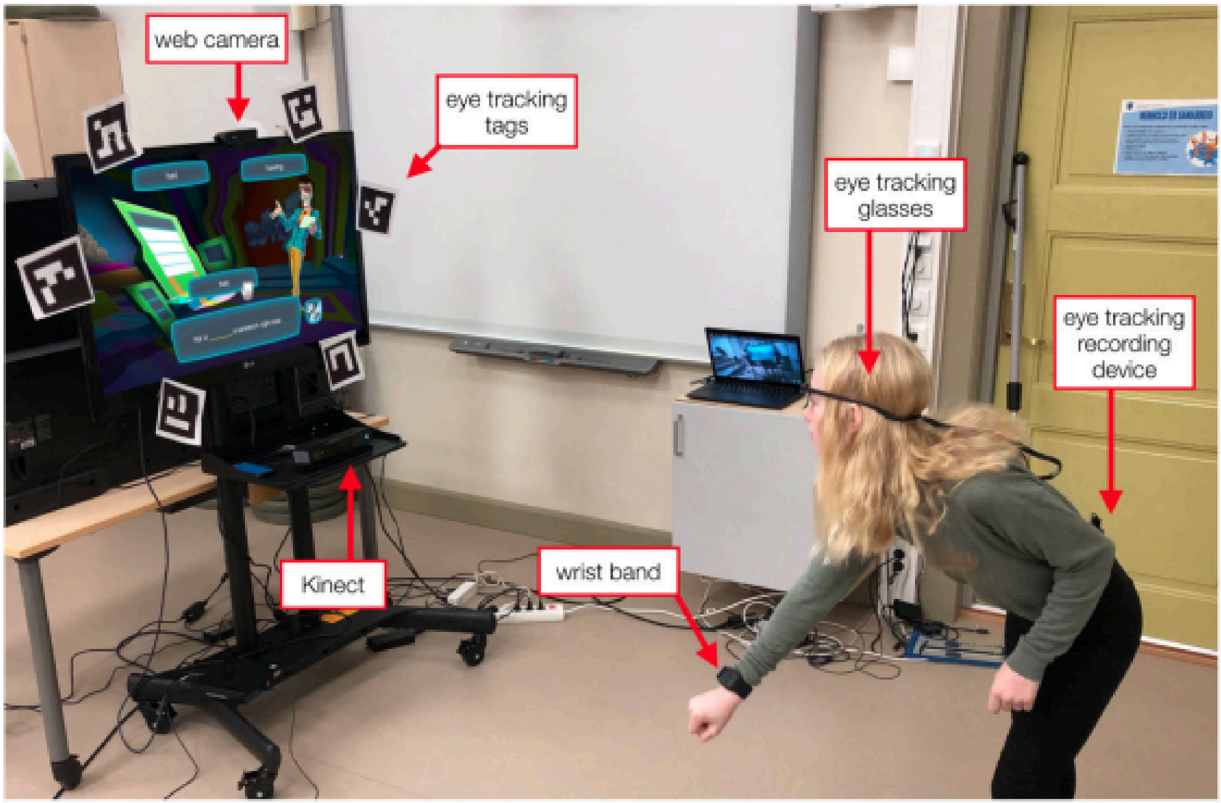
Technique #1 Name - Eye Tracking (Cami Lacy)

This particular technique explores how eye tracking and additional sensory-based analytics (SBA) may establish themselves as useful tools in educational settings both for research and technology implementations. This motivation further demonstrates how students involved have the potential to benefit from these studies while the learning systems are simultaneously improved upon according to the multimodal data that is collected by the research teams. Ultimately, this is needed because it provides a more comprehensive investigation into the benefits of multimodal data collection to further develop the learning systems assisting new generations of young students to learn better. As a cohesive unit, this technique explores how SBA largely contributes to this idea of a four-phase cycle: interaction, sensing, sense-making, and enhancing. This sequence is utilized to follow those steps beginning with the users engaging with the learning system and performing tasks appropriately. As those tasks are being completed, the sensory devices like “cameras, wearable sensors, and eye-tracking glasses” each gather information about the experimental sessions. Moreover, the sense-making phase is designated for taking the information collected and to “explain, diagnose, and guide actions, and to enrich existing practices and routines.” This is really the analysis stage before continuing to implement the new ideas for improvement during the enhancing step. This final process builds upon the previous research knowledge to actually improve the design systems to be more usable and effective for the students using them.

While our project is concentrating on the educational field with regards to artificial intelligence and implementation via headwear, we found that it was helpful to consider how this study incorporated data collection of the entire body. The eye-tracking was followed through special Tobii brand “glasses at 50-Hz sampling rate and one-point calibration, to capture students’ eye data.” Concurrently, Empatica E4 wristbands were used to measure the user’s “heart rate variability (HRV) (1 Hz), electrodermal activity (EDA) (64 Hz), skin temperature (4 Hz), and Blood Volume Pulse (4 Hz).” Furthermore, Microsoft’s Kinect Skeleton “was recorded at a sampling rate of 1 Hz, and consisted of the 3-D position for 20 joints.” While wearing these sensory devices, the user was tasked with completing a motion-based educational literacy game called Suffizz, and they had to solve various problems through several different movements. Some important design features of this technique included non-intrusive features and a simplified experience for the users. Because the target audience is exclusively elementary students for this study, it was essential to utilize devices that would not hinder the learning process in any way. This means they are easy to wear, easy to use, and easy to integrate into the learning system. Primarily, these are the key insights that make the technique “work” for the target audience involved with this system. These insights take care to focus on human interactions and usability so that the learning systems may continue advancing with each phase of the cycle.

Figure 1: “Experimental setup of a student playing Suffizz, while SBA is collected.”



This technique was evaluated through a sample of 45 students conducting trials playing the Suffizz literacy learning game, and the average age of these users was about 10.6 years old ranging from 8 years old to 12 years old. The sessions were voluntary and thoroughly detailed for each participant which required legal guardian consent. Moreover, “thirty of them participated at the science center, and 15 at the elementary school(located in Trondheim, Norway), each student played circa 10 minutes and received a gift card for participating.” In-game problems included fill-in-the-blank questions involving multiple choice options which required the student “to perform gestures to interact with on-screen content, relocate, and reposition their body within the physical game space.” As for the metrics, “all of the measures were calculated for a fixed time window of 30s and represented via a time series.” This was further examined by the research team applying “an analysis of variances (ANOVA) that investigates the relation between high- and low-performance students with respect to their stress, arousal, amount of movement, fatigue, and cognitive load (performance is based on the correctness of answers within the game).”

Using these findings, they are compared throughout the four-phase cycle as discussed earlier to encourage growth and development for the learning systems being used by educators. Additional factors that the research team took into consideration for balancing their data were utilizing methods like the Shapiro-Wilk test, Beusch-Pagan test, and Welch correction. After completing the investigations, the main findings consisted of discovering how this real-time information gathering of cognitive and affective processes helped better understand the energy and engagement levels of students so that the learning systems and educators may adjust their lesson plans appropriately. One example it gave was assisting stressed students by offering more hints or further explanations after noticing that their current state was agitated. The ultimate outcome of this research technique believes that “SBA holds the potential to significantly augment advanced learning technologies by providing novel design capabilities, such as real time, intelligent, adaptive, individualized scaffolding, and feedback to address learners' states and needs."

Figure 2: SBA Measurements

Dep. Var.	Data source	Definition
Movement (meters)	Kinect	The total distance travelled by each joint in the skeleton data, averaged over the whole body.
Fatigue (meter/ second^3)	Kinect	Fatigue is proportional to energy spent. For moving objects, it can be shown that thetrajectory with the lowest jerk (rate of change of acceleration) is least energy consuming. Hence, greater jerk leads to greater fatigue.
Arousal micro Simens	Wristband	Arousal is computed by the increasing slope of the EDA.[16] The more positive the slope of the EDA in a given time window, the higher the arousal.
Stress (celsius)	Wristband	Stress is computed as temperature's decreasing slope.[17] The more negative the slope of the temperature is in a given time window, the higher the stress.
Cognitiveload	Eye-tracker	The cognitive load is calculated using the variations in the pupil dilation for each student.[11]

While SBA provides an extensive range of data collection for studying new ways to improve learning systems, there are still various opportunities to improve or consider during the utilization of this method. For instance, the capability of the model may be restricted as it is important for researchers to note sampling frequency influences and varied numbers between accuracy and errors. This may further be discussed with regards to group-level capabilities as well "to support team-based learning and group-level feedback and awareness." It was also interesting to explore how these interdependencies between the data collected need to be reviewed between varied components for the purpose of securing a more appropriate assessment. Some final considerations for this technique contend with providing more real-time feedback rather than short trials in addition to the "ethics, privacy, and trust issues" regarding this type of data collection. It is a great deal of responsibility for teams to uphold transparency, actively avoid bias in their results, and reliably protect user data. Although these are apparent issues concerned with this technique, there is still substantial potential for growth in these educational systems given that researchers are provided access to this data and are able to advance the contemporary learning systems to best assist future generations to come.

M. N. Giannakos, S. Lee-Cultura and K. Sharma, "Sensing-Based Analytics in Education: The Rise of Multimodal Data Enabled Learning Systems," in IT Professional, vol. 23, no. 6, pp. 31-38, 1 Nov.-Dec. 2021, doi: 10.1109/MITP.2021.3089659. keywords: {Learning systems;Education;Human computer interaction;Adaptation models;Predictive models;Reflection;Real-time systems}

Technique #2 Name - Assistive Technology for Sunglasses (Grace Setiawan)

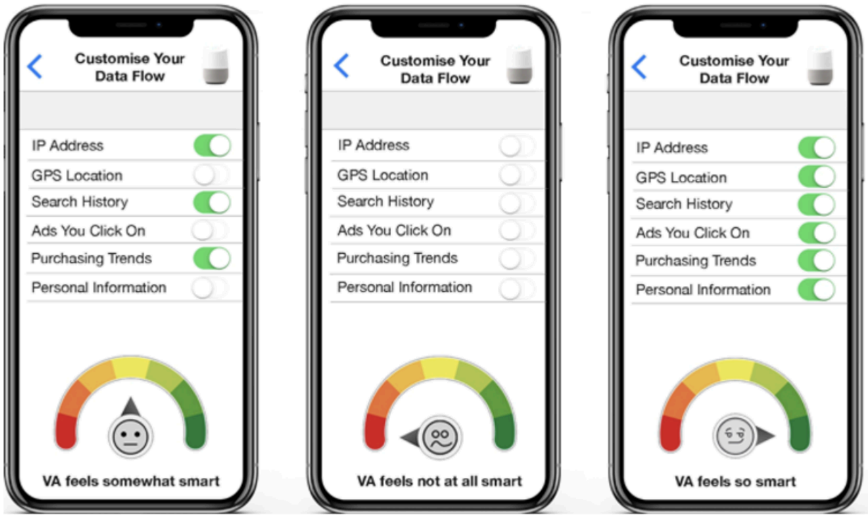
The motivation behind this technique is to enhance the functionality and accessibility of the sunglasses by integrating assistive technologies such as virtual assistants. By providing users with features like voice commands, information retrieval, and task assistance, sunglasses can become more versatile tools for various applications, including navigation, education, healthcare, and we still focus on education. The target audience for this technique includes individuals who can benefit from augmented reality glasses and assistive technologies, such as people with visual impairments, mobility limitations, or cognitive disabilities, as well as professionals in industries like healthcare and manufacturing. Also, this is a benefit for people who prefer hands-free interaction or have limited visual capabilities. It helps the users who seek assistance in various tasks without relying on visual feedback.

How it works is that the integration of the virtual assistant into the sunglasses allows users to initiate conversations and issue voice commands. Users can activate the assistant by using a wake word or a designated gesture, after which they can interact with it solely through voice input. The design features of this technique include accessibility features, such as voice commands and audio feedback, customizable user interfaces, smooth integration with existing technologies and accessibility standards, and possible support for third-party applications and services.

For the evaluation, we can do a user study focusing on different aspects of usability and user satisfaction: first, we will measure how accurate the system is to understand user commands and it involves testing whether the system correctly transcribes spoken input and executes what the user wants. Second, we want to evaluate how quickly the virtual assistant responds to user commands. We will measure the time it takes to give a command to get the response. Third, it’s important to have feedback from the users about the overall satisfaction and this also will help us understand user perceptions, preferences, and areas for improvement. The evaluation compared the performance of the voice interaction system against the “traditional” input methods such as buttons.

To measure the accuracy and responsiveness, we use other metrics such as satisfaction surveys, task completion rates, and error analysis. User feedback provides an important insight into the usability and task completion rates will provide the system’s ability to execute the command without errors. Overall, assistive technology for the sunglasses will give benefits in terms of usability and accessibility. The hands-free interaction will give users more convenience since it allows for seamless access to information and assistance. However, we understand that we will have some issues such as privacy and noise problems that need to be addressed for users to have a smooth experience.

Figure 1: The data flow customization prototype. This image shows that you can customize the prototype if you want to include your GPS location or IP address, etc.



Michael Shorter, Bettina Minder, Jon Rogers, Matthias Baldauf, Aurelio Todisco, Sabine Junginger, Aysun Aytaç, and Patricia Wolf. 2022. Materialising the Immaterial: Provotyping to Explore Voice Assistant Complexities. In Proceedings of the 2022 ACM Designing Interactive Systems Conference (DIS '22). Association for Computing Machinery, New York, NY, USA, 1512–1524. https://doi-org.offcampus.lib.washington.edu/10.1145/3532106.3533519

Technique #3 Name - Voice Input Tutoring System for Older Adults using Input Stumble Detection (Tushar Thonupunoori)

The increasing use of smartphones among older adults highlights the need for effective support in learning smartphone functionalities. Text input, particularly through voice, presents challenges for beginner older users. There are many older adults who have a desire to use smartphones but struggle with self-instruction, especially for text input. While voice input offers a solution, it presents difficulties such as input stumbles. This tutoring system aims to provide personalized guidance to improve the efficiency of voice input for older adults.

This paper provides a tutoring system tailored for voice input, addressing input stumbles encountered by older adults. In this study, they developed a tutoring system for voice input that detects input stumbles using a statistical approach, and then provides instructions that help users resolve input stumbles independently. The research follows a structured approach, starting with a user study to identify challenges faced by older adults with voice input. Based on these findings, the authors design a tutoring system tailored to address these challenges, employing machine learning algorithms for input stumble detection and clear instructional delivery. Finally, the system's effectiveness is evaluated through experiments comparing performance metrics between users with and without the tutoring system.

In terms of the target audience, they are older adults who are new to smartphones, particularly those unfamiliar with text input methods. The system is designed to assist them in mastering voice input, enabling them to make full use of smartphone functionalities. The technique remains relevant in modern systems as older adults continue to face challenges with smartphones. While newer voice recognition technologies may have improved, the need for tailored support for older adults still exists.

In smartphones, there is automatic speech recognition (ASR), which is an option for text input and this is where older adults struggle since they are not accustomed to the usage of this feature in smartphones. However, this tutoring system detects input stumbles during voice input using machine learning algorithms and a statistical approach. It provides instructions through text and key highlighting based on the observed effective instructions from human tutors. The system aims to improve user efficiency and confidence in voice input.

Some important design features are machine learning-based detection of input stumbles, personalized instructions through text and key highlighting, and a user-friendly interface for clear guidance and feedback. The personalized and adaptive nature of the tutoring system, providing tailored guidance based on individual challenges, enhances user learning and efficiency in voice input.

The system was evaluated with 20 older adults, comparing performance with and without the tutoring system. Findings show significant improvements in completion time, number of input stumbles, and user feedback regarding ease of correction and learning confidence.

Personally, I think this is a good approach to train older adults who want to use more advanced technology and utilize all the features a smartphone provides. I like that based on the input, more instructions are given or highlighted on the screen which makes it easier for adults to see and follow instructions that are simple. Some advantages I thought were that it seems simple, has a personalized approach for each individual, and can improve user confidence. While the system offers personalized instructions, its reliance on text and key highlighting may pose challenges for older adults with visual impairments or those who like other approaches such as audio feedback.

Figure 1: Screenshot showing instructions after input stumble

Toshiyuki Hagiya, Keiichiro Hoashi, and Tatsuya Kawahara. 2018. Voice Input Tutoring System for Older Adults using Input Stumble Detection. In Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI '18). Association for Computing Machinery, New York, NY, USA, 415–419. https://doi-org.offcampus.lib.washington.edu/10.1145/3172944.3172995

Technique #4 Name - FlexType: Flexible Text Input with a Small Set of Input Gestures (Sovannara Tav)

Inputting text on touchscreen mobile devices such as smartphones, tablets, and smartwatches involves the use of an on-screen virtual keyboard and tapping on specific keys to produce your desired word. However, given that touchscreen mobile devices are shaped in small form factors, the keys on the on-screen virtual keyboard of these types of devices are smaller compared to the keyboard keys on laptops and computers, requiring more precision and attention when tapping each specific key. This presents challenges for certain types of people, thus the motivation and why the technique of inputting text with the use of finger gestures is needed is to allow people to input text on their touchscreen mobile devices in situations where precise tapping on the keys of the on-screen virtual keyboard is impractical or unavailable.

As for who the technique would be good for and its target audience, this includes people with various degrees of visual impairment such as those who are blind and those who can't see small keys on the on-screen virtual keyboard well enough due to their age. This also includes people who can't communicate with their voice and rely on text-to-speech software as well as people who are limited to only moving their fingers when inputting text.

The technique utilizes an ambiguous keyboard with four-character groups which are (a, b, c, d, e), (f, g, h, i, j, k, l, m), (n, o, p, q, r), and (s, t, u, v, w, x, y, z,'). Upon entering a text entry input field, the user taps between one to four fingers to designate one of the four character groups. From there, the user selects a character from the corresponding character group by tapping in a sequence. To backspace a single character, they swipe left with one finger. To backspace all characters of the current word, they swipe left with two fingers. To indicate that the current word is finished, they swipe right with one finger. When this occurs, the user can swipe up with one finger to replace the current word with the next most likely word from a list or they can swipe down with one finger to replace the current word with the previous most likely word from the list. To indicate they are done inputting into the text entry field, the user swipes down with two fingers.

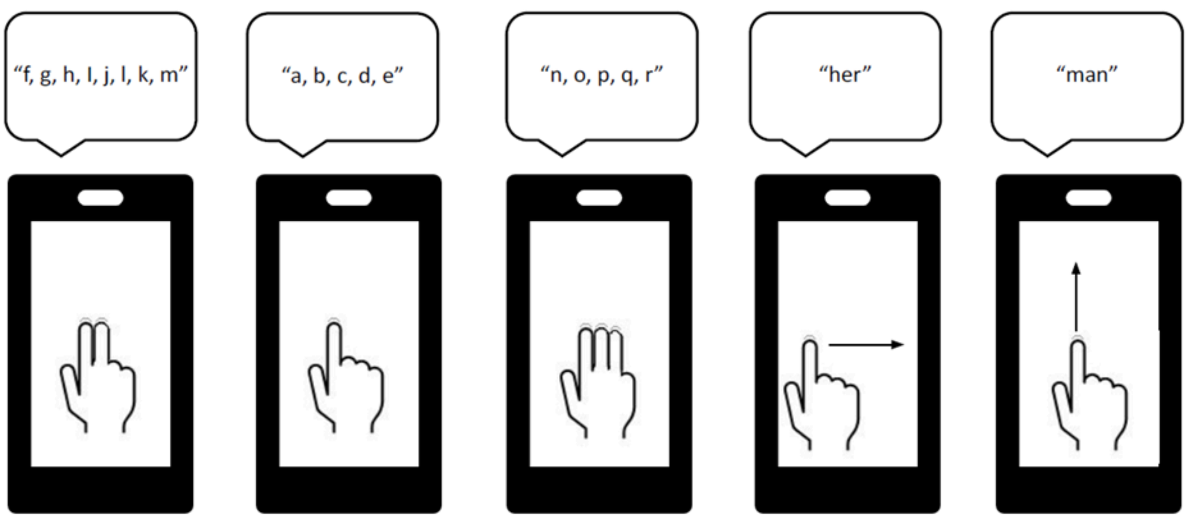
Important design features of this technique are upon entering a text entry input field, the FlexType interface appears as a solid black background taking up the entire size of the touchscreen mobile device to indicate to the user they can provide input with finger gestures. In addition, text-to-speech is incorporated to read out the characters to again accommodate those with various degrees of visual impairment. What makes this technique "work" is that people have already built muscle memory when interacting and navigating with touchscreen interfaces such as moving left, right, up, and down by swiping horizontally and vertically and performing actions by tapping, holding, and moving while holding onto the touchscreen.

For the participants using this technique, they completed a total of eight sessions, each lasting an hour and conducted on a separate day. Within each session, participants took a break approximately every ten minutes to reduce fatigue after using their fingers extensively. For the first session, participants received single-letter prompts meant for them to learn the character groups individually. In the second session, participants received single-word prompts to learn using the character groups together. In the third session, participants received phrase prompts containing no more than four words, and each word no longer than six characters. For sessions four to eight, phrases were only restricted to be no longer than six words.

This technique compared between a constrained character group layout mentioned earlier and an unconstrainted character group layout: (a, d, f, h, k, q, y, '), (b, c, e, i, j, n, x), (g, l, o, s, v, w), and (m, p, r, t, u, z). The main findings were for a constrained character group layout, it achieved an average of 12.0 words per minute (WPM) with an average 2.03% character error rate compared to an average 13.5 WPM and an average of 1.81% character error rate for the unconstrained character group layout. Though the unconstrained character group layout performed slightly better, the researchers concluded the preference for the constrained character group layout to avoid the high barrier to entry that the unconstrained character group layout creates. Other metrics apart from WPM for speed and CER for errors that were also measured include backspaces per character (BPC), the total number of up and down swipes when replacing the current word with the next or previous most likely word, the tap sequence and context to determine the distribution of the target word in the n-best list, and responses from the participants on their thoughts of the technique.

In assessing the technique of inputting text with the use of finger gestures, a positive aspect of this technique is it accommodates those with various degrees of visual, speech, and movement impairments. Another positive aspect is using a structure people are familiar with when interacting and navigating touchscreen interfaces such as swiping horizontally and vertically as well as tapping, holding, and moving while holding onto the touchscreen. A negative aspect of this technique upon entering a text entry input field, the FlexType interface appears as a solid black background taking up the entire size of the touchscreen mobile device which I think can be inconvenient as there can be some cases where it may be faster to dedicate half the screen for displaying text and the other half for the interface rather than the interface taking up the entire screen and causing the user to rely on the text-to-speech to identify on what to type. Another negative aspect is no form of customization is mentioned regarding the figure gestures for people who are left-hand dominated, being able to add additional figure gestures, disable or delete current figure gestures, and many more forms of customization.

Figure 1: An example input sequence for the word ‘man’. From left to right, the user enters each character by tapping with two fingers, one finger, and then three fingers. The letters corresponding to a group are read out after each tap. The user then swipes to the right and the word ‘her’ is recognized as most likely and read out. The user then swipes up to change ‘her’ to the second most likely word, ‘man’.



Dylan Gaines, Mackenzie M Baker, and Keith Vertanen. 2023. FlexType: Flexible Text Input with a Small Set of Input Gestures. In Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23). Association for Computing Machinery, New York, NY, USA, 584–594. https://doi-org.offcampus.lib.washington.edu/10.1145/3581641.3584077

Bibliography

Dylan Gaines, Mackenzie M Baker, and Keith Vertanen. 2023. FlexType: Flexible Text Input with a Small Set of Input Gestures. In Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23). Association for Computing Machinery, New York, NY, USA, 584–594. https://doi-org.offcampus.lib.washington.edu/10.1145/3581641.3584077

Michael Shorter, Bettina Minder, Jon Rogers, Matthias Baldauf, Aurelio Todisco, Sabine Junginger, Aysun Aytaç, and Patricia Wolf. 2022. Materialising the Immaterial: Provotyping to Explore Voice Assistant Complexities. In Proceedings of the 2022 ACM Designing Interactive Systems Conference (DIS '22). Association for Computing Machinery, New York, NY, USA, 1512–1524. https://doi-org.offcampus.lib.washington.edu/10.1145/3532106.3533519

M. N. Giannakos, S. Lee-Cultura and K. Sharma, "Sensing-Based Analytics in Education: The Rise of Multimodal Data Enabled Learning Systems," in IT Professional, vol. 23, no. 6, pp. 31-38, 1 Nov.-Dec. 2021, doi: 10.1109/MITP.2021.3089659. keywords: {Learning systems;Education;Human computer interaction;Adaptation models;Predictive models;Reflection;Real-time systems}

Toshiyuki Hagiya, Keiichiro Hoashi, and Tatsuya Kawahara. 2018. Voice Input Tutoring System for Older Adults using Input Stumble Detection. In Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI '18). Association for Computing Machinery, New York, NY, USA, 415–419. https://doi-org.offcampus.lib.washington.edu/10.1145/3172944.3172995