From Accessibility to Empowerment: Rethinking the Role of Voice Control In GIS

Barry J. Kronenfeld [1]*, Gaurav Sinha [2]

[1] Geology and Geography, Eastern Illinois University, Charleston, USA
[2] Department of Geography, Ohio University, Athens, USA
* bjkronenfeld@eiu.edu

Keywords: accessible computing, voice control, GIS, productivity

Introduction

This study poses the question, can a modern GIS be controlled entirely by voice?

Voice control (VC) has been studied as an accessibility tool for people with repetitive strain injuries (RSI) and other disabilities who are unable to use keyboard and mouse (KM) (Bhalerao et al. 2017; Pant & Panigrahi 2022). VC is simpler and more versatile than other accessible input mechanisms such as wrist bands, eye gaze and head movement tracking sensors because it builds upon cognitive language skills and requires only a microphone. Focusing only on physically disabled users is limiting, however, as it assumes a false dichotomy between disabled and able-bodied people (Guelke 2003) when in fact most technology workers experience some degree of RSI (Kayode, Adewale & Lawal 2015; Singh & Borkar 2020; Zheng et al. 2023). VC is also an essential component of multimodal and hands-free computational environments (e.g., mobile device operation, AR/VR environments, gaming, field data collection).

While there has been intermittent interest in voice enabled geospatial interfaces, prior research and applications have focused on a limited set of functions for specific circumstances such as field data collection (Kemper and Ballieu 2003) and mobile navigation (e.g. voice control for Google Maps). Recent interest in AI agents in GIS that can process natural language input is also promising, but is presently focused on developing specialized agents for domains for which AI capabilities are most promising (Kang 2025). This does not directly address the goal of avoiding KM usage, since in many cases even interaction with the agent requires KM. The authors are not aware of sustained effort to design a VC first interface for GIS and mapping software that aims for 100% parity with KM, i.e. enabling every single operation to be performed by voice without any KM usage.

Both authors have experienced computer-related RSI for many years. In 2022, the first author began exploring VC for personal and professional computational experiences. After experimenting with Dragon Naturally Speaking, Windows 11 Voice Access and Window Speech Recognition, the decision was made to work with Talon Voice (https://talonvoice.com) due to its sophisticated grammar capabilities and high degree of extensibility. While other systems excel at dictation or provide basic assistance with general, predefined computer navigation tasks, Talon Voice enables users to construct their own voice control command repositories with fine control over command grammar and functionality limited only by the reach of the two integrated scripting languages (TalonScript & Python). After developing competence with basic computer control, the idea was developed to create a Talon repository to perform KM-intensive GIS work, and Talon GIS (TaGIS) was born.

Through the TaGIS project we explore the question: can the entire corpus of a modern GIS be made accessible through a VC system without requiring KM? What would such a system look like, and can productivity approach that of standard KM workflows? While we envision the eventual integration of multiple modalities, combining VC, KM, eye tracking and other input mechanisms to create a truly multimodal GIS, for now our research focuses on an exclusive VC system to assess the possibilities and limits of voice as an input mechanism.

TaGIS System Description

TaGIS is constructed using Talon Voice, a proprietary application and scripting framework that "aims to bring programming, realtime video gaming, command line, and full desktop computer proficiency to people who have limited or no use of their hands, and vastly improve productivity and wow-factor of anyone who can use a computer". Talon does not actually include any built-in voice commands, but enables users to construct their own command grammar using a combination of Talon's own scripting language and python. A vibrant user community exists, with a large corpus of spoken commands made available through a "Community" repository (https://github.com/talonhub/community).

Building on the talon community repository, TaGIS implements actions triggered by spoken commands that control ArcGIS Pro through talon and python scripts. Actions are translated by the Talon API into sequences of keyboard, mouse and Windows UIAutomation (Microsoft n.d.) operations (Figure 1).

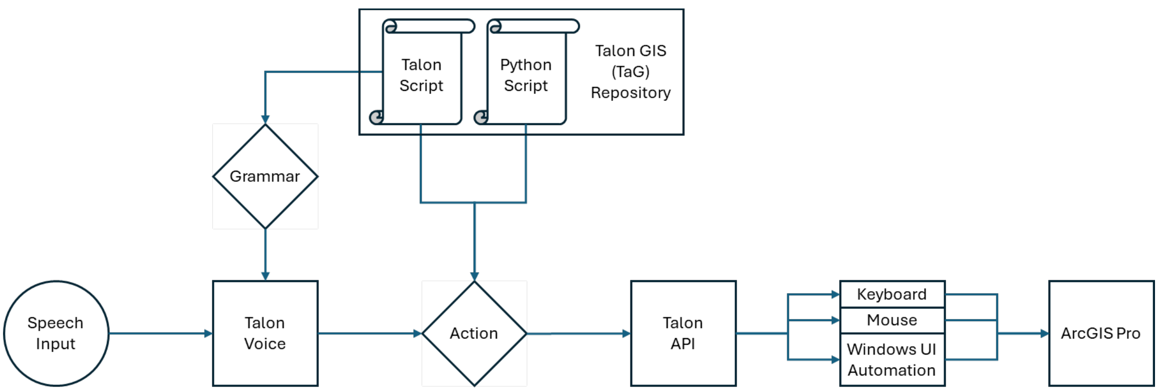

Figure 1: Simplified schematic of Talon GIS (TaGIS) processing of speech input.

To illustrate how TaGIS works, consider the process of selecting different "list by" views of the Contents pane in ArcGIS Pro (Figure 2). To begin, the user might search for commands containing the words 'list by' with the help command "help search 'list by'" (from the Talon Community repository). This brings up an overlay help dialog displaying seven matching spoken commands (Figure 2a).

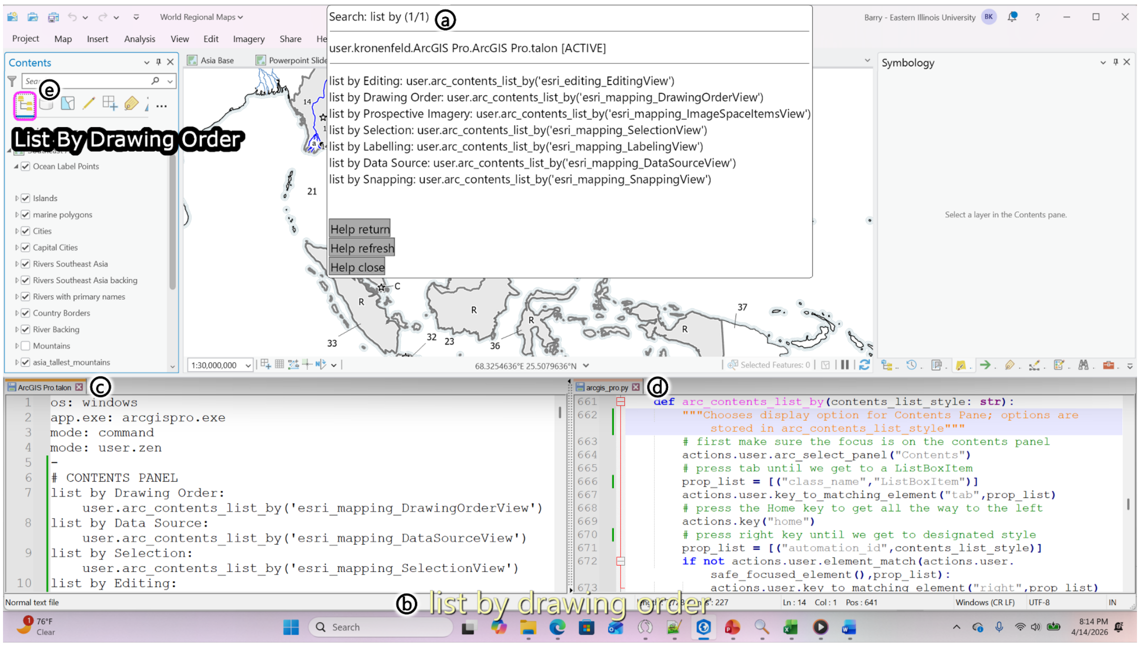

Figure 2: Screenshot of elements of TaGIS related to the action of selecting a different "list by" view option for the Contents pane in ArcGIS Pro.

Once the command "list by drawing order" is spoken and recognized, the command is displayed in a subtitle near the bottom of the screen (Figure 2b). All talon commands consist of spoken forms and corresponding actions defined in *.talon files. Figure 2c shows the talon file defining the command "list by drawing order" (line 7). The action triggered by this command is a user-defined python function *arc_contents_list_by()*, the definition of which resides in a separate python file (Figure 2d). This function calls other user defined functions including *arc_select_panel* which navigates the ArcGIS Pro interface using UIAutomation, and *key_to_matching_element* which presses a specified key (e.g. tab, right-arrow) until a UI element with specified properties is reached.

To help the user keep track of the software state, TaGIS polls Windows UIAutomation multiple times per second for the currently focused element. This element is automatically highlighted with a rounded magenta rectangle and optionally labelled. In this example, the highlight and label confirms successful selection of the "List by Drawing Order" toggle button to the user (Figure 2e).

Some other grammar patterns implemented in the TaGIS repository are listed in Table 1. Grammar patterns are defined modularly with parameters to allow efficient definition and implementation. For example, in the pattern "panel <pane_name>" the parameter <pane_name> can take on many values including "contents," "catalog," "symbology," etc. This allows code reuse, so a single command definition encapsulates many distinct spoken commands. Grammar patterns also facilitate learning and memorizing the available spoken forms, as these will be identical except for the variable parameter (e.g. the name of the pane). Screen recordings of several representative grammar patterns are available for the curious reader [here](#).

Command Group	Sample Grammar Pattern	Sample Spoken Command	Corresponding Action
UI Nav	panel <pane_name>	"panel contents"	place focus on Contents pane
	layer select <lyr_name>	"layer select rivers"	place focus on Contents pane, select layer with alias beginning with "rivers"
Ribbon	menu <rbn_hdr>	"menu insert"	expand Insert ribbon
	<rbn_hdr> <rbn_ctrl>	"insert legend"	expand Insert ribbon, invoke Legend button
Symbology	symbol <sym_grp>	"symbol structure"	navigate to Symbology pane, place focus on Structure tab
	symbol <sym_ctrl> <val>	"symbol fill color purple"	navigate to Symbology pane, open fill color dialog, set hex code value, close dialog
Query	add <conjunction> clause	"add OR clause"	invoke Add Clause button, select "Or" in conjunction combo box
	clause <conjunction> <field> <comparison> <constructed_text>	"clause or name is equal river Danube"	set values of combo boxes in current query clause
Mouse Sim	map go to <lon> <lat>	"map go to 90 east 25 north"	expand Map ribbon, invoke *Go To XY* button, enter lon/lat values into corresponding text input boxes
	place on layout	"place on layout"	move mouse to drag rectangle onto current layout (to place map frame, legend, etc)

Table 1: Examples of command groups, grammatical patterns, commands and actions implemented in the TaGIS repository.

As a preliminary assessment of productivity potential, TaGIS was used to construct a series of map quizzes and study guides for six different regions for an undergraduate course in World Regional Geography over the span of several semesters. The question was whether persistent use of TaGIS without KM was sustainable in a real-world work setting. The task of map construction was selected because cartographic production is typically mouse-and-keyboard intensive, so the ability to meet deadlines and teaching standards without reverting to KM would indicate potential productivity parity between these input mechanisms.

Not every GIS operation was equally efficient using VC as with KM. It was difficult to replicate the fine motor control of mouse movements using VC, so digitizing and construction of ad hoc polygons required significantly more time with VC. However, this was offset by the ability to chain together multiple KM and UIAutomation actions within a single voice command. For example, designating the fill color of a polygon layer typically requires seven mouse clicks but could be triggered with a single spoken command that was logical, easy to remember and enabled specification of any color in the standard matplotlib list of named colors.

Overall, the lead author was able to complete the development of all course materials entirely by voice as initially intended, despite time spent simultaneously on repository development and usage. Although not a rigorous quantitative analysis, the experience lends confidence to the thesis that it is not only possible to control virtually every aspect of a modern GIS using VC, but that productivity levels can potentially approach or exceed that of standard KM input.

Discussion

While our research illustrates the possibility of 100% parity between GIS functions accessible by VC & KM, realizing this potential will require effort and coordination within the GIS community and beyond. In addition to numerous technical challenges, work is needed to improve efficiency, reduced learning curves and promote a wider community of use. We are working to develop standardized, logical semantic grammar patterns for controlling GIS that can be standardized across applications. The visual GIS interface should also be augmented or even reimagined to support VC by making available commands easy to identify and commit to memory. Other areas for future research include development of mechanisms for audio communication of GIS state and map elements to support monitor-free GIS, as well as mechanisms for fine spatial control using voice. VC already relies on AI for voice recognition, but additional research into AI models to identify and avoid sources of ambiguity and misrecognition in spoken grammar and to translate pure natural language into likely intended deterministic commands would further enhance the user experience. Numerous other minor obstacles to VC execution will require the attention of GIS software and OS vendors, some of whom are already moving in this direction. By our assessment, however, technical capabilities are already sufficient for productive control of GIS using voice. We intend to continue to develop the TaGIS system and make it available to the broader GIS community, to begin laying the theoretical groundwork for a hands-free cartographic future.

References

Bhalerao, A., Bhilare, S., Bondade, A. & Shingade, M. (2017). Smart Voice Assistant: a universal voice control solution for non-visual access to the android operating system. *International Research Journal of Engineering and Technology*, 4(1):1713-1720.

Guelke, J.K. (2003). Road-kill on the information highway: repetitive strain injury in the academy. *The Canadian Geographer*, 47(4):386-399. doi: 10.1111/j.0008-3658.2003.00031.x

Hammersley, B. (2020) A Voice User Interface for ArcGIS using REST and Alexa. Esri Developer Blog. https://resource.esriuk.com/blog/a-voice-user-interface-for-the-arcgis-platform-using-rest-and-alexa

Kang, Y. & Wang, C. (2025). Envisioning Generative Artificial Intelligence in Cartography and Mapmaking. International Cartography Conference, Vancouver, August, 2025. https://arxiv.org/html/2508.09028v1

Kayode, A.A., Adewale, J.A., Lawal, N.T.A. (2015). An exploration of prevalence of repetitive stress injuries among computer operators in Nigeria. *International Journal of Computer Applications*, 109(1):1-5. doi: 10.5120/19149-0568

Kemper, I.C. & Ballieu, B. (2003). Voice, A New Interface to GIS. Proceedings, Esri User Conference, 2003. https://proceedings.esri.com/library/userconf/proc03/p0959.pdf

Microsoft. (n.d.). UI Automation. Microsoft Learn. https://learn.microsoft.com/en-us/windows/win32/winauto/entry-uiauto-win32

Pant, S. & Panigrahi, N. (2022). Designing smart voice command interface for geographic information system. In M. Panda et al. (eds.), ICIICC 2022: Innovations in Intelligent Computing and Communication, CCIS 1737, pp. 95–112. doi: 10.1007/978-3-031-23233-6_7

Ruan, S., Wobbrock, J.O., Liou, K., Ng,A. & Landay, J.A. (2018). Comparing speech and keyboard text entry for short messages in two languages on touchscreen phones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1 (4), 23pp. doi: 10.1145/3161187

Singh, R.M. & Borkar P. (2020). Prevalence of work-related musculoskeletal disorders among IT professionals in India - A literature review. *International Journal of Research in Medical Science*, 8(10):3765-3771. doi: 10.18203/2320-6012.ijrms20204271

Wang, C., Kang Y., Gong, Z., Zhao, P., Feng, Y., Zhang, W. & Li G. (2025). CartoAgent: a multimodal large language model-powered multi-agent cartographic framework for map style transfer and evaluation. *International Journal of Geographical Information Science*, 39(9):1904-1937. doi: 10.1080/13658816.2025.2507844

Zheng, B., Chen, F., Wang, J., Deng, H., Li, J., Zhou, C. & Ye, M. (2023). The Prevalence and Correlated Factors of Occupational Stress, Cumulative Fatigue, and Musculoskeletal Disorders among Information Technology Workers: A Cross-Sectional Study in Chongqing, China. *Healthcare*, 11:2322, 15pp. doi: 10.3390/healthcare11162322