

Cognitive Accessibility Issue Papers - Conversational Voice Systems

29 July 2024

More details about this document

[Copyright](#) © 2024 [World Wide Web Consortium](#). W3C® [liability](#), [trademark](#) and [permissive document license](#) rules apply.

Abstract

EDITOR'S NOTE

This text is an early draft. We intend to improve it before publication. We are still

This document is part of set of papers that describe accessibility issues for users with various cognitive or learning disabilities and mental health issues. See [cognitive or learning disabilities issue papers](#) for other modules.

This document is part of a set of related informative publications from the [Cognitive and Learning Disabilities Accessibility Task Force](#) (COGA TF), a joint task force of the [Accessible Platform Architectures Working Group](#) (APA WG) and the [Accessibility Guidelines Working Group](#) (AG WG) of the [Web Accessibility Initiative](#).

Status of This Document

This document is merely a W3C-internal document. It has no official standing of any kind and does not represent consensus of the W3C Membership.

Feedback on any aspect of the document is welcome. The Working Groups particularly seeks feedback on the following questions:

- Are issues for the included user groups well covered?
- Do the issue papers effectively bridge between research and priorities for new work?
- Should additional user groups should be addressed?

To comment, [file an issue in the W3C coga GitHub repository](#). You can also send an email to public-coga-comments@w3.org ([comment archive](#)). Comments are requested by **16 July 2024**. In-progress updates to the document may be viewed in the [publicly visible editors' draft](#).

table of contents

[Abstract](#)

[Status of This Document](#)

[1.](#)

[Introduction](#)

[2.](#)

[Challenges for People with Cognitive Disabilities](#)

[2.1](#)

[Memory: Recalling and Responding to Prompts](#)

[2.2](#)

[Executive function: Deciding When to Respond](#)

[2.3](#)

[Reasoning: Making the Correct Selection](#)

[2.4](#)

[Knowledge: Interpreting Responses](#)

[2.5](#)

[Processing Speed](#)

[2.6](#)

[Attention: Making Correct Selections](#)

[2.7](#)

[Language and Auditory Perception: Correctly Interpreting Spoken Information](#)

[2.8](#)

[Speech and Language Production: Responding to Voice and Speech](#)

[Recognition Systems](#)

[2.9](#)

[Mental Health: Interacting with Conversational Voice Systems](#)

[3.](#)

[Proposed Solutions Based on Current Research](#)

[3.1](#)

[Ensure Human Backup is Available and Reachable](#)

[3.2](#)

[Allow Users To Change Settings](#)

[3.3](#)

Ensure Voice and Speech Recognition for All Users

3.4

Follow Best Practices in General Voice Interaction Design

3.5

Implement Best Practices in Cognitively Accessible VUI Design

3.6

Follow Appropriate Legislative Requirements

3.7

Integrate New Technology-Based Solutions

3.8

Status of these solutions

4.

References

1. Introduction

In this issue paper, we address issues for users with cognitive disabilities using conversational voice systems. Conversational voice systems, including voice menu systems and voice user interfaces, are systems with bi-directional communication in which a user:

- Supplies a spoken prompt and receives a spoken response from the system
- Receives a spoken prompt from the system and supplies either a spoken or manual reply (for example, a button push)

Conversational Voice Systems may include:

- **Voice Menu Systems** used in telephone self-service applications that require the user to respond by pressing keys on a telephone keypad or verbally responding (or both).
- **Voice User Interfaces** (VUIs) or conversational interfaces such as Siri, Cortana, Amazon Alexa or Google Assistant. These systems use voice recognition and artificial intelligence to allow users to interact with them via spoken commands.

It is worth noting that many crucial systems integrate voice systems, including emergency notifications, healthcare scheduling, prescription refilling, and more. With this in mind, full accessibility needs to be supported.

An example use case of a voice system used in telephone self-service may be as follows:

- Upon dialing the number, the user may be asked "Press 1 for hours of operation, Press 2 to refill your prescription, press 3 to speak to the pharmacist." The system then waits for a response.

An example of a use case for a voice user interface may be as follows:

- The user says the activating phrase to begin an interaction with the VUI such as “Hello, [Name of System].” The user then asks, “What is the weather today?” The VUI provides a verbal response.

Voice systems are often implemented with the [W3C VoiceXML 2.0](#) standard and supporting standards from the [Voice Browser Working Group](#).

VoiceXML 2.0 has an appendix regarding accessibility that briefly discusses use by users with hearing and speech disabilities as well as WCAG and WAI specifications. Cognitive disability is mentioned once in regards to allowing users to control the length of time before time out. See [VoiceXML 2.0#accessibility](#) for more.

Do we need to include something about training AI for VUIs or is this covered elsewhere?

2. Challenges for People with Cognitive Disabilities

Voice technology can create a number of challenges for people with cognitive disabilities, due to its heavy demands on memory and on the ability to understand and produce speech in real time. In particular, voice systems and VUIs can be inaccessible to people with cognitive disabilities that affect:

- Memory
- Executive Function
- Reasoning
- Knowledge
- Processing Speed
- Attention
- Language and Auditory Perception
- Speech and Language Production
- Mental Health

2.1 Memory: Recalling and Responding to Prompts

General: The user needs to recall information needed to successfully interact with the system, such as activating phrases, as well as information presented by the system during the interaction

Voice Menu Systems: Systems that rely on menus that present several choices at once may pose challenges for users with disabilities related to working memory. Such systems require users to hold multiple pieces of information, such as the number associated with an option, while processing the terms that follow. This is true of systems that require either a voice response or a key press.

Extensive lists that require strong working memory may result in users with cognitive disabilities making the wrong selection. Holding a list of 5-7 items (which is often considered the standard amount to use in working memory) may be difficult for someone with cognitive impairment/working memory impairment.

Voice User Interfaces: For VUIs, users may be required to remember key phrases (such as activating phrases like “Hey Google”) in order to operate successfully. Users who may not be able to recall such phrases due to long-term memory impairments may not be able to operate the system.

2.2 Executive function: Deciding When to Respond

General: When interacting with a system, if a system response is too slow, a user with disabilities related to executive functioning may not know that their input has registered, and may press the key or speak again.

Voice Menu Systems: The user needs to be able to decide when to act on a menu choice. If the user does not know how many options will be presented or if the system presents them too slowly, they may make an incorrect choice based on partial information.

2.3 Reasoning: Making the Correct Selection

Voice Menu Systems: The user may need to compare similar options such as "billing", "accounts," "sales" and decide which is the service that is best suited to solve the issue at hand. Without additional context or prior knowledge, the user is likely to select the wrong menu option.

2.4 Knowledge: Interpreting Responses

General: If responses produced by a system are not provided in clear and accessible language, the user may have difficulty interpreting them, meaning that even if the request is appropriately processed by the system, the response may be inaccessible to the user.

2.5 Processing Speed

General: Systems that time out may not give users with cognitive disabilities affecting processing speed sufficient time to interpret information and formulate a response. Advertisements and additional, unrequested information also increase the amount of processing required.

2.6 Attention: Making Correct Selections

Voice Menu Systems: The user needs to focus on the different options and select the correct one. Having long or multi-level spoken menus without written counterparts, inserting

advertisements, or otherwise including additional, unrequested information may make it harder to retain attention for users with attention-related disabilities.

2.7 Language and Auditory Perception: Correctly Interpreting Spoken Information

General: The user needs to interpret the correct terms and match them to their needs within a certain time limit. This involves speech perception and language understanding: sounds of language are heard, interpreted and understood, within a given time.

Users with disabilities related to language and auditory perception may make mistakes in interpretation due to auditory-only input.

2.8 Speech and Language Production: Responding to Voice and Speech Recognition Systems

General: The user needs to be able to formulate a spoken response to the prompt before the system "times out" and generates another prompt.

Users who utilize assistive and augmentative communication devices (AAC) or speech-to-speech technologies may require additional time to respond before the system times out.

In directed dialog systems the user only needs to be able to speak a word or short phrase. However, increasingly, natural language systems allow the user to describe their issue in their own words. While this feature is an advantage for some users because it does not require them to remember menu options, it can be problematic for users with disabilities that impact their ability to produce spontaneous speech, such as people with aphasia, or autistic people for whom stress may impact spoken communication.

Speech recognition systems may not recognize and be responsive to inputs from users with disabilities that impact the intelligibility of their speech, such as people with Down syndrome.

Users may not be able to interact with a system that requires a verbal input but does not recognize their speech.

2.9 Mental Health: Interacting with Conversational Voice Systems

General: Mental health, such as anxiety levels, may also impact a user's ability to interact with a conversational voice system. High demands to cognitive load, negative experiences with conversational systems, and interruptions can exacerbate anxiety or frustration, and decrease a user's ability to interact with a system.

3. Proposed Solutions Based on Current Research

3.1 Ensure Human Backup is Available and Reachable

Voice Menu Systems: For users who are unable to use the automated system, it must be possible to reach a human through an easy transfer process (e.g., not by being directed to call another phone number).

For telephone self-service systems, there should be a reserved digit for requesting a human operator. The most common digit used for this purpose is "0"; however, if another digit is already in widespread use in a particular country, then that digit should always be available to get to a human agent. Systems especially should not attempt to make it difficult for users to reach an agent through the use of complex digit combinations. This could be enforced by requiring implementations to not allow the reserved digit to mean anything other than going to an operator. Other digits similarly could be used for specific reserved functions, keeping in mind that too many reserved digits will be confusing and difficult to learn. Remembering more than one or two reserved digits may be problematic for some users, but repeated verbal recitals of the reserved digits will also be distracting.

3.2 Allow Users To Change Settings

User-specific settings can be used to customize the voice user interface, keeping in mind that the available mechanisms for invoking user-specific settings are minimal in a voice interface (speech or DTMF tones). If it is difficult to set user preferences, they won't be used. Setting preferences by natural language is the most natural ("slow down!") but is not currently very common. Examples of customization include:

General:

1. Extra time should be a user setting for both the speed of speech and ability for the user to define if they need a slower speech or more input time etc.
2. Timed text should be adjustable (as with all accessible media).
3. The user should be able to extend or disable time out as a system default on their device
4. Advertisement and other information should not be read as it can confuse the user, increase cognitive load, and can make it harder to retain attention.
5. Terms used should be as simple as possible. Uncommon and unfamiliar terms should be explained.

Voice Menu Systems:

1. Error recovery should be simple, and take you to a human operator. Error response should not end the interaction or send them to a more complex menu. For telephone-based systems, they should use a reserved digit.

2. Examples and advice should be given on how to build a prompt that reduces the cognitive load
 1. Example 1: Reducing cognitive load: The prompt "*press 1 for the secretary,*" requires the user to remember the digit 1 while interpreting the term secretary. It is less good than the prompt "*for the secretary (pause): press 1*" or "*for the secretary (pause) or for more help (pause): press 1*"
 2. Example 2: Setting a default for a human operator as the number 0

Voice User Interfaces

1. Consider customizable activation phrases for VUIs.
2. Consider customizable voices, including regional and standard accents to create familiarity for the user.

3.3 Ensure Voice and Speech Recognition for All Users

1. For speech recognition based systems, an existing ETSI standard for voice commands for many European languages exists and should be used where possible [ETSI 202 076], keeping in mind that expecting people to learn more than a few commands places a burden on the user.
2. Natural language understanding systems allow users to state their requests in their own words, and can be useful for users who have difficulty remembering menu options, or who have difficulty mapping the offered menu options to their goals. However, natural language interfaces can be difficult to use for users who have difficulty producing speech or language. Directed dialog (menu-based) fallback or transfer to an agent should be provided.
3. Test speech recognition for users with non-normative speech patterns or speech impairments related to disability. Include contingencies for pauses, use of incorrect terms, and mistakes. Ensure alternative access for users whose speech is not recognized.

3.4 Follow Best Practices in General Voice Interaction Design

Standard best practices in voice user interface apply to users with cognitive disabilities, and should be followed, such as those provided by the Association for Voice Interaction Design Wiki [[AVIXD](#)] or [ETSI ETR 096](#). Some examples of generally accepted best practices in voice user interface design:

1. Pauses between phrases to allow processing time.
2. Long time outs.
3. Simple error recovery, taking user to human operator if error persists.
4. No advertisements or extraneous information presented.
5. Jargon-free and simple terms.

6. Ability to review the conversation as it is happening, such as through a visual log.

See the AVIXD wiki cited above for additional recommendation and detail.

3.5 Implement Best Practices in Cognitively Accessible VUI Design

Some specific best practices for users with cognitive disabilities include:

1. **Manageable Lists:** Repeat and/or chunk lists into manageable groups. Each list item should be a single idea.
2. **Reminders:** Provide cues as to what the user needs to do and when.
3. **Guided Instructions:** Provide guided instructions to walk user through the steps in the process
4. **Task History:** Include methods that inform the user of what you have done recently to help prevent from repeating tasks and to allow the user to know where they are in the process.
5. **Timers:** Provide timers to track the user's time on a step and invoke cues or reminders if thresholds of time are crossed
6. **Task Overview:** Provide a way to attain a list of all the steps required and information needed to complete the task or process.

3.6 Follow Appropriate Legislative Requirements

For example, the U.S. Telecommunications Act Section 255 Accessibility Guidelines [Section255] paragraph 1193.41 Input, control, and mechanical functions, clauses (g), (h) and (i) apply to cognitive disabilities and require that equipment should be operable without time-dependent controls, the ability to speak, and should be operable by persons with limited cognitive skills.

3.7 Integrate New Technology-Based Solutions

Recent technological developments may be helpful for users with cognitive disabilities.

1. **Visual VUI:** When a call comes in on a smartphone, the system can ask the user if they want to switch over to a visual interface which mirrors the voice interface. This allows a user to see the prompts instead of having to remember them. If and where possible, supplementing text with images, symbols, or icons can increase understandability.
2. **Adaptive voice interface:** This is a technology that is sensitive to the user's behavior and changes the voice interface dynamically. For example, it can slow down or speed up to match the user's speech rate.
3. **Tapered prompts:** Best practices in voice user interface design include providing several different prompts for each point in the interaction. The different prompts are used based

on the user's behavior. For example, if the user takes a long time to respond to a prompt, a simpler or more explanatory version of the prompt may be used instead of the default.

3.8 Status of these solutions

Note: The above proposed solutions have been tested for users in the general population and have been shown to improve the usability of voice systems. Some of these solutions have been tested with users with cognitive disabilities, primarily in academic research contexts.

Currently VoiceXML does not directly enforce accessibility for people with cognitive disabilities. However, a considerable literature on voice user interface design exists and is in many cases very applicable to cognitive accessibility for voice systems. Developers must become aware of these resources and of the need to design systems with these users in mind.

4. References

1. [AVIXD] [The Association for Voice Interaction Design Wiki](#)
2. [ETSI 202 076] [ETSI ES 202 076 V2.1.1 \(2009-06\)](#)
3. [ETSI ETR 096] [ETSI ETR 096 Human factors guidelines for the design of minimum phone based user interface to computer services](#)
4. [Section255] [Telecommunications Act Section 255 Accessibility Guidelines](#)
5. [Adaptive] [Adaptive Voice White Paper](#)
6. Clark, L. et al. (2019). The State of Speech in HCI: Trends, Themes, and Challenges. *Interacting with Computers*, 31 (4). DOI: 10.1093/iwc/iwz016
7. Ferland, L., et al. (2018). Assistive AI for Coping with Memory Loss. *Association for the Advancement of Artificial Intelligence*.
<https://www-users.cse.umn.edu/~gini/publications/papers/ferland-aaai18w.pdf>
8. Dahl, D. (2015). Voice Should Be User-Friendly--to All Users. *Speech Technology Magazine*, 20(4), 8.
9. Braun, M., Wolfel, M., Renner, G., & Menschik, C. (2020). Accessibility of Different Natural User Interfaces for People with Intellectual Disabilities. *2020 International Conference on Cyberworlds (CW), Cyberworlds (CW), 2020 International Conference on, CW*, 211–218. <https://doi.org/10.1109/CW49994.2020.00041>.
10. A. Pradhan, K. Mehta, and L. Findlater, "“Accessibility Came by Accident”: Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. Montreal QC, Canada: ACM Press, 2018, pp. 1–13.
11. Wolters MK, Kelly F, Kilgour J. Designing a spoken dialogue interface to an intelligent cognitive assistant for people with dementia. *Health Informatics Journal*. 2016;22(4):854-866. doi: [10.1177/1460458215593329](https://doi.org/10.1177/1460458215593329).

Search terms in World Cat:

1. Voice User Interface + cognitive disabilit* OR intellectual disabilit* OR learning disabilit*
2. Conversational voice system + cognitive disabilit* OR intellectual disabilit* OR learning disabilit*
3. Conversational voice system + intellectual disabilit* OR learning disabilit*
4. Voice response system + cognitive disabilit* OR intellectual disabilit* OR learning disabilit*
5. Conversational assistant + cognitive disabilit* OR intellectual disabilit* OR learning disabilit*

