxy-user evidence concerning the Turkish voice-command module and should not be interpreted as target-user validation of the complete multimodal system (Keleş, 2024).

2.2. Formative needs elicitation

Before system development, a formative needs-elicitation stage was conducted to explore perceived barriers to computer access and expectations regarding alternative interaction methods. The consultation group comprised two individuals with upper-limb impairments, one individual with a visual impairment, and six non-disabled stakeholders.

An initial feedback form was prepared using Google Forms to introduce the proposed system and collect preliminary views. Participants who indicated an interest in seeing the preliminary prototype were subsequently contacted. Consultations were conducted either face-to-face or online, depending on participant availability (Keleş, 2024).

The participants with upper-limb impairments described the proposed system as potentially relevant to computer-access needs. Their suggestions included allowing users to choose between alternative face-based cursor-control options, such as eye- or head-based movement, where technically feasible. Audible system feedback was also identified as a potentially useful feature. The participant with a visual impairment similarly emphasised the potential value of audible feedback. The non-disabled stakeholders provided additional observations concerning the possible usefulness and broader application of the proposed approach (Keleş, 2024).

These consultations informed the initial system requirements and design direction. However, they were conducted before completion of the system and did not constitute a usability or effectiveness evaluation. The views reported in this subsection should therefore be interpreted as formative stakeholder input rather than evidence of performance among the intended user population.

2.3. System architecture and evaluated scope

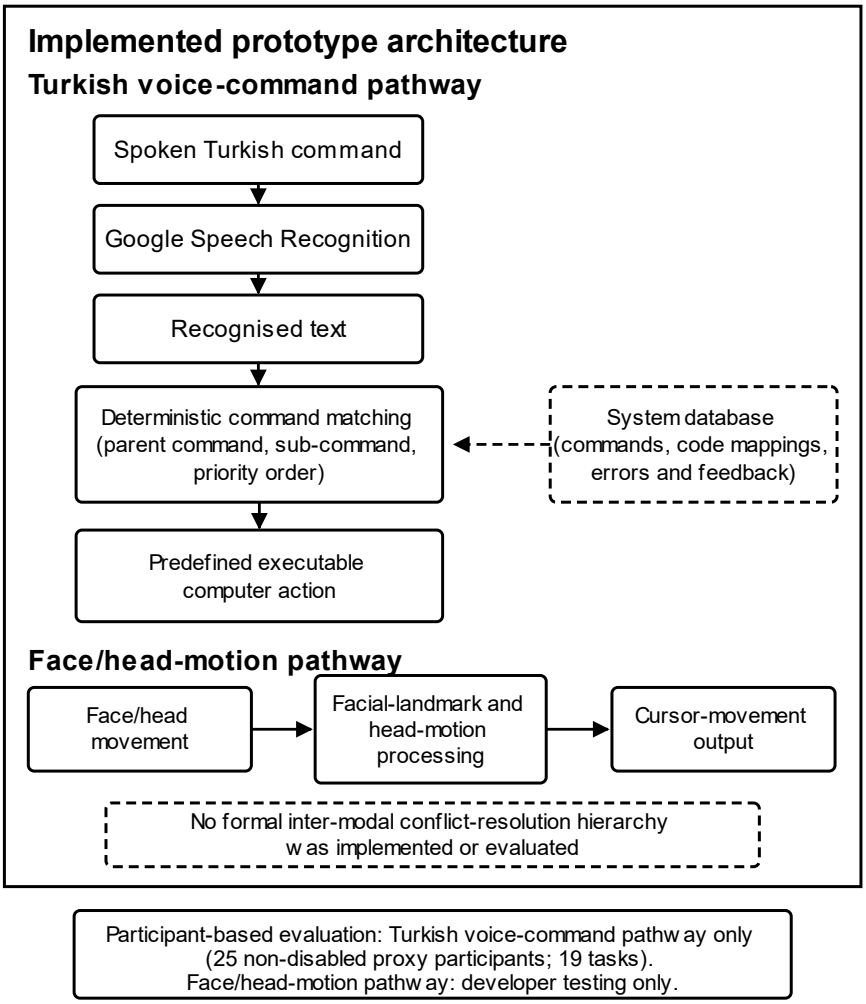
The developed system comprised two complementary interaction modules. The Turkish voice-command module was designed to execute discrete computer operations, including application and browser actions, text entry and editing, file operations, keyboard functions, and system-volume controls. Spoken Turkish utterances were converted into text using Google Speech Recognition and were subsequently processed through deterministic command-matching rules (Keleş, 2024).

The voice-command architecture was supported by a database containing user, command, parent-command, sub-command, code, error-category, feedback, and feedback-category records. Commands were organised hierarchically through parent-command and sub-command relationships. Priority fields determined the order in which command conditions were checked, while error and feedback records supported the identification of recurring interaction problems and subsequent system refinement (Keleş, 2024).

The second module used facial landmarks and head movement to support cursor navigation. Although this module formed part of the complete system architecture, it was not included in the 25-participant usability evaluation. The participant-based findings reported in this article therefore concern the Turkish voice-command module only. The face/head-motion module is described as a developed system component, and no participant-level performance or usability claims are made for it.

Figure 1 summarises the implemented processing pathways and distinguishes the component included in the participant-based evaluation from the module examined only during development.

Figure 1. Implemented system architecture and scope of the participant-based evaluation.



2.4. Rule-based voice-command processing

Following speech-to-text conversion, the recognised utterance was processed through a deterministic, hierarchical command structure. The system first identified the relevant parent command and then examined the associated sub-command conditions to determine the intended computer action. The resulting action was executed through predefined code mappings for operations such as cursor movement, clicking, application control, text entry, file management, and system-volume adjustment (Keleş, 2024).

Parent commands, sub-commands, executable code mappings, and command priorities were stored in the system database. Priority values determined the order in which command conditions were examined. This database-driven structure also enabled new commands or command variants to be incorporated by updating the corresponding command and code records rather than redesigning the entire processing architecture (Keleş, 2024).

The rule-based approach was selected in the original development because the system was intended to perform a bounded set of explicitly defined computer tasks for a specific accessibility context. The approach provided direct control over the relationship between recognised expressions and resulting actions. However, it did not learn from new utterances automatically, and extensions to the command vocabulary required manual intervention. These limitations are

considered in relation to contemporary adaptive and machine-learning-based interfaces in the Discussion section.

2.5. Preliminary usability participants

The post-development usability evaluation involved 25 non-disabled participants who acted as proxy users for a preliminary technical and interaction-level assessment. Participants were recruited through convenience sampling. They did not simulate upper-limb impairment during the sessions. The sample comprised 14 men and 11 women, with a mean age of 28 years. Participants had varied educational and professional backgrounds, including individuals with and without experience in information technology (Keleş, 2024).

None of the 25 participants was reported as having an upper-limb impairment. The purpose of this sample was to identify preliminary problems involving command use, task sequences, feedback, and facilitator-assistance requirements before conducting validation with the intended user population.

The findings from this sample should therefore not be interpreted as evidence of usability, accessibility, or effectiveness for individuals with upper-limb impairments. Participant information was reported only in aggregated or anonymised form.

Table 1. Characteristics of the preliminary usability participants.

Characteristic	Description
Number of participants	25
Participant role	Non-disabled proxy participants
Simulation of impairment	No impairment was simulated
Gender	14 men and 11 women
Mean age	28 years
Educational and professional background	Varied; included participants with and without information-technology experience
Reported upper-limb impairment	None
Evaluated module	Turkish voice-command module
Face/head-motion module used in sessions	No

The proxy-participant sample was used to identify preliminary technical and interaction-level problems before target-user validation. It was not intended to represent the functional diversity or lived experience of individuals with upper-limb impairments.

2.6. Task scenario

A predefined scenario comprising 19 sequential tasks was used to examine the preliminary usability of the Turkish voice-command module. The tasks represented common computer operations involving browser use, tab navigation, web-content interaction, text manipulation, file operations, and system-audio control. Participants completed the tasks in the same order, and each task had a predefined observable outcome (Keleş, 2024).

The scenario was designed as a procedural usability assessment rather than as a controlled pointing-device experiment. Target distance, target width, and movement amplitude were not

systematically manipulated or recorded. Consequently, the scenario does not provide the variables required for a retrospective Fitts's Law analysis.

Table 2. Tasks included in the preliminary usability scenario.

Task	Predefined task
1	Open a web browser.
2	Perform a search using the web browser.
3	Open a new browser tab.
4	Switch to the previous tab.
5	Switch to the next tab.
6	Close the newly opened tab.
7	Select the fifth link in the search results.
8	Select a paragraph on the opened web page.
9	Copy the selected paragraph.
10	Return to the desktop.
11	Create a new file.
12	Name the file using the participant's own name.
13	Open the file.
14	Paste the copied text into the file.
15	Move to a new line and enter the participant's own name.
16	Save the file.
17	Increase the system volume.
18	Decrease the system volume.
19	Mute the system audio.

The scenario progressed from browser initiation and tab management (Tasks 1–6) to web-content selection and copying (Tasks 7–9), desktop and file operations (Tasks 10–16), and system-audio control (Tasks 17–19). Several later tasks depended on outputs produced in earlier tasks; for example, the paragraph copied in Task 9 was pasted into the newly created file in Task 14.

2.7. Measures

Three task-level measures were recorded during the preliminary usability evaluation.

Task completion indicated whether the participant reached the predefined outcome for each task. Because verbal assistance could be provided when a participant reached an impasse, task completion should be interpreted within the assisted study protocol rather than as a measure of fully independent completion.

Completion time was recorded in seconds for each task. Timing began when the participant started the relevant task and ended when the predefined outcome was reached. Task-level times were used descriptively to compare durations across the scenario and to highlight operations that may warrant further investigation.

Facilitator assistance indicated whether the participant required a verbal hint to continue after reaching an impasse. Assistance consisted of verbal guidance and did not involve the facilitator completing the task on the participant's behalf. The number of participants requiring assistance was summarised separately for each task (Keleş, 2024).

After completing the scenario, participants provided comments and suggestions concerning the voice-command interaction. These comments addressed issues such as command phrasing, command coverage, feedback visibility, and difficulties encountered during multi-step operations. The feedback was used descriptively to support the interpretation of the task-level results; it was not collected using a standardised usability or cognitive-load scale.

2.8. Procedure

The usability sessions were conducted after completion of the initial software-development stage. Participants were informed about the Turkish voice-command interaction and the sequence of activities included in the scenario. They were then asked to complete the 19 tasks in the predefined order using the voice-command module.

For each task, the evaluator recorded the completion time and whether the participant required assistance to proceed. When a participant reached an impasse, brief verbal guidance was provided to enable continuation of the scenario. The evaluator did not complete the task on the participant's behalf. Assistance events were recorded because completion achieved after guidance could not be interpreted as fully independent task completion.

After completing the scenario, participants were asked to comment on the strengths and limitations of the interaction and to suggest possible improvements. These comments were collected as descriptive feedback; no standardised usability, workload, fatigue, or cognitive-load questionnaire was administered (Keleş, 2024).

The face/head-motion cursor-control module was not used during the participant sessions. It was examined only through developer testing during the implementation stage and is therefore excluded from the participant-based usability and performance findings reported in this article.

2.9. Analysis plan

The analysis was descriptive and was conducted at both task and scenario levels. At the scenario level, the number of completed tasks and the overall distribution of facilitator-assistance events were summarised. At the task level, completion time and the number of participants requiring assistance were examined to identify tasks associated with comparatively greater interaction difficulty (Keleş, 2024).

Task results were interpreted according to the operations required in the scenario, including browser initiation, tab navigation, web-content selection, text manipulation, file operations, and system-audio control. Because the participant sessions evaluated the Turkish voice-command module, the analysis focused on command use, task sequencing, interaction feedback, and multi-step workflow demands. No participant-level analysis of face/head-motion cursor performance was conducted.

Post-scenario comments were reviewed descriptively and grouped according to recurring issues explicitly raised by participants, such as command phrasing, command coverage, feedback visibility, and suggestions for simplifying multi-step operations. These comments were used to

explain or contextualise the task-level findings. No formal qualitative coding framework, inferential statistical analysis, or standardised cognitive-load analysis was applied.

2.10. Ethical considerations

All participants provided informed consent before taking part in the usability evaluation, and participation was voluntary. Participants could withdraw at any time without penalty. No personally identifying information is reported in this article, and the findings are presented in anonymised or aggregated form.

3. Results

3.1. Evaluation sample and scope

A total of 25 non-disabled proxy participants completed the preliminary scenario-based usability evaluation of the Turkish voice-command module. The participant group included 14 men and 11 women, with a mean age of 28 years. The face/head-motion module was not used during the participant sessions and is not represented in the results reported in this section. The reported outcomes comprise task completion, task completion time, facilitator assistance, and descriptive post-scenario feedback (Keleş, 2024).

3.2. Task completion and facilitator assistance

All 25 participants reached the predefined outcome for each of the 19 tasks. However, task completion occurred within an assisted evaluation protocol and therefore does not represent fully independent completion. When participants reached an impasse, brief verbal guidance was provided so that they could continue the scenario (Keleş, 2024).

Facilitator assistance was recorded in six of the 19 tasks, corresponding to 28 task-participant instances. The highest assistance frequency occurred in Task 9, copying the selected paragraph, for which 8 participants required guidance. This was followed by Task 7, selecting the fifth search-result link, and Task 11, creating a new file, with 6 participants each. Assistance was also required in Task 1 by 5 participants, Task 10 by 2 participants, and Task 8 by 1 participant. No assistance was recorded for the remaining 13 tasks (Keleş, 2024).

Table 3. Task-level facilitator-assistance frequency.

Task	Task operation	Participants requiring assistance, n	Percentage of sample
1	Open a web browser	5	20%
7	Select the fifth search-result link	6	24%
8	Select a paragraph on the web page	1	4%
9	Copy the selected paragraph	8	32%
10	Return to the desktop	2	8%
11	Create a new file	6	24%

The distribution suggests that some operations were more likely than others to produce an impasse during the preliminary voice-command evaluation. Because the sample comprised non-disabled proxy participants, these observations point to interaction points that may warrant further examination but do not establish the accessibility barriers or assistance requirements experienced by individuals with upper-limb impairments.

Percentages were calculated using the full sample of 25 participants. Assistance was recorded as a task-level occurrence indicating that verbal guidance was required for the participant to continue. The table does not indicate the severity or cause of the impasse.

3.3. Task completion time profile

Mean completion times varied across the 19 tasks, ranging from 9 seconds for Task 6 to 240 seconds for Task 7. The longest mean completion times were recorded for Task 7 (240 s), Task 1 (222 s), Task 11 (214 s), Task 9 (161 s), Task 8 (141 s), and Task 10 (139 s). By contrast, Tasks 6, 18, and 19 had mean completion times of 9, 10, and 10 seconds, respectively (Keleş, 2024).

Table 4. Mean task-completion times in the preliminary usability scenario.

Task	Task operation	Mean completion time, s
1	Open a web browser	222
2	Perform a web search	19
3	Open a new browser tab	41
4	Switch to the previous tab	13
5	Switch to the next tab	11
6	Close the newly opened tab	9
7	Select the fifth search-result link	240
8	Select a paragraph on the web page	141
9	Copy the selected paragraph	161
10	Return to the desktop	139
11	Create a new file	214
12	Name the file	24
13	Open the file	17
14	Paste the copied text into the file	96
15	Enter the participant’s name on a new line	38
16	Save the file	14
17	Increase the system volume	11
18	Decrease the system volume	10
19	Mute the system audio	10

These values describe the time required to reach each predefined outcome within the assisted proxy-user evaluation. Because the tasks differed in their content, number of steps, and dependence on previous scenario actions, the mean times should not be interpreted as

controlled comparisons of command efficiency. They also do not provide direct evidence of fatigue, cognitive load, accessibility for individuals with upper-limb impairments, or performance relative to conventional mouse-and-keyboard input.

Mean times represent complete task durations within the sequential scenario and may include time spent interpreting the task, formulating or repeating a command, waiting for system response, and receiving verbal assistance where required. They are not isolated speech-recognition latency or movement-time measurements.

4. Discussion

4.1. Interpreting task-level impasses in the preliminary voice-command evaluation

Although all 25 participants reached the predefined outcome for every task, facilitator assistance was required in 28 task-participant instances across six tasks: Tasks 1, 7, 8, 9, 10, and 11. The concentration of assistance in a limited subset of the scenario distinguishes assisted completion from independent operability. In particular, Task 9 produced the highest assistance frequency, followed by Tasks 7 and 11.

The affected tasks involved different forms of interaction, including initiating an application, selecting web content, copying text, returning to the desktop, and creating a file. The available records do not support a definitive explanation for every impasse because assistance was not classified using a formal error taxonomy. Nevertheless, the pattern points to voice-command operations and task transitions that may warrant closer examination in subsequent design iterations. Because the evaluation involved non-disabled proxy participants, these impasses should not be interpreted as direct evidence of the barriers experienced by individuals with upper-limb impairments. They instead suggest areas in which the current voice-command implementation might benefit from clearer command phrasing, improved feedback, or simplified workflow support.

4.2. Interpreting task-completion time in the preliminary evaluation

Task-completion times varied considerably across the scenario, ranging from 9 to 240 seconds. The longer mean durations recorded for Tasks 1, 7, 8, 9, 10, and 11 indicate that these operations required more time within the assisted evaluation protocol. However, the tasks differed in their number of steps, command requirements, dependence on previous actions, and likelihood of facilitator assistance. The observed time differences therefore cannot be attributed to a single interaction problem.

The task durations are useful as descriptive indicators of operations requiring further investigation, but they should not be interpreted as direct measures of efficiency, accessibility, fatigue, frustration, cognitive load, or willingness to adopt the system. No standardised workload, fatigue, satisfaction, or cognitive-load instrument was administered, and the study did not include a comparison condition using conventional mouse-and-keyboard interaction.

Cognitive-load theory distinguishes demands inherent in a complex task from avoidable demands introduced by the way information and interaction are structured (Paas & van Merriënboer, 2020). In the present scenario, command formulation, state tracking, error recovery, and

transitions across multi-step operations could plausibly contribute to cognitive demand. This interpretation remains theoretical because cognitive load was not measured directly.

Future evaluations should separate speech-recognition latency, command-formulation and repetition time, system-response time, facilitator assistance, and independent task-execution time. Studies involving the intended user population should also combine performance measures with validated self-report instruments and qualitative accounts of effort, comfort, and interaction experience.

4.3. Positioning the rule-based approach in relation to adaptive AI

The deterministic architecture used in the original development should not be interpreted as a claim that rule-based systems are technically superior to machine-learning or generative-AI approaches. The prototype was designed to execute a bounded set of computer operations through explicit parent-command, sub-command, priority, and executable-code mappings. For this early-stage implementation, the architecture allowed the command logic and action mappings to be inspected directly, did not require a task-specific training dataset, and permitted the command vocabulary to be extended through database updates (Keleş, 2024).

These properties were accompanied by important limitations. The system depended on predefined expressions, required manual intervention when new commands or alternative formulations were added, and could not learn automatically from user interaction. Recent voice-interface research includes more adaptive alternatives. AutoVCI, for example, generates voice-command mappings from smartphone operation sequences and uses accumulated interaction information to improve semantic understanding. More recently, Savant has used large language models to enable blind screen-reader users to issue flexible natural-language commands across different computer applications (Kodandaram et al., 2024; Pan et al., 2022).

One possible future direction would be a hybrid architecture in which an adaptive language model interprets paraphrases or higher-level intentions, while a deterministic execution layer restricts the resulting action to a predefined set of validated computer operations. Such an architecture could request confirmation when the interpreted intention is ambiguous and could use logged corrections to support subsequent personalisation. This hybrid proposal was not implemented or evaluated in the present study. Because no direct comparison with machine-learning or large-language-model approaches was conducted, the current findings do not establish the superiority of the rule-based architecture.

4.4. Implications for Design for All and inclusive interaction design

Design for All requires more than the provision of an alternative input channel. It requires interaction methods that accommodate diverse abilities and preferences, provide meaningful choices, and avoid placing the full burden of adaptation on the user. The Universal Design principles most relevant to the present system include equitable use, flexibility in use, simple and intuitive use, tolerance for error, and low physical effort. Ability-Based Design similarly emphasises what individual users can do and how an interactive system can adapt to those abilities rather than designing solely around diagnostic categories (Center for Universal Design, 1997; Wobbrock et al., 2011).

Within the limited scope of the preliminary voice-command evaluation, the observations suggest the following considerations for future design:

- Provide choice without requiring every user to employ every modality. A future multimodal version should allow users to select voice commands, face/head-based control, or other compatible access methods according to their abilities, preferences, context, and task. The modalities should offer alternative or complementary routes rather than requiring a fixed combination. Command vocabulary, confirmation requirements, and interaction pace should also be adjustable where technically feasible.
- Simplify and clarify multi-step workflows. The concentration of facilitator assistance in selected browser, text-manipulation, and file-operation tasks indicates that users may benefit from explicit step prompts, clear information about the current system state, and confirmation before consequential actions. These features would support simple and intuitive use without assuming that all users have the same technical experience.
- Make recognition, errors, and recovery visible. The interface should communicate what expression was recognised, what action will be performed, and why an action could not be completed. Functions such as repeating the last command, displaying available commands, cancelling an action, and undoing the previous operation could help reduce reliance on external assistance. Audible and visual feedback should be considered as complementary options rather than as interchangeable requirements for every user.
- Support personalisation where feasible. Differences in motor control, speech, preferred interaction pace, fatigue, prior computer experience, and environmental conditions may affect which interaction configuration is appropriate. Future versions could therefore include user-adjustable settings rather than assuming that one default configuration will serve all intended users.
- Involve intended users throughout design and validation. Proxy-user evaluation may identify preliminary technical and workflow problems, but it cannot establish equitable use or accessibility for individuals with upper-limb impairments. Subsequent development should involve people with diverse upper-limb access requirements in requirements elicitation, iterative co-design, usability testing, and longer-term evaluation.

These considerations translate the preliminary observations into prospective inclusive-design considerations rather than validated accessibility outcomes. The present study points to interaction points that may warrant refinement within the Turkish voice-command module; it does not demonstrate that the current prototype already achieves equitable use, low physical effort, or effective personalisation for the intended user population.

4.5. Accessibility as a socio-technical and rights-based concern

Disability is not treated in this study solely as an individual deficit to be compensated for through technology. The social model of disability directs attention towards barriers created by environments, institutions, and design decisions. Disability Studies has similarly encouraged assistive-technology researchers to examine how problems are defined, whose expertise is prioritised, and whether disabled people participate meaningfully in research and design (Mankoff et al., 2010; Oliver, 1990). In the context of computer access, exclusion may emerge from the interaction among individual motor or speech characteristics, hardware, operating-system conventions, command vocabulary, application design, recognition performance, environmental conditions, and available support.

The World Wide Web Consortium (2024) also presents accessibility as the outcome of interdependent components rather than of content alone. These components include content,

user agents, assistive technologies, users' knowledge and strategies, developers, authoring tools, and evaluation tools. WCAG primarily addresses web content, whereas the wider W3C framework also includes guidance for user agents and authoring tools. Although the present prototype was not evaluated for conformance with WCAG, UAAG, or ATAG, this wider perspective is relevant because successful computer access depends on how the speech-recognition service, operating system, applications, feedback mechanisms, access technologies, and user strategies function together.

Article 9 of the Convention on the Rights of Persons with Disabilities frames access to information and communication technologies as a condition for independent living and participation on an equal basis with others (United Nations, 2006). Accordingly, future evaluation should consider not only whether a predefined task can eventually be completed, but also whether users can interact with autonomy, choice, privacy, acceptable effort, and reduced dependence on third-party assistance. These questions require direct participation by people with diverse upper-limb access requirements and evaluation within realistic contexts of use.

Within this socio-technical and rights-based framing, the current proxy-user study should be understood as a preliminary engineering and workflow evaluation. Non-disabled proxy participants may help identify initial technical problems and unclear task transitions, but they cannot substitute for the lived expertise of intended users or establish equitable access, independent participation, or social inclusion. The scope of the present evaluation is therefore limited to documenting the prototype, reporting preliminary interaction points that may warrant refinement, and outlining considerations for subsequent co-design and target-user validation.

4.6. Limitations

The most important limitation is that the post-development evaluation involved 25 non-disabled proxy participants rather than individuals with upper-limb impairments. Although the earlier formative consultations included two individuals with upper-limb impairments and one individual with a visual impairment, those consultations occurred before completion of the system and did not constitute usability or performance testing. The findings therefore cannot establish accessibility, usability, effectiveness, autonomy, or reduced physical effort for the intended user population.

The participant-based evaluation concerned only the Turkish voice-command module. The face/head-motion cursor-control module was examined through developer testing during implementation but was not used by the 25 participants. Consequently, the study does not provide participant-level evidence concerning cursor accuracy, pointing performance, multimodal interaction, modality switching, or conflict resolution between voice and head-motion input.

Task completion occurred within an assisted protocol. All participants eventually reached the predefined outcome for every task, but verbal guidance was provided when they reached an impasse. Completion should therefore not be interpreted as fully independent performance. Assistance was recorded as a task-level occurrence but was not classified using a formal error or assistance taxonomy, which limits conclusions concerning the precise causes and severity of the observed difficulties.

The study used a small and heterogeneous convenience sample. Prior experience with voice interfaces, familiarity with the applications used in the scenario, digital proficiency, age-related

factors, and individual command strategies were not systematically measured or controlled. These limitations restrict the interpretation and generalisability of the descriptive results.

The reported completion times represented whole-task durations within a sequential scenario. They may include task interpretation, command formulation or repetition, speech-recognition and system-response time, dependence on earlier tasks, facilitator assistance, and task execution. Fitts's Law requires movement amplitude and target width, whereas Steering Law requires path geometry such as path width and length (Accot & Zhai, 1997; Soukoreff & MacKenzie, 2004). Because target width, movement amplitude, target coordinates, isolated movement time, and path geometry were not recorded, the available data do not support a valid Fitts's Law or Steering Law analysis.

No conventional mouse-and-keyboard condition, alternative voice interface, machine-learning system, or large-language-model-based interface was included for comparison. The findings consequently do not demonstrate that the developed rule-based architecture is more efficient, usable, accurate, or accessible than contemporary alternatives.

No standardised instrument was used to measure usability, satisfaction, workload, cognitive load, fatigue, comfort, perceived autonomy, or intention to use the system. Post-scenario comments were collected and interpreted descriptively rather than through a predefined qualitative coding framework. The study was also limited to short, researcher-observed sessions and did not examine longitudinal use, learning effects, personalisation, environmental variation, privacy concerns, or integration into participants' everyday computer activities.

Speech-to-text conversion relied on Google Speech Recognition. Recognition performance may therefore be influenced by the external service, microphone characteristics, network conditions, background noise, pronunciation, and changes to the service beyond the control of the developed system. The present study did not isolate or systematically compare these factors.

4.7. Future work

Future research should directly involve individuals with diverse upper-limb access requirements. Intended users should participate not only as final evaluators but also in requirements elicitation, iterative co-design, prototype refinement, and the selection of meaningful outcome measures. Participant reporting should describe relevant functional abilities, preferred access methods, prior experience, and contextual factors without reducing participants to diagnostic labels.

The Turkish voice-command and face/head-motion modules should initially be evaluated separately. For the voice-command module, future studies should record recognition accuracy, command-success rate, response latency, repetitions, reformulations, errors, recovery actions, independent task completion, and facilitator assistance. Testing should also examine variation in pronunciation, speech characteristics, background noise, microphone quality, network conditions, and familiarity with the command vocabulary.

The face/head-motion module requires a controlled participant-based pointing evaluation. Target width, movement amplitude, target location, movement time, selection errors, and calibration requirements should be recorded systematically. These data would permit appropriate Fitts's Law analyses (Soukoreff & MacKenzie, 2004) and, where path-constrained movement is examined, Steering Law analyses (Accot & Zhai, 1997). The evaluation should also consider comfort, physical effort, unintended movement, fatigue, and the need for adjustable sensitivity, dwell time, and confirmation mechanisms.

After each input method has been evaluated independently, the complete multimodal system should be assessed to explore whether the modalities function as useful alternatives or complementary forms of interaction. Future work should define how users select or switch modalities, how ambiguous input is handled, whether confirmation is required before consequential actions, and how the system responds when the modalities produce conflicting signals. These mechanisms should be evaluated rather than assumed from the existence of multiple input channels.

Multimodal interfaces may use complementary input signals to reduce recognition ambiguity, but this benefit depends on explicitly designed temporal integration, confirmation, and conflict-resolution rules (Oviatt, 2000). The present prototype did not implement or evaluate such mutual-disambiguation mechanisms.

Comparative studies should examine the rule-based system alongside conventional mouse-and-keyboard interaction, existing accessibility tools, and adaptive voice-interface approaches. A hybrid architecture could use an adaptive language model to interpret paraphrases or higher-level intentions while retaining a deterministic execution layer that restricts actions to a validated command set. Such a design would require explicit confirmation, transparent feedback, error recovery, privacy protection, and safeguards against unintended computer actions. Its effectiveness should be established through direct comparison rather than theoretical claims.

Future evaluations should combine objective performance data with validated measures of usability, workload, cognitive load, comfort, satisfaction, perceived autonomy, and intention to use. Qualitative interviews or think-aloud sessions could provide further insight into command formulation, trust, privacy concerns, preferred feedback, and dependence on external assistance. Any qualitative analysis should use a clearly described coding and reliability procedure.

Finally, longitudinal studies in realistic contexts are needed to examine learning, personalisation, sustained use, environmental variation, and integration into everyday computer activities. Evaluation should consider not only task completion and speed but also autonomy, user choice, acceptable effort, privacy, reliance on third-party assistance, and participation in education, employment, communication, and other areas of digital life.

5. Conclusion

This study documented an early-stage computer-access prototype comprising a Turkish voice-command module and a face/head-motion cursor-control module, and reported a preliminary scenario-based evaluation limited to the voice-command module. Twenty-five non-disabled proxy participants completed 19 tasks within an assisted protocol. All participants reached the predefined outcome for each task, but facilitator assistance was recorded in 28 task-participant instances across six tasks, and mean task-completion times ranged from 9 to 240 seconds. These observations point to operations that may warrant further examination, particularly web-content interaction, text manipulation, file creation, and application navigation.

Within its limited scope, the study documents an early-stage Turkish computer-access prototype and reports preliminary interaction problems observed in the voice-command evaluation. The observations suggest possible areas for future development, including clearer command and system-state feedback, accessible error recovery, personalisation, user choice, and direct involvement of intended users in co-design and evaluation. Future studies could evaluate the two modules independently before examining their combined use, recruit participants with diverse upper-limb access requirements, collect the variables required for appropriate pointing-

performance analyses, and compare the deterministic architecture with adaptive alternatives. The present results should be treated as preliminary and should not be interpreted as evidence that the system has achieved effective or equitable computer access.

6. Acknowledgements

Funding. This research received no external funding. The authors declare that they have no financial or other substantive conflicts of interest.

AI disclosure. During the revision of this manuscript, the authors used ChatGPT (OpenAI) to assist with language editing, restructuring, and consistency of selected passages in response to the reviewers' comments. The tool was not used to generate or modify the study design, data, analyses, or reported results. All AI-assisted text was critically reviewed, verified, and edited by the authors, who take full responsibility for the accuracy and integrity of the final manuscript.

7. Bibliography

Accot, J., & Zhai, S. (1997). Beyond Fitts' law: Models for trajectory-based HCI tasks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '97) (pp. 295–302). Association for Computing Machinery. https://doi.org/10.1145/258549.258760

Center for Universal Design. (1997). The principles of universal design (Version 2.0). North Carolina State University. https://design.ncsu.edu/wp-content/uploads/2022/11/principles-of-universal-design.pdf

Hawley, M. S. (2002). Speech recognition as an input to electronic assistive technology. British Journal of Occupational Therapy, 65(1), 15–20. https://doi.org/10.1177/030802260206500104

International Organization for Standardization. (2025). Ergonomics of human-system interaction—Part 171: Software accessibility (ISO Standard No. 9241-171:2025). https://www.iso.org/standard/86308.html

Kabir, M. R., Mahmud, H., & Hasan, M. K. (2023). Acceptability of a head-mounted assistive mouse controller for people with upper limb disability: An empirical study using the technology acceptance model. PLOS ONE, 18(10), e0293608. https://doi.org/10.1371/journal.pone.0293608

Keleş, M. S. (2024). El veya kollarında engeli bulunan bireylerin bilgisayar kullanabilmelerini sağlayabilecek bilgisayar sistemi (Tez No. 874956) [Yüksek lisans tezi, İzmir Kâtip Çelebi Üniversitesi]. Ulusal Tez Merkezi

Kodandaram, S. R., Uckun, U., Bi, X., Ramakrishnan, I. V., & Ashok, V. (2024). Enabling uniform computer interaction experience for blind users through large language models. In Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '24) (Article 73, pp. 1–14). Association for Computing Machinery. https://doi.org/10.1145/3663548.3675605

Mandak, K., Light, J., & Brittlebank-Douglas, S. (2022). Exploration of multimodal alternative access for individuals with severe motor impairments: Proof of concept. Assistive Technology, 34(6), 674–683. https://doi.org/10.1080/10400435.2021.1910375

Mankoff, J., Hayes, G. R., & Kasnitz, D. (2010). Disability studies as a source of critical inquiry for the field of assistive technology. In Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '10) (pp. 3–10). Association for Computing Machinery. https://doi.org/10.1145/1878803.1878807

Oliver, M. (1990). The politics of disablement. Macmillan Education. https://doi.org/10.1007/978-1-349-20895-1

Oviatt, S. (2000). Taming recognition errors with a multimodal interface. Communications of the ACM, 43(9), 45–51. https://doi.org/10.1145/348941.348979

Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. Current Directions in Psychological Science, 29(4), 394–398. https://doi.org/10.1177/0963721420922183

Pan, L., Yu, C., Li, J. H., Huang, T., Bi, X., & Shi, Y. (2022). Automatically generating and improving voice command interface from operation sequences on smartphones. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22) (Article 208, pp. 1–21). Association for Computing Machinery. https://doi.org/10.1145/3491102.3517459

Soukoreff, R. W., & MacKenzie, I. S. (2004). Towards a standard for pointing device evaluation: Perspectives on 27 years of Fitts' law research in HCI. International Journal of Human-Computer Studies, 61(6), 751–789. https://doi.org/10.1016/j.ijhcs.2004.09.001

United Nations. (2006). Convention on the Rights of Persons with Disabilities. https://www.un.org/disabilities/documents/convention/convoptprot-e.pdf

Wobbrock, J. O., Kane, S. K., Gajos, K. Z., Harada, S., & Froehlich, J. (2011). Ability-based design: Concept, principles and examples. ACM Transactions on Accessible Computing, 3(3), Article 9, 1–27. https://doi.org/10.1145/1952383.1952384

World Health Organization. (2023, March 7). Disability and health. https://www.who.int/news-room/fact-sheets/detail/disability-and-health

World Wide Web Consortium. (2023). Web Content Accessibility Guidelines (WCAG) 2.2. https://www.w3.org/TR/WCAG22

World Wide Web Consortium. (2024, March 14). Essential components of web accessibility. https://www.w3.org/WAI/fundamentals/components

World Wide Web Consortium. (2025). Guidance on applying WCAG 2 to non-web information and communications technologies (WCAG2ICT). https://www.w3.org/TR/wcag2ict

© Journal of Accessibility and Design for All (JACCES), Volume 16, Issue 2, 2026, ISSN: 2013-7087
DOI: https://doi.org/10.17411/jacces.v16i2.727

How to cite this article

Keleş M.S., Alaybeyoğlu A. (2026). A rule-based multimodal computer-access system for upper-limb accessibility. Design and preliminary usability evaluation of the Turkish voice-command module. Journal of Accessibility and Design for All, 16 (2), 48-65. https://doi.org/10.17411/jacces.v16i2.727

© Journal of Accessibility and Design for All (JACCES), ISSN 2013-7087, is published by the Universitat Politècnica de Catalunya, Barcelona Tech, with the sponsoring of ONCE Foundation for Cooperation and Social Inclusion of People with Disabilities. This issue is free of charge and is available in electronic format.

This work is licensed under an Attribution-Non-Commercial 4.0 International Creative Commons License. Readers are allowed to read, download, copy, redistribute, print, search, or link to the full texts of the articles or use them for any other lawful purpose, giving appropriate credit. It must not be used for commercial purposes. To see the complete license contents, please visit http://creativecommons.org/licenses/by-nc/4.0.

Jacces is committed to providing accessible publication to all, regardless of technology or ability. The present document grants vital accessibility since it applies to WCAG 2.2 and PDF/UA recommendations. The evaluation tool used has been Adobe Acrobat® Accessibility Checker. If you encounter problems accessing the content of this document, you can contact us at jacces@catac.upc.edu

Jacces is published under the DOAS (Diamond Open Access Standard) framework of the DIAMAS project, which certifies the quality and excellence of academic journals. It has been evaluated using the Self-assessment tool provided by FECYT, from the European Research Area (ERA). This article meets the requirements of the Diamond standard https://diamas.fecyt.es