Finding a New Voice: Transitioning Designers from GUI to VUI Design

Christine Murad
University of Toronto, Toronto, Canada
christine.murad@mail.utoronto.ca.

Cosmin Munteanu
University of Toronto Mississauga, Mississauga, Canada
cosmin.munteanu@utoronto.ca

Benjamin R. Cowan
University College Dublin, Dublin, Ireland
benjamin.cowan@ucd.ie

Leigh Clark
Swansea University, Swansea, UK
l.m.h.clark@swansea.ac.uk

ABSTRACT

As Voice User Interfaces (VUIs) become widely popular, designers must handle new usability challenges. However, compared to other established domains such as Graphical User Interfaces (GUIs), VUI designers have fewer resources (training support, usability heuristics, design patterns) to guide them. On the other hand, GUI-trained designers may also be solicited upon to design VUIs given the increased demand for such interfaces. This raises the question: how can we best support such designers as they transition from GUI to VUI design? To answer this, we focus on usability heuristics as a key resource, and conduct several workshops with GUI design experts, exploring how they map their design experience onto VUI design. Based on this, we suggest that the "path of least resistance" to transitioning designers from GUI to VUI may be the adaptation of familiar resources and concepts (such as GUI heuristics) to the VUI design space, instead of the imposition of novel VUI-specific heuristics on GUI-trained designers. This finding can inform the development of design resources that can support the increase demand for VUIs.

CCS CONCEPTS

• **Human-centred computing**; • **Human computer interaction (HCI)**;

KEYWORDS

Voice user interfaces, Usability heuristics, Speech Interfaces, Design guidelines

1 INTRODUCTION

Voice user interfaces (VUI) are increasing in popularity, with devices such as Amazon Alexa, Google Home, and Apple Homepod becoming widely available consumer products. Voice has been described as a natural form of interacting with technology [18, 38], with long-running predictions of VUIs becoming the future of interaction [6]. However, as their popularity grows, designing interactions with VUIs has progressed only incrementally, despite significant engineering advances (such deep learning approaches applied to speech recognition) [14]. Currently, VUIs fall short of the advertised "natural" and "conversational" functionality that users have come to expect from these devices [1, 10, 12, 25, 32]. Issues such as a lack of system feedback and difficulties in correcting errors are still prevalent within current VUI systems [12, 25]. This causes current VUIs to not match high user expectations, or in the case of many underrepresented groups, increases the barriers to adoption [12, 18, 20].

VUIs are still relatively new as mass-market consumer products. Consequently, established standards for VUI design are lacking in this research area [10]. As such, a large majority of established interface designers may not have been formally trained in VUI design [22, 23]. Many designers currently tasked with developing VUIs or components of VUIs (e.g. Alexa skills) are likely to be relying on established practices adopted from Graphical User Interface (GUI) design, which has encompassed the majority of HCI design training [21]. However, designing for voice interaction poses many new usability challenges that do not exist in GUIs [25, 46], with audio and speech being the primarily (sometimes the only) medium of interaction. With the increasing demand for the development and integration of VUI interfaces into people's daily lives, the demand is growing rapidly for proper tools, methods, and practices is growing rapidly that can aid designers in developing for a VUI paradigm.

Usability Heuristics are an example of a widely-used and adopted tool for designing user interfaces. While heuristics already exist for the design of GUIs [27, 30, 37], VUI heuristics are not yet widely available - or at best, are in their infancy. The few available VUI heuristics are not yet settled into a widely adopted and universally understood canon (with research on what such VUI heuristics may consist of also not yet completed [24]). While summative usability evaluation methods are often used for VUIs, heuristics allow designers to ground the entire process (e.g. design, formative research, summative evaluations) in a more comprehensive framework, which necessitates the exploration of developing widely-adoptable VUI heuristics. Though research shows that existing

GUI-based heuristics cannot be mapped directly "as is" to VUI interaction [24, 35, 46], recent empirical work on VUI usability suggests that some existing GUI heuristics can provide a groundwork for the development of VUI heuristics [44]. Therefore, VUI heuristics may not need to be developed from scratch. There are also critical arguments in recent speech-related HCI literature that GUI heuristics can evolve or be adapted to VUIs [24, 25].

The commercial interest in developing "voice-first interactions" is increasing at a large rate, forcing a very large number of GUI-trained designers to adapt "on the fly" to designing VUIs – mostly without any guidance on how best to do so [22]. This strongly suggests that, while it is important to continue research on developing VUI heuristics from scratch, we may need to offer a "path of least resistance" that leverages the vast body of knowledge related to GUI heuristics, as opposed to creating VUI heuristics primarily from scratch. We explore this "path of least resistance" by answering the following questions in this paper:

Given that most designers have currently been trained in GUI design, how do they interact with and handle VUI usability issues, given their existing GUI training and experience with existing tools (e.g. GUI heuristics)?

What barriers do GUI-trained designers face when trying to interact with, understand, and apply existing heuristics (both GUI and VUI) to VUI usability issues?

What is the "path of least resistance" to helping designers to transition into the VUI design space, and how should we create tools (such as heuristics) to aid these designers?

The workshops demonstrated that there is a path of least resistance that can be followed to help designers transition from GUI to CUI design. This path uses both existing GUI heuristics and experience from GUI design training in order to help inform VUI heuristics that are easy to understand and adopt in practical design. Rather than presenting designers with completely novel tools, we found having heuristics that follow similar language and concepts that are found in GUI heuristics and methods (that designers are already trained in) provides the easiest path for transition GUI designers into the VUI design space – the broad academic and industry practice of designing voice interfaces.

1.1 Research Goals and Contributions

Previous VUI heuristics in literature have often been developed from scratch, by collecting empirical VUI usability issues and designing heuristics around 'solutions' for these issues. This can be problematic for designers who are transitioning into the VUI design space, as this can impose the adoption of tools that require a familiarity in VUI design that many current designers do not have. This has been shown to be an ineffective approach to developing new heuristics for VUIs in previous research, particularly for designers who are not already well-versed in speech interfaces [43]. In this paper, we show instead that a "path of least resistance" exists by exploring the design knowledge and experience of current GUI designers and using it to form VUI heuristics. Due to a higher familiarity factor for current GUI-trained designers, this may be a better venue to explore for creating VUI heuristics that can be widely validated and adopted after their creation. To do this, we present the findings of seven participatory workshops with

usability experts, consisting of cognitive walkthroughs, heuristic evaluations, and participatory design. These workshop activities aimed to learn how GUI experts interact with VUI usability issues and apply existing GUI experience and heuristics to them, and how GUI experts conceptualize and understand recently proposed VUI heuristics in connection with existing familiar GUI heuristics. From these, we provide recommendations for how VUI heuristics should be developed in order to provide such "path of least resistance" for transitioning designers. Our paper proposes a new way to better support designers as they transition from one paradigm to another, using empirical evidence from the seven workshops we conducted with designers. This has the potential to address several of the barriers toward the use of heuristics for VUI design.

Our primary contributions in this paper are as follows:

- By exploring how GUI designers interact with and handle VUI usability issues (RQ1), we contribute an understanding of the challenges that GUI-trained designers have when transitioning to handling VUI usability issues;
- By identifying barriers that GUI-trained designers experience while interacting with applying existing heuristics to VUI issues (RQ2), we provide insights into their ability to engage with GUI and VUI heuristics as a design / evaluation resource for VUI issues, & what pathways may be feasible in supporting such transitions for GUI-trained designers, including re-framing the development of new VUI heuristics into more familiar constructs;
- Using results from the two contributions above, we provide recommendations for how to best support designers as they transition to the VUI design space, through following a "path of least resistance" in the development and adoption of VUI-specific heuristics

1.2 Related Work

Research on designing usable VUIs is more limited than for other interaction types, as only recently have significant inroads been made in the VUI consumer space [9]. To better contextualize the central argument of this paper, we survey the literature on the usability and design of VUIs, as well as efforts to develop VUI-specific heuristics. However, before we explore the VUI heuristic space, a broader discussion of heuristics and other design resources is needed, along with current efforts to explore and improve VUI interaction.

1.3 Usability Heuristics

Many resources exist to help designers create usable interfaces. One such resource are heuristics, used both during design (to "help choose among alternative designs") and during formative and summative usability assessments (to "find and justify interface problems") [48]. Some of the most established ones in GUI design are the ones by Nielsen [27], Norman [30], and Shneiderman [37]. They are important on a fundamental level and can help both conceptualize a system and evaluate its usability. They can also provide the theoretical foundation on which to develop more practical tools (e.g. design patterns) that help build and assess interfaces.

For example, the heuristic of "Visibility of System Status" [27] provides a conceptual design guideline that applies to an entire

interaction – specifically, always providing feedback of what the system is doing to the user. In contrast, other resources such as design patterns or "how-to" guides [2, 16] are concerned with the more specific "building blocks" of designing an interface.

Our work focuses on the former – design and usability heuristics. We acknowledge that both these resources are important to the designing any interaction. However, we argue that it is important to have a conceptual framework of high-level principles for VUIs, like the ones that exist for GUIs. The creation and understanding of tools such as heuristics can provide the foundation that designers can use to aid them in designing better voice interfaces – particularly in that period of transitioning from a GUI design paradigm to a VUI design paradigm.

1.4 Usability of Voice Interfaces

Previous research has greatly explored the usability of VUIs, particularly in the last 5-10 years. Infrequent users of personal voice assistants encounter numerous usability challenges [12]. These include difficulties with the amount of information that can be remembered [36], system feedback [12, 18] and recognition errors [12]. Many difficulties are a consequence of users' expectations of interacting with VUIs versus the reality of using existing commercial VUIs [18]. Learnability – learning how to use a VUI - is another pressing issue in current VUI design [13, 26, 46]. Schneiderman [36] highlighted key issues such as the cognitive effort required to interact with them, problems with presenting information through speech, and resolving errors due to speech recognition. Nielsen [29] and colleagues [7] highlights that the usability of VUIs limits their usefulness, and that visual interfaces surpass voice interfaces in usability for most cases (besides hands-free interaction or for people with physical disabilities). Even with research and technological advances over the past decade, users continue to experience many usability issues with VUIs, which often lead to lack of adoption or abandonment [12, 18, 20].

1.5 Improving VUI Interaction Design

Much work has been performed over the past few decades that attempts to improve VUI Interaction design. Some researchers have explored the attitudes, perceptions, and experiences of VUI usability by current users [12, 18, 32]. Others have attempted to address VUI usability issues directly by working to develop solutions to specific VUI usability issues, such as creating tutorials and contextual help to improve learnability [11, 13, 46], and reducing the memory load required in VUI interaction [45]. Others have tried addressing VUI design more holistically, by developing design guidelines, heuristics, and methods for VUI systems [24, 40, 43]. Recent research has highlighted the importance of training VUI designers as a means to improve VUI usability issues and have begun to explore VUI design training in HCI Education [22]. As experienced designers transition from a GUI to a VUI design space, and new designers enter the VUI industry, there currently lacks appropriate training and education for VUI design. A previous syllabi review and meta-analysis showed that there is very little discussion and training on VUIs in current HCI curricula [22]. This in particular makes it very difficult for designers to transition appropriately from the familiar space of designing GUIs to designing VUIs.

1.6 Developing Usability Heuristics for VUIs

Usability heuristics are useful for evaluating and designing new UIs. As new interaction domains have arisen, researchers have developed appropriate new heuristics tailored to them. One method of developing usability heuristics for new interaction paradigms is exploring documented usability issues within that paradigm and generating design heuristics that encapsulate and seek to resolve these issues. This method was used in creating video-game specific heuristics, using usability issues identified in game reviews [31], and for telephone-dialogue interfaces, using a database of usability issues and design solutions [40]. Another method (addressed in this paper) is to take established usability heuristics and adapt them to a new interaction context. These are often grounded in Nielsen's [27] established heuristics for user interface design. This has been conducted in many fields, including for usable web pages [4], virtual reality applications [41], ambient displays [19], and touch-screen mobile devices [15].

Previous literature has explored developing design guidelines for VUIs. Rudnicky [34] developed seven guidelines for speech user interfaces (SUIs), however, they were directed at SUIs that were incorporated into graphical-first dialogue interfaces. Recently, two sets of preliminary guidelines have been proposed for designing voice-dedicated interfaces (which are the focus of our study). One is a set of telephone dialogue design guidelines, that was created from a database of usability issues and respective "design solutions" [40]. Another is a set of 17 VUI usability heuristics was created to inform the design and evaluation of speech-based smart devices [43], expanding on previous efforts (e.g. [40]). These have primarily followed a "top-down" approach of heuristic development, by taking documented issues and developing design heuristics around these issues, vs. a more bottom-up approach grounded in existing heuristics (such as GUI heuristics).

1.7 From GUI to VUI: Investigating the Path of Least Resistance

Little work explores how existing GUI heuristics may map over to VUI usability issues, and in turn how they may be used as a grounding for VUI heuristics. VUI interaction is very different from GUI interaction [16, 19, 37]. Thus, GUI interaction principles cannot be mapped directly. However, a recent review of speech work in HCI found that discussion of design issues often aligned significantly with GUI heuristics [24], though the literature does not explicitly discuss GUI heuristics themselves. Previous research has also identified the difficulty for designers not thoroughly experienced with speech design to adopt VUI heuristics that have been developed from scratch [43]. Given that current (primarily GUI-trained) designers in industry are increasingly being asked to develop voice interfaces, this poses a large problem for voice interfaces being developed today – of which the damage can be seen based on current research on the many usability issues still existent in VUIs today [12, 18, 26, 32]. Our paper seeks to address this critical design gap. The literature surveyed here suggests that, while developing VUI heuristics from scratch may still be necessary, it may not be sufficient to address this gap, given how widespread and entrenched GUI-specific training and GUI heuristics are. Hence, we seek to investigate a "path of least resistance" that more easily

Table 1: Participant demographics and experience with VUIs and GUI heuristics

ID	Gender	Position	Used VUIs	Designed for VUIs	Familiar with GUI heuristics
P1	M	Assistant Professor			✓
P2	F	Assistant Professor			✓
P3	F	Industry Designer		✓	
P4	M	Industry Designer	✓		✓
P5	M	Industry Designer			✓
P6	F	Post-Doctoral Fellow	✓		✓
P7	M	UX Developer	✓	✓	✓

bridges the GUI design resource space (e.g. heuristic principles) to the VUI space. The exploration of directly soliciting and incorporating design knowledge and experience of existing designers (many GUI-trained) has not previously been thoroughly explored in VUI heuristic development, particularly in early heuristic brainstorming and development [34, 40, 43]. Our study aims to do exactly this – to solicit early design knowledge and experience from GUI-trained designers to better inform the development of future VUI heuristics.

2 FROM GUI TO VUI HEURISTICS: AN EMPIRICAL ASSESSMENT

As the commercial development of VUIs grows, existing GUI trained designers may lack tools on how to approach the design of such an interaction, relying on adaptation of existing knowledge to inform design [22]. Our work aims to explore the feasibility of using GUI heuristics to scaffold the development of VUI heuristics, scoping the feasibility and challenges faced when using GUI heuristic tools to inform VUI design. This may have significant benefit for current designers as it will allow them to use existing knowledge to ease into the VUI design space. To explore this, we conducted one-on-one workshops with seven design experts, consisting of several usability activities with VUI devices, as well as participatory design sessions. These are described in detail below.

2.1 Participants

Seven usability experts participated in the workshops. Snowball sampling and social media was used to recruit both academic and industry design experts from the UX community in our area (Table 1). This sample is in line with typical in-depth, extensive qualitative research [8]. We consider this suitable for an investigation focused on exploring pathways toward developing VUI heuristics that current designers could adopt (instead of a formal validation of a proposed set of heuristics, which will be conducted after such pathways are identified). The study was approved as low srisk by our university's ethics board. All but one of the participants were familiar with GUI heuristics, and over half of the experts had either used VUIs before or had designed for them.

2.2 Methods and Instruments

We conducted seven individual one-on-one workshops with participants, that consisted of a variety of usability and brainstorming activities. Each workshop involved four phases:

Phase 1 - Cognitive Walkthrough with Commercial VUIs: This phase involved a cognitive walkthrough [28] with the Amazon Echo and Google Home. These devices were chosen due to current market presence and the push by various companies to develop apps or skills for them. Experts were given a list of tasks that people commonly perform with commercial VUIs like Amazon Alexa, such as controlling music, calendar appointments, setting alarms and general search [12, 17, 18, 42]. The given tasks were:

- Play "Imagine" by the Beatles
- Turn the volume up
- Turn the volume back down
- Pause the music
- Set a timer for 10 seconds
- Set an alarm for ten seconds from now
- Add an event for Lunch on September 27th at 1 PM
- Find out who the 15th prime minister of Canada is
- Find out how many milliliters are in 5 cups
- Get the temperature in Illinois, in both C and F

They were asked to perform the tasks in any order they chose with both devices, using whatever language they wanted to perform the task. While performing the tasks, they were asked to think of any usability issues that they encountered (they were provided with sticky notes to record notes for later discussion). This phase was very familiar in terms of setup and execution to what usability experts normally do in their line of work.

Phase 2 - Heuristic Evaluation using Existing GUI Heuristics: Experts were then asked to identify the VUI usability issues that they encountered while performing the tasks. They were given the opportunity to try new tasks they came up with on their own if they wished. After identifying these issues, they were asked to go through each issue and specify whether it violated one or more GUI heuristics.

We used a merged list [24] of Nielsen's [27], Schneiderman's [37], and Norman's [30] GUI heuristics, as these are well-established and widely used:

- G1: Visibility/Feedback of System Status
- G2: Mapping Between System and Real World
- G3: User Control and Freedom
- G4: Consistency throughout the Interface
- G5: Helping to Prevent User Errors
- G6: Recognition Rather than Recall
- G7: Flexibility and Efficiency
- G8: Minimalism in Design and Dialogue

Table 2: Wei/Landay (A) and Suhm (B) VUI Heuristics

ID	Heuristic title
A1	Give the agent a persona through language, sounds, and other styles
A2	Make the system status clear
A3	Speak the user's language
A4	Start and stop conversations
A5	Pay attention to what the user said and respect the user's context
A6	Use spoken language characteristics
A7	Make conversation a back-and-forth exchange
A8	Adapt agent style to who users are, how they speak, and how they are feeling
A9	Guide users through a conversation so they are not easily lost
A10	Use responses to help users discover what is possible
A11	Keep feedback and prompts short
A12	Confirm input intelligently
A13	Use speech-recognition system confidence to drive feedback style
A14	Use multimodal feedback when available
A15	Avoid cascading correction errors
A16	Use normal language in communication errors
A17	Allow users to exit from errors or a mistaken conversation
B1	Keep it simple
B2	Carefully control the amount of spoken output
B3	Word options the way users think
B4	Minimize acoustic confusability of vocabulary
B5	Provide carefully designed feedback
B6	Abide by natural turn-taking protocol
B7	Coach a little at a time
B8	Offer alternative input modalities
B9	Yes/no queries can be very robust
B10	Carefully select the appropriate persona

- G9: Allow Users to Recognize & Recover from Errors
- G10: Providing Help and Documentation

Phase 3: Brainstorming Potential New VUI Heuristics: After us-ability experts had gone through each of the VUI usability issues and identified violations based on GUI heuristics, they were then asked to engage in a participatory-design activity. It is important to note that the goal of this phase was not the creation of new VUI heuristics, but to explore how GUI experts engage in VUI heuristic development. They were asked to take the usability issues that were not categorized (i.e. found to not violate any of the GUI heuristics), and to brainstorm potential new heuristics that could encompass these issues. As discussed in the Related Work section, such heuris-tics have been proposed before, although these were not grounded in a collaborative approach that involved design experts from early on in the design process. Instead, we propose here pathways or approaches to develop VUI-specific heuristics. As part of exploring these pathways, we included this brainstorming phase with our expert participants.

Phase 4: Exploring Wei/Landay and Suhm's VUI Heuristics: Ex-perts were presented with both Wei/Landay's [43] and Suhm's [40] heuristics in their original format, which included the top-level de-scription of the heuristic (Table 2). Our goal was to explore whether GUI experts can understand and apply these existing VUI heuristics, giving much needed empirical validation and evaluation of the VUI

heuristics currently available. We also wanted to explore how these existing heuristics resonated with predominantly-GUI designers. Lack of knowledge of guidelines [47] and lack of time [33, 47] are often reasons for not following guidelines. Hence, we consider that engaging our expert participants around the top-level description of VUI heuristics (instead of one that may include extensive and long explanations) is ecologically more valid, as it mimics more closely what GUI-trained designers may face when asked to design a VUI. However, verbal clarification was provided to experts if asked.

First, they were asked to go through each of the two sets of VUI heuristics and identify which (if any) GUI heuristics overlapped with them. They were then asked to comment on how useful the VUI heuristics were for designing VUIs.

3 FINDINGS

We describe a qualitative analysis of the observations and direct par-ticipant transcripts from the workshops. Audio-recorded transcripts from the sessions were coded and analyzed for themes through inductive thematic analysis, using the framework approach [39]. Thematic analysis is a rigorous interpretative method that allows researchers to uncover themes from qualitative data [5].

The first few transcripts were coded to identify initial observa-tions, and then were organized into an analytic framework based on categories emerging from the initial codes. These were then
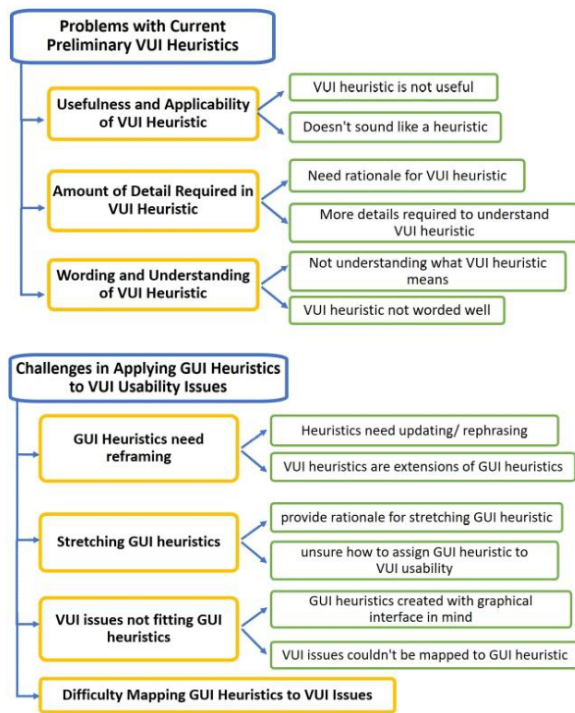
Figure 1: Themes (Blue), Sub-themes (Orange), and Codes (Green) Identified during Thematic Analysis.

applied to code the rest of the transcripts. This framework was guided by our research questions – the potential applicability of current GUI heuristics to VUI usability issues, and the exploration of existing VUI heuristics. After coding all the transcripts, codes were grouped into distinct sub-themes, and then summarized into two main themes using thematic analysis (as illustrated in Figure 1). Quantitative results were also collected, such as the amount of GUI heuristics that were violated by VUI usability issues and the number of overlaps between GUI & VUI heuristics. While the order we present the findings differs from the order in which we conducted the phases, we have chosen to present the findings in the following way based on the themes gathered from our thematic analysis (which sometimes bridge across several phases), and to illustrate the experience of experts applying heuristics to VUI issues (and vice versa) in a narratively coherent manner.

3.1 How GUI Designers Understand and Apply Existing VUI Heuristics

We first describe how GUI designers interpret existing recently-proposed VUI heuristics, and the challenges experts found with applying them to VUI issues (RQ2). These findings come from the activity performed in Phase 4 of the participatory workshops. The existing VUI heuristics we are referencing here are the ones listed in Table 2. The sub-themes described here are those encompassed by the theme "Problems with Current Preliminary VUI heuristics" in Figure 1

3.1.1 Usefulness & Applicability of VUI Heuristics to Design. Experts noted that some of the VUI heuristics were not very useful for designing a VUI. Sometimes it was not obvious to experts how they would be able to use an existing VUI heuristic: "Use multimodal feedback when available. It seems to me that it's always available. So, I don't know how I would use this one, because when I turn down the volume, I don't want it to tell me, because I already hear it." (P2)

Experts also expressed that some of the VUI heuristics were "not a design heuristic". Experts often expressed a perception of the format of a design heuristic and applied this to the VUI heuristics. This was applied with B9 - "Yes/no queries can be very robust": "It's just telling me that yes or no isn't very, it's too simple. That doesn't seem like a heuristic" (P7). This also often occurred with A4 – "Start and Stop Conversations": "Start and stop conversations...why is that a heuristic?" (P6).

Some heuristics were also noted to not really apply to the actual design of an interface: "Now carefully designed, that's not really a heuristic because what you're saying is, 'apply effort to the design of your interface'." (P5).

3.1.2 Wording and Understanding of VUI Heuristics. Experts mentioned that some of the VUI heuristics were hard to understand: "Confirm input intelligently, I don't know what intelligently means" (P5). This stood out with A4 – "Start and Stop Conversations": "I don't know exactly what it means. And actually, what is conversation, because it's not always clear. . ." (P2).

Often, this was because the VUI heuristic was not worded or phrased well to be applied usefully: "I guess queries from the system can be very robust . . . but it should be rephrased" (P2). Sometimes the wording of a heuristic left much to be interpreted: "...so keep it simple, what does that mean for somebody?" (P6). Some were noted to require re-wording: "Use speech recognition confidence, so the spirit of that heuristic is useful, the way it's worded is not." (P5).

3.1.3 Amount of Detail Required in VUI Heuristic. Experts noted that they wanted more details to understand the purpose and applicability of the VUI heuristics: "Keep it simple, that's ... it doesn't give me enough detail." (P5). Experts often asked for verbal clarification if they found it difficult to understand. Some of the heuristics were noted as simplistic and vague: "Provide carefully designed feedback, yes, but it's very vague. It should be a bit more refined." (P2).

To rectify these issues, experts wanted distinct descriptions and examples to assist them in understanding and applying the heuristic: "So I would like to have its rationale, why we need a persona, or why this persona has to be consistent . . . before I would use that as a design guideline." (P2). Some wanted a format that's reminiscent of existing familiar GUI heuristics: "That would be the title of that guideline. Then . . . a paragraph explaining what you mean by keeping it simple .... Followed by a couple of examples." (P1).

3.2 Types of VUI Usability Issues

Next, we describe the types of VUI usability issues that these GUI experts found while interacting with Google Home and Amazon
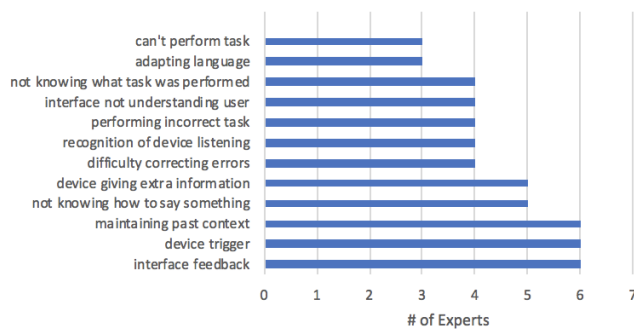
Figure 2: Types of VUI Usability Issues Mentioned

Echo. The findings in this section come from Phase 1, where experts were asked to identify VUI usability issues while performing a cognitive walkthrough with Google Home and Amazon Echo. Figure 2 illustrates the number of experts that mentioned a particular VUI issue in Phase 2. The top three categories were "interface feedback", "device trigger", and "maintaining past context". All issues mentioned in the chart are ones that have been identified in previous literature exploring usability issues with Voice User Interfaces [3, 12, 18, 26, 36].

3.2.1 Interface Feedback. Experts noted many usability issues with the audio and visual feedback of the two devices. While both were display-less, coloured indicators provided visual feedback for certain actions, such as volume changes, identifying when the interface was listening, or processing user input. Experts noted that both visual & audio feedback was sometimes hard to understand. In some cases, experts could not understand what the visual indicators were indicating: "The lights are doing something, but I don't know how to record that yet." (P2). In other cases, audio feedback did not provide enough information to the user about the task that was performed, such as the temperature degree scale used.

Experts noted that there wasn't enough feedback to know if the device understood or performed a task properly. In some cases, no feedback was provided at all: "Feedback is...they assume self-evident. So, if I ask Google to turn the volume up, and the volume goes up a little, they assume that's enough for me." (P5). At other times, device feedback was provided to show that an error occurred when performing a task, but experts did not know what the issue was: "Alexa. . . it gave a funny noise. . .when it couldn't perform the task I asked it to. . .[I] didn't know exactly what the problem was." (P6).

Another problem expressed was that visual feedback was not useful unless someone is looking at the device. P1 expressed their preference for not needing to look at the device: "Visibility & feedback ...I don't wanna be looking at it. . .they don't need to provide any visual feedback."

3.2.2 Maintaining Past Context. Experts identified issues concerning maintaining memory or context of previous requests. The most noted issue was that both devices contained no memory of previous requests: "The continuity is not there yet. It doesn't relate what I just asked back to what I asked a second ago" (P3). Due to this

issue, one cannot make requests or ask questions based on previous things that they had asked: ". . . there doesn't seem to be any memory. . . They don't expect that there might be follow ups" (P5).

Another issue discussed was the lack of undo capability: "See, without a universal undo command . . . The fact that you don't have a memory for these things, at least like, 'No, no, no. Take it back'" (P5).

3.2.3 Device Trigger. Experts experienced issues when dealing with trigger phrases ("Okay Google" and "Alexa") for the device to listen for input. Experts often had to ask what the triggers were for each device – especially if they were not already familiar with the device.

Experts also found it frustrating to have to trigger the device every time they wanted to speak to it: "It's like, 'Okay, Google,' every time . . . just gets annoying" (P7). Experts often forgot to say the trigger phrase first before speaking to the device. P1 expressed that it was not a habit to repeat someone's name each time they spoke to them. P5 also expressed frustration with this.

3.3 Applying GUI Heuristics to VUI Usability Issues & VUI Heuristics

Next, we describe how our experts applied GUI heuristics to both VUI issues and existing VUI heuristics. Here we report quantitative results illustrating how experts applied GUI heuristics to VUI usability issues, and how experts conceptualized and understood existing VUI heuristics from literature, along with how they connected GUI heuristics to them (RQ2). The findings in this section come from the activity performed in Phase 2 (VUI issues) and Phase 4 (VUI heuristics) of the participatory workshops.

3.3.1 Applying GUI Heuristics to VUI Usability Issues. Table 3a) shows how experts applied GUI heuristics to VUI usability issues. We found that almost every usability issue that was brainstormed by an expert was said to violate at least one of each GUI heuristic.

G1, G2, G3, and G7 were found to be violated the most across all experts. The VUI issues that these heuristics were found to violate (from Table 2) were:

G1: interface feedback; recognition of listening

G2: maintaining past context; adapting language

G3: not knowing how to say something; maintaining past context; performing incorrect task

G7: adapting language; maintaining past context

Two top usability issues from Figure 2 - "interface feedback" & "maintaining past context" – are represented in the four GUI heuristics mentioned above. While "device trigger" was an issue that experts struggled with, it was not often directly identified by experts as a usability issue – informing experts what the triggers were was often enough to solve that issue.

3.3.2 Connecting GUI Heuristics to Existing VUI Heuristics. Experts identified many overlaps between GUI heuristics and existing VUI heuristics. Almost every GUI heuristic overlapped with one VUI heuristic, except for G4 (Table 3b). G1, G2, and G9 were found to have the most overlap with VUI heuristics. Many overlapping themes were connected to usability issue themes identified in Figure 2 (A* and B* as defined in Table 4):

G1: interface feedback/system status (A1, A12, A14, B5)

Table 3: a) # of VUI issues that experts identified as violating GUI (G*) heuristics; b) # of VUI Heuristics that experts identified as overlapping with GUI Heuristics.

	Total
G1: Visibility/Feedback Of System Status	22
G2: Mapping Between System And Real World	20
G3: User Control And Freedom	20
G4: Consistency Throughout The Interface	9
G5: Helping To Prevent User Errors	4
G6: Recognition Rather Than Recall	15
G7: Flexibility And Efficiency	21
G8: Minimalism In Design And Dialogue	3
G9: Allow Users To Recognize And Recover From Errors	13
G10: Providing Help And Documentation	11

	Total
G1: Visibility/Feedback Of System Status	19
G2: Mapping Between System And Real World	31
G3: User Control And Freedom	10
G4: Consistency Throughout The Interface	0
G5: Helping To Prevent User Errors	14
G6: Recognition Rather Than Recall	1
G7: Flexibility And Efficiency	15
G8: Minimalism In Design And Dialogue	14
G9: Allow Users To Recognize And Recover From Errors	23
G10: Providing Help And Documentation	13

G2: conversational protocol & providing personas (A1, A3, A6 – A8, A16, B3, B6)

G9: error correction/communication (A14, A16, A17).

P7 was an outlier, finding a significant amount of VUI heuristics overlapping with G9. This is because P7 also interpreted heuristics that speak about proper feedback as overlapping with G9 (A6, A9, A13, B5).

VUI heuristics that overlapped with GUI heuristics often had very similar phrasing or wording. (e.g. "feedback" for G1 and "correction" or "error" for G9). The ones overlapping with G2 also maintained a theme of "conversational/spoken language constructs" or "personas". On the other hand, some of the GUI heuristics that mapped to VUI usability issues did not often overlap with VUI heuristics. One example is G3, where it was found to violate VUI usability issues 20 times, but only found to overlap with VUI heuristics 10 times. Another example is G4, where it was found to violate VUI usability issues 9 times but was not found to be overlapping with any VUI heuristics.

The top six VUI heuristics that overlapped with GUI heuristics were A7, A9, A14, A17, B6, and B7 (Table 4).

3.4 How GUI Designers Deal with VUI Usability Issues

Here, we describe in more qualitative detail how expert designers currently deal with VUI usability issues, particularly when trying to apply heuristics (GUI or VUI) to them. We also describe how their existing experience and knowledge of familiar established

GUI heuristics affects how they apply heuristics to VUI issues in this new paradigm (RQ1). The findings in this section come from the activity performed in Phase 2 of the participatory workshops. The sub-themes described here are those encompassed by theme "Challenges Applying GUI Heuristics to VUI Usability Issues" in Figure 1.

3.4.1 Usability Issues Not Fitting GUI Heuristics. Experts mentioned that some VUI usability issues did not fit well with GUI heuristics. These often existed because of the conversational nature of these devices. Certain types of non-verbal communication that accompany voice (such as gestures) cannot be captured by VUIs: "If I'm trying to describe something and I'm using either my hand to accentuate something. . .that's lost." (P1). Other issues that did not fit well were ones that involved devices being aware of their surrounding environment: "This environmental context. I couldn't categorize this in that at all" (P3).

Experts noted how GUI heuristics were designed with a non-audio interface in mind, making issues that were based on the medium of voice difficult to translate to GUI heuristics. P3 noted the time-based element of audio output being non-existent in graphical interfaces: "... [GUI heuristics] feel like they go to a 'web' kind of environment . . . if I was looking at a screen I could just ignore what I'm looking at, or I could scan through what I'm looking at. Whereas, with the device, it's more like it comes back with a verbal response, and there is a time element to a verbal response". P5 expressed the differences in how VUIs listen for input: "I mean, it's always listening to something, so this requires a different set of

Table 4: Number of GUI Heuristics Overlapping with VUI Heuristics (Suhm: A* and Wei/Landay: B*)

VUI Heuristic ID	VUI Heuristic	# GUI Heuristics
A1	Give the agent a persona through language, sounds, and other styles	1
A2	Make the system status clear	7
A3	Speak the user's language	5
A4	Start and stop conversations	6
A5	Pay attention to what the user said and respect the user's context	3
A6	Use spoken language characteristics	3
A7	Make conversation a back-and-forth exchange	9
A8	Adapt agent style to who users are, how they speak, and how they are feeling	2
A9	Guide users through a conversation so they are not easily lost	10
A10	Use responses to help users discover what is possible	5
A11	Keep feedback and prompts short	7
A12	Confirm input intelligently	4
A13	Use speech-recognition system confidence to drive feedback style	3
A14	Use multimodal feedback when available	8
A15	A15: Avoid cascading correction errors	7
A16	A16: Use normal language in communication errors	6
A17	A17: Allow users to exit from errors or a mistaken conversation	9
B1	B1: Keep it simple	5
B2	B2: Carefully control the amount of spoken output	4
B3	B3: Word options the way users think	5
B4	B4: Minimize acoustic confusability of vocabulary	3
B5	B5: Provide carefully designed feedback	7
B6	B6: Abide by natural turn-taking protocol	8
B7	B7; Coach a little at a time	8
B8	B8: Offer alternative input modalities	5
B9	B9: Yes/no queries can be very robust	5
B10	B10: Carefully select the appropriate persona	1

considerations than something that is being operated through a keyboard. . . because a keyboard doesn't listen".

3.4.2 Stretching GUI Heuristics. Experts often assigned a GUI heuristic to a VUI issue even if they were not sure that it fit. Experts would use words like "maybe", "I think", or "I guess" when mapping GUI heuristics to VUI issues: "This whole stop listening one, maybe user control freedom, then, maybe." (P4). This occurred often with G2, as the devices were conversational in nature: "I'm not used to the need to trigger it. Maybe that's just a little bit of G2 going on, the real-world thing, but it just felt a little stilted." (P3).

Experts would often explain their reasoning when they were unsure of whether a GUI heuristic fit. Their rationale would stretch or generalize the interpretation of the heuristic in question. This occurred with G7, as what it means for a VUI to be "flexible" & "efficient" was interpreted differently: "I think to adapt requests. I guess flexibility and efficiency, but it's only a result of the fact that it already didn't get what I asked" (P7). This also happened for G2: ". . .It's only mapping between the system and the real world in that it needs contextual knowledge from the world in order to understand it. . . But I want to leave that as a G2." (P5)

3.4.3 Reframing GUI Heuristics. Experts often expressed that GUI heuristics may need to be updated or rephrased for VUIs: "I would want to reframe or rephrase some of the G ones . . .. I could relate

them but it's a bit of a force with a couple of them" (P3). It was noted that some VUI heuristics sounded like extended versions of GUI heuristics: "this is just G6, but with a twist for audio." (P5). P2 expressed a similar sentiment: "So it will be G1 plus plus, right? It's improved version of G1." As noted earlier, many GUI heuristics that overlapped with VUI heuristics often had very similar wordings, which supports this observation.

3.4.4 Process of Mapping GUI Heuristics to VUI Usability Issues. Some experts found the process of mapping GUI heuristics onto VUIs difficult. Experts found that it was often hard to adjust to applying GUI heuristics to VUI issues. As P1 noted: ". . . these heuristics were developed with certain devices in mind, they may not map as easily onto these new devices". P4 also expressed difficulty in reconciling the visual-based aspects of GUI heuristics with audio-based VUI interactions: "This talks about visual, but I think it should be audible. I mean, yeah, it's not providing the system status audibly until it's already made. No, this is hard."

3.5 Revisiting Heuristics Within the GUI to VUI Transition

Finally, we describe the process of brainstorming potentially new VUI heuristics with our experts. In Phase 3 (Brainstorming Potential New VUI Heuristics), GUI experts were asked to propose potential

"new heuristics" to solve issues that did not fit with any of the existing GUI heuristics. These new heuristics were not meant to stand as a formal proposal. Instead, they were meant to engage GUI Designers in exploring solutions for VUI usability issues that did not align directly with usability issues they may have previously experienced with GUIs in their careers (RQ3). Experts proposed heuristics that fit into three main themes: 1) Adaptation/Personalization, 2) Matching Conversational Constructs, and 3) Interacting with VUIs.

While these are themes that supposedly do not connect to existing GUI heuristics (as they were brainstormed from VUI issues that did not violate existing GUI heuristics), we found that they are still in fact grounded in GUI-design concepts, based on how experts interacted with VUI issues and GUI heuristics in earlier phases. For example, "Match Conversational Constructs" is a theme that is connected to G2 – Mapping Between System and Real World, and which appeared both when experts assigned GUI heuristics to VUI usability issues, and when experts identified overlaps between GUI and VUI heuristics. Many proposed heuristics in "Interaction with VUIs" are associated with interface output & feedback, which experts previously connected with G1 – Visibility/Feedback of System Status. This activity illustrates how, even with novel issues which our experts were not very familiar with, their previous GUI design training and experience influenced the solutions and heuristics they proposed.

This also shows that such participatory approaches that engage expert GUI designers may help re-purpose, re-frame, evolve, and adapt existing GUI heuristics to the VUI space. This contrasts with the challenges our expert participants had when handling new "purpose-built" VUI heuristics. Further, this suggests that a top-down "imposition" of new design principles may not be desirable, and in fact, a grassroots approach to co-creating new principles by anchoring them in familiar concepts may result in principles that help GUI designers transition to the VUI space.

4 DISCUSSION

Grounded in the analysis of data from our expert workshops, we return to answer the research questions we formulated at the beginning of this paper (Section 1.1).

RQ1. Given that most designers have currently been trained in GUI design, how do they interact with and handle VUI usability issues, given their existing GUI training and experience with existing tools (e.g. GUI heuristics)?

In general, we found that experts applied many GUI heuristics to the VUI usability issues that were brainstormed - as many of the VUI issues are similar to ones found in badly-designed GUIs. While experts did encounter difficulties in applying GUI heuristics to the VUI application space, they were still able to identify at least one VUI issue that violates a GUI heuristic for almost every VUI usability issue that was brainstormed. Since GUI experts were also involved in the brainstorming of the VUI issues themselves, it is possible that experts' previous GUI training also affected their identification of issues, and therefore found that they could apply many existing GUI heuristics to them. While this is so, all the issues that were brainstormed were issues that have been identified in previous literature about VUI usability [3, 12, 18, 26, 36]. Particularly, experts commented that a reframing or adjustment of a GUI heuristic could help make it applicable to the VUI design space.

This suggests that while GUI heuristics cannot be directly mapped to the VUIs, existing GUI heuristics can be adjusted to deal with the challenges that are specific to the new paradigm of voice interaction. The fact that experts were still able to apply the GUI heuristics show that they can still be useful as a high-level groundwork for developing VUI heuristics. Even during our workshops, usability experts incorporated their conceptual knowledge of GUI heuristics into the VUI heuristics that they proposed. Taking advantage of this previous experience and training can provide a pathway to transition more easily into VUI design. In particular, reframing GUI heuristics can help pave a pathway for an easy transition from GUI to VUI design.

RQ2. What barriers do GUI-trained designers face when trying to interact with, understand, and apply existing heuristics (both GUI and VUI) to VUI usability issues?

As mentioned earlier, experts did encounter some difficulties in applying GUI heuristics to the VUI application space, though they were still able to apply them to almost every VUI usability issue that they brainstormed. GUI experts primarily struggled to frame and interpret proposed VUI heuristics in literature, noting factors such as unfamiliar terminology or lack of clarity around how to apply the VUI heuristics to VUI usability issues. We observed a disconnect between their current understanding and familiarity of design heuristics, and how existing VUI heuristics were proposed. This caused participants to have some difficulty assessing the applicability of proposed VUI heuristics. This suggests that providing GUI-trained designers with novel tools (such as heuristics) may represent an unnecessary hurdle in both conceptualizing and applying them to VUI issues. This is congruent with previous research [43].

RQ3. What is the "path of least resistance" to helping designers to transition into the VUI design space, and how should we create tools to aid these designers?

One important observation was that, among both Wei/Landay's [43] and Suhm's [40] VUI heuristics, those that overlapped the most with GUI heuristics were the easiest to understand. These happened to be the same GUI heuristics that were easiest to apply to VUI usability issues in the earlier phases (G1, G2, G3, G7, G9). We found here that usability experts use and apply their familiarity of GUI heuristics to VUI design. This suggests that developing a common language, anchoring, or transition mapping between GUI heuristics and VUI heuristics may be the key to aiding designers to build better VUIs. This then further supports the need to base new VUI heuristics in familiar GUI heuristics. It is thus possible that GUI heuristics could be extended or reengineered to capture VUI usability issues. Furthermore, the observed challenges (such as stretching existing GUI heuristics) can guide us in identifying the kinds of modifications that are needed in order to adapt them to a VUI design space. Developing VUI heuristics in this way can provide a pathway of transition into VUI design that has a lower learning curve and is based in existing established design concepts.

5 TOWARDS USABILITY HEURISTICS FOR VOICE

As we have discussed in this paper, since current usability experts are trained primarily in GUI design principles and methods, it is

important to provide a "path of least resistance" to aid in transitioning from GUI to VUI design. Even new designers that are entering industry need to be trained in appropriate tools and methods - and often such training is provided by experts or educators who themselves may be more familiar with GUIs [22].

It is likely that the maturing of VUI design space will lead to opportunities for the next generation of designers to receive VUI-focused training. However, currently there is limited evidence that such opportunities will materialize soon enough [21] without a pathway for designers to make this transition. We must be able to provide appropriate tools to aid them as they transition from GUI to VUI design. Facilitating such transitions by creating heuristics that are anchored in prior experience is also likely to benefit new designers (as it is reasonable to assume they may have to handle both GUI and VUI designs in the future, or hybrid, multimodal interfaces).

One main observation was that experts fell back to their familiarity with GUI heuristics, which led to them reframing GUI heuristics to apply them to VUI usability issues. This leads to strong considerations when moving forward and can lead to solutions for developing VUI-based heuristics. Based on our observations and challenges identified, we suggest three considerations for designing VUI heuristics.

Reframe Existing GUI Heuristics for Voice. Usability experts were found to stretch definitions of GUI heuristics to apply them to VUI usability issues and VUI heuristics. They also commented on how many existing GUI heuristics can be reframed for voice. As well, experts found VUI heuristics more difficult to understand and use and found ones that shared common language with GUI heuristics the easiest to apply. Therefore, it would be useful to start with GUI heuristics and re-frame/adapt them for voice-dedicated devices, considering VUI-specific requirements.

Use Familiar Concepts and Wording/Terminology. We found that usability experts used concepts and terminology that they were familiar with from GUI heuristics (such as "interface feedback" and "correcting errors"). Therefore, maintaining that common language when designing VUI heuristics is important, especially for GUI designers who may be transitioning to the field of voice.

Follow Heuristic Language Format. Usability experts have existing interpretations of how a heuristic should be phrased and what it requires (such as a description and examples), which are grounded in existing GUI heuristics. Our workshops revealed the importance of providing information about the scope and applicability of new heuristics, in a language that design experts can understand and relate to. This should be grounded in their expectations about how that heuristic guideline is assisting them in their design activities. This will facilitate the adoption of such new heuristics by usability experts.

6 CONCLUSION

In this paper, through a set of seven participatory workshops with usability experts, we show that a pathway for helping GUI designers transition into VUI design can be grounded in developing VUI heuristics that are based on GUI design knowledge, concepts, and practices. This requires a bottom-up process that involves design experts from the beginning and throughout the entire process of VUI-heuristic development. Our results suggest that adapting or reframing existing GUI heuristics for a VUI paradigm can help develop a "path of least resistance" for existing (primarily GUI-trained) designers to smoothly transition into the VUI design space. Using participatory approaches such as the ones used in our workshops, that directly engage expert GUI designers, may help re-purpose, reframe, evolve, and adapt existing GUI heuristics to the VUI space. A top-down approach of developing new design heuristics (a method that has been followed by previous attempts to develop VUI heuristics [40, 43]) may need to be discarded in favour of a grassroots approach to co-creating new heuristics by anchoring them in familiar GUI concepts. In this way, we can identify a common language or anchoring between GUI heuristics and potential new VUI heuristics, which may help develop this transitional pathway to VUI design.

As designers transition from graphical-based interfaces into the new field of voice, grounding new VUI heuristics in existing GUI heuristics and developing a common language between the two design paradigms will allow new VUI designers to make the transition smoothly, and to facilitate the improvement of voice interface design. Based on the findings of this paper, we believe that future work should focus on gathering both GUI-trained designers and VUI-trained designers together to explore how to create a concrete set of heuristics that make particular use of GUI designer knowledge and experience. We hope that this work can be a launching pad for new methods in developing heuristics that may lead to better and more adoptable VUI design.

REFERENCES

[1] Matthew P. Aylett, Per Ola Kristensson, Steve Whittaker, and Yolanda Vazquez-Alvarez. 2014. None of a CHInd. Proc. of CHI EA '14: 749–760. https://doi.org/10.1145/2559206.2578868

[2] Bruce Balentine and David P. Morgan. 1999. How to Build a Speech Recognition Application: A Style Guide for Telephony Dialogues.

[3] Dan Bohus and Alexander I Rudnicky. 2005. Sorry, I Didn't Catch That! – An Investigation of Non-understanding Errors and Recovery Strategies. 12.

[4] Jose A Borges, Israel Morales, and Nkstor J Rodriguez. 1996. Guidelines for Designing Usable World Wide Web Pages. Retrieved from http://delivery.acm.org/10.1145/260000/257320/p277-borges.pdf?ip=174.112.248.232&id=257320&acc=ACTIVE SERVICE&key=FD0067F557510FFB.148C9AE997532579.2370BB3FAC5962EF.4D4702B0C3E38B35&__acm__=1537318774_5ad8eb060e1eefa36cb28cf6616570b5

[5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. Qualitative Research in Psychology 3, 2: 77–101. https://doi.org/10.1191/1478088706qp063oa

[6] Kim Brunhuber and 2018. 2018. The hottest thing in the world of technology: your voice | CBC News. CBC. Retrieved July 9, 2019 from https://www.cbc.ca/news/technology/brunhuber-ces-voice-activated-1.4483912

[7] Raluca Budiu and Page Laubheimer. Intelligent Assistants Have Poor Usability: A User Study of Alexa, Google Assistant, and Siri. Nielsen Norman Group. Retrieved October 26, 2018 from https://www.nngroup.com/articles/intelligent-assistant-usability/

[8] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16), 981–992. https://doi.org/10.1145/2858036.2858498

[9] Leigh Clark, Philip Doyle, Diego Garaialde, Emer Gilmartin, Stephan Schlögl, Jens Edlund, Matthew Aylett, João Cabral, Cosmin Munteanu, Justin Edwards, and Benjamin R Cowan. 2019. The State of Speech in HCI: Trends, Themes and Challenges. Interacting with Computers 31, 4: 349–371. https://doi.org/10.1093/iwc/iwz016

[10] Leigh Clark, Nadia Pantidi, Orla Cooney, Philip Doyle, Diego Garaialde, Justin Edwards, Brendan Spillane, Emer Gilmartin, Christine Murad, Cosmin Munteanu, Vincent Wade, and Benjamin R. Cowan. 2019. What Makes a Good Conversation? Challenges in Designing Truly Conversational Agents. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), 1–12. https://doi.org/10.1145/3290605.3300705

[11] Eric Corbett and Astrid Weber. 2016. What Can I Say? Addressing User Experience Challenges of a Mobile Voice User Interface for Accessibility. In Proceedings of

the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16), 72–82. https://doi.org/10.1145/2935334. 2935386

[12] Benjamin R Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. "What Can I Help You With?": Infrequent Users' Experiences of Intelligent Personal Assistants. In Proc. of MobileHCI '17, 1–12. https://doi.org/10.1145/3098279.3098539

[13] Anushay Furqan, Chelsea Myers, and Jichen Zhu. 2017. Learnability through Adaptive Discovery Tools in Voice User Interfaces. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17), 1617–1623. https://doi.org/10.1145/3027063.3053166

[14] Alex Graves, Abdel-rahman Mohamed, and Geoffrey Hinton. 2013. Speech recognition with deep recurrent neural networks. In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 6645–6649. https://doi.org/10. 1109/ICASSP.2013.6638947

[15] Rodolfo Inostroza, Cristian Rusu, Silvana Roncagliolo, Cristhy Jimenez, and Virginica Rusu. 2012. Usability Heuristics for Touchscreen-based Mobile Devices. In 2012 Ninth International Conference on Information Technology - New Generations, 662–667. https://doi.org/10.1109/ITNG.2012.134

[16] James R. Lewis. 2010. Practical Speech User Interface Design. CRC Press, Inc. Retrieved September 20, 2019 from http://dl.acm.org/citation.cfm?id=1965361

[17] Irene Lopatovska, Katrina Rink, Ian Knight, Kieran Raines, Kevin Cosenza, Harriet Williams, David Hirsch, Qi Li, and Adrianna Martinez. 2018. Talk to me: Exploring user interactions with the Amazon Alexa. https://doi.org/10.1177/0961000618759414

[18] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16, 5286–5297. https://doi.org/10.1145/2858036.2858288

[19] Jennifer Mankoff, Anind K Dey, Gary Hsieh, Julie Kientz, Scott Lederer, and Morgan Ames. 2003. Heuristic Evaluation of Ambient Displays. NEW HORIZONS, 5: 8.

[20] Christine Murad and Cosmin Munteanu. 2019. "I Don't Know What You're Talking About, HALexa": The Case for Voice User Interface Guidelines. In Proc. of CUI '19 (CUI '19), 9:1–9:3. https://doi.org/10.1145/3342775.3342795

[21] Christine Murad and Cosmin Munteanu. 2019. Teaching for Voice: The State of VUI Design in HCI Education. Proceedings of EduCHI 2019 Symposium.

[22] Christine Murad and Cosmin Munteanu. 2020. Designing Voice Interfaces: Back to the (Curriculum) Basics. In Proc. of CHI '20 (CHI '20).

[23] Christine Murad and Cosmin Munteanu. 2020. Alexa, How do I Build a VUI Curriculum? In Proc. of CUI 2020.

[24] Christine Murad, Cosmin Munteanu, Leigh Clark, and Benjamin R. Cowan. 2018. Design guidelines for hands-free speech interaction. In Proc. of MobileHCI '18, 269–276. https://doi.org/10.1145/3236112.3236149

[25] Christine Murad, Cosmin Munteanu, Benjamin R. Cowan, and Leigh Clark. 2019. Revolution or Evolution? Speech Interaction and HCI Design Guidelines. IEEE Pervasive Computing 18, 2: 33–45. https://doi.org/10.1109/MPRV.2019.2906991

[26] Chelsea M. Myers, Anushay Furqan, and Jichen Zhu. 2019. The Impact of User Characteristics and Preferences on Performance with an Unfamiliar Voice User Interface. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), 1–9. https://doi.org/10.1145/3290605.3300277

[27] Jakob Nielsen. 1994. Enhancing the explanatory power of usability heuristics. Proc. of CHI '94: 152–158. https://doi.org/10.1145/191666.191729

[28] Jakob Nielsen. 1994. Usability Inspection Methods. In Conference Companion on Human Factors in Computing Systems (CHI '94), 413–414. https://doi.org/10. 1145/259963.260531

[29] Jakob Nielsen. 2003. Voice Interfaces: Assessing the Potential. Nielsen Norman Group. Retrieved November 20, 2018 from https://www.nngroup.com/articles/voice-interfaces-assessing-the-potential/

[30] Donald Norman. 1988. The Design of Everyday Things. Doubled Currency.

[31] David Pinelle, Nelson Wong, and Tadeusz Stach. 2008. Heuristic evaluation for games: usability principles for video game design. Proceedings of SIGCHI Conference on Human Factors in Computing Systems: 1453–1462. https://doi.org/10. 1145/1357054.1357282

[32] Martin Porcheron, Joel E. Fischer, Stuart Reeves, and Sarah Sharples. 2018. Voice Interfaces in Everyday Life. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), 1–12. https://doi.org/10.1145/3173574. 3174214

[33] Bob Regan. 2004. Accessibility and Design: A Failure of the Imagination. In Proceedings of the 2004 International Cross-disciplinary Workshop on Web Accessibility (W4A) (W4A '04), 29–37. https://doi.org/10.1145/990657.990663

[34] Alexander I Rudnicky. 1996. Speech Interface Guidelines. Retrieved from www.speech.cs.cmu.eduu

[35] J Sherwani, Dong Yu, and Tim Paek. 2007. Voicepedia: towards speech-based access to unstructured information. Interspeech: 2–5.

[36] Ben Shneiderman. 2000. The limits of speech recognition. Communications of the ACM 43, 9: 63–65. https://doi.org/10.1145/348941.348990

[37] Ben Shneiderman and Catherine Plaisant. 2010. Designing the User Interface: Strategies for Effective Human-Computer Interaction. https://doi.org/10.1016/0166-3615(93)90066-A

[38] Anjeli Singh, Andrea Johnson, Hanan Alnizami, and Juan E Gilbert. 2011. The Potential Benefits of Multi_Modal Social Interaction on the Web for Senior Users. J. Comput. Sci. Coll. 27, 2: 135–141.

[39] Joanna Smith and Jill Firth. 2011. Qualitative data analysis: the framework approach. Nurse Researcher 18, 2: 52–62. https://doi.org/10.7748/nr2011.01.18.2.52. c8284

[40] Bernhard Suhm. 2003. Towards Best Practices for Speech User Interface Design. In Proc. of EuroSpeech '03, 2217–2220.

[41] Alistair Sutcliffe and Brian Gault. 2004. Heuristic evaluation of virtual reality applications. Interacting with Computers 16, 4: 831–849. https://doi.org/10.1016/j.intcom.2004.05.001

[42] Janice Y Tsai, Jofish Kaye, Tawfiq Ammari, and Abraham Wallin. 2018. Alexa, play some music: Categorization of Alexa Commands. Workshop on Voice-based Conversational UX Studies and Design.

[43] Zhuxiaona Wei and James Landay. 2018. Evaluating Speech-Based Smart Devices Using New Usability Heuristics. IEEE Pervasive Computing 17, 2: 84–96. https://doi.org/10.1109/MPRV.2018.022511249

[44] Kathryn Whitenton. 2016. Voice Interaction UX: Brave New World...Same Old Story. Nielsen Norman Group. Retrieved September 21, 2018 from https://www.nngroup.com/articles/voice-interaction-ux/

[45] Maria Wolters, Kallirroi Georgila, Johanna D. Moore, Robert H. Logie, Sarah E. MacPherson, and Matthew Watson. 2009. Reducing working memory load in spoken dialogue systems. Interacting with Computers 21, 4: 276–287. https://doi.org/10.1016/j.intcom.2009.05.009

[46] Nicole Yankelovich, Gina-Anne Levow, and Matt Marx. 1995. Designing SpeechActs: Issues in Speech User Interfaces. In Proc. of CHI '95, 369–376. https://doi.org/10.1145/223904.223952

[47] 2005. Analysis of the ENABLED Web Developer Survey. Retrieved from http://www.enabledweb.org/public_results/survey_results/analysis.html

[48] Reading 20: Heuristic Evaluation. Retrieved September 13, 2019 from http://web.mit.edu/6.813/www/sp16/classes/20-heuristic-evaluation/