

Unpacking Practitioners' Attitudes Towards Codifications of Design Knowledge for Voice User Interfaces

KRISHIKA HARESH KHEMANI

Horizon Digital Economy Research, University of Nottingham, krishika.khemani@nottingham.ac.uk

STUART REEVES

Mixed Reality Lab, School of Computer Science, University of Nottingham, stuart.reeves@nottingham.ac.uk

Recent HCI research has sought to develop guidelines—'heuristics', 'best practices', 'principles' and so on—for voice user interfaces (VUI) to aid both practitioners and researchers in improving the quality of VUI-based design. However, limited research is available on how such design knowledge is conceptualised and used by industry practitioners. We present a small interview-based study conducted with 9 experienced VUI industry practitioners. Their concerns range from terminological challenges associated with VUI design knowledge, the role of codifications of such knowledge like design guidelines alongside their practical design work, through to their views on the value of 'harmonisation' of VUI design knowledge. Given the complex—albeit preliminary—picture that emerges, we argue for HCI's deeper consideration of how design knowledge meshes with the contingencies of practice, so that VUI design knowledge—such as design guidelines developed in HCI—delivers the most potential value for industry practice.

CCS CONCEPTS • Human-centered computing • Human computer interaction • Empirical studies in HCI

Additional Keywords and Phrases: Design guidelines, voice user interfaces, industry practitioners

ACM Reference Format:

Krishika Haresh Khemani and Stuart Reeves. 2022. Unpacking Practitioners' Attitudes Towards Codifications of Design Knowledge for Voice User Interfaces. In CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3491102.3517623>.

1 INTRODUCTION

A key part of HCI's programme has been in the development of tools to generalise and codify (i.e., document or formalise) design knowledge and provide reliable resources to support design work [36]. This includes—variously—heuristics, guidelines, frameworks, best practices, principles, and so on (see [36]). Recent research interest in conversational user interfaces has (perhaps somewhat naturally) also led to the development of design tools like guidelines for voice user interfaces (VUI) [9, 13, 14, 15, 37, 41, 42]. And like more general HCI design knowledge that is presented as guidelines etc., codified VUI design knowledge also tend to be presented as application-oriented and framed for audiences of both practitioners (e.g., VUI designers in industry) and researchers.

At the same time, industry UX and design practice in VUI—strongly oriented to specific platforms like Alexa, Siri or Google Assistant which tend to define and delimit what is meant by 'VUI'—is expanding, and maturing. Like academic HCI, it too is seeking ways to preserve, codify, and share generalised design knowledge in the form of practitioner-originated guidelines, best practices, principles and so on (e.g., see Deibel and Evanhoe's recent book [21]). This move reflects practitioners' prior familiarity with codified design knowledge and experience of their value within design work for more traditional interfaces—desktop, web, mobile—for which significant efforts have been made in, e.g., influential style guides [7, 24].

However, it is unclear how these various emerging forms of codified design knowledge for voice user interfaces, particularly from research, will meet the needs of practice [5] given that we do not know much about how VUI practitioners conceptualise them or indeed actually use them in the first place. To begin sketching out the dimensions of this problem space, we conducted a survey and interviewed a number of VUI design practitioners to discuss their work with respect to what *they* consider to be relevant 'design guidelines', and uncover some initial concerns with which to orient future HCI

research on VUI practitioners' design practices. We focussed on 'design guidelines' as a way into conversations with them rather than a more abstract notion like 'codified design knowledge', which we use in this paper to refer to such resources.

In this paper, at first, we very briefly examine how codified design knowledge has been discussed within relevant research communities (e.g., HCI, conversational user interface (CUI) research, etc.). We also look at the role of this in GUI, interactive voice response (IVR) and VUI design and its impact on UX. Then we discuss platform-specific VUI 'design guidelines' published by industry practitioners, alongside contributions made by academic research. Turning to our empirical studies, we describe the approach we used to explore how practitioners treat such codified design knowledge, then present our findings structured around the following: variation in use of terminology, integration of guidelines and considerations in practitioners' design process, and lastly harmonisation of design guidelines. By 'harmonisation' we mean the notion that various codifications may be brought together somehow for varied potential purposes: e.g., increasing consistency, establishing standards, or generating community agreement on design. We conclude with a discussion on the salient issues arising from our findings with respect to HCI research and implications of this for potential areas of further exploration.

2 DESIGN PRINCIPLES, GUIDELINES, HEURISTICS AND BEST PRACTICES

We first need to lay some groundwork for our interview study's findings, firstly by scoping the domains of conversational UI, voice UI, and so on. Then we examine the language of codifications of design knowledge, their expression within prior HCI and interaction design, and review emerging discussions on the development of design knowledge codifications for VUI design research and practice.

2.1 Scoping conversational user interfaces (CUI) and voice interfaces

Here we want to provide some context for our interest in voice user interfaces specifically, and very briefly tease out the complexities of this space, particularly given that some of our interviewees had experience in, say, IVRs or chatbots (which were not the topic of our investigation but nevertheless related). The notion of interfaces that engage in 'dialogue' or 'conversation' has of course had a long history and goes by various names—conversational interfaces, speech user interfaces, IVRs, voice assistants, smart speakers, etc. Ultimately, however, the interactional modes of such interfaces tend to either be text-based ('chatbots') or voice-based. CUIs can be defined as interfaces where the "user interaction is conducted primarily through words: typed or spoken." [41]. As implied by the name, voice user interfaces use 'voice' as a mode of interaction [32]. Interactive voice response (IVR) systems and voice assistants (e.g., Siri, Alexa, etc.) are the two main types of voice user interfaces. IVRs—developed since the 1970s—are mainly integrated with telephone systems and use a menu-based approach to carry out a limited number of tasks for the user, often integrating speech recognition since the 1990s. Voice assistants on the other hand have been designed with speech recognition as a given, and offer a much wider range of tasks.

Ultimately, however, text-based and voice-based conversational interfaces have a similar architecture for handling input and output. The main difference is that voice-based interfaces employ Automatic Speech Recognition (ASR) and Speech-to-Text (STT) components [39], but both interfaces are often implemented using similar natural language processing engines and back-end dialogue management systems to process data. Our paper therefore, while focusing on modern, voice assistant platforms that are popular with UX practitioners, may have some implications for conversational UI in general.

2.2 Codifications of design knowledge: Design guidelines and affiliated concepts

Prior work has sought to clarify and define what is meant by (e.g.) design principles, design guidelines and heuristics [29, 30, 33, 35]. HCI researchers and practitioners suggest that guidelines and other forms of HCI knowledge like principles or heuristics are derived from empirical research and practical experiences [29, 30, 35]. One view distinguishes *design principles* as "laws" that drive the design process forward to produce effective solutions and *design guidelines* as "context-dependent directives" [35] or recommendations on how design principles should be applied [30, 33].

The term *heuristics* is also often used, with some arguing that although they are similar in nature to design guidelines as context-dependent rules, they are usually less prescriptive in nature than design guidelines and can also be descriptive [35]. In this view, heuristics are based on “intuition, tacit knowledge, or experiential understanding” [35]. Others argue that heuristics may be thought instead as “rules of thumb” or (circularly) “principles” [30, 33] because they are thought to be not as specific as design guidelines. Another term in use is *best practices*. Suhm argues that best practices are simply design guidelines which produce a “positive impact” [10] which can be evaluated using an appropriate metric.

As may be clear at this point, the discourse of HCI ‘guidelines’ and affiliated concepts—as codified design knowledge—is sometimes contradictory and ambiguous. We will see later how this in some senses mirrors the picture painted by VUI design practitioners.

2.3 From classic HCI guidelines to VUI guidelines

HCI research has often suggested (e.g., within standard textbooks [45]) that the various codifications of design knowledge as described above are crucial to successful implementation of a design solution. Designers and developers reportedly use them as tools to guide implementation of various interfaces. Therefore, we refer to literature on human-centered and UX design to ground our understanding of how they are embedded into GUI, IVR and VUI design. Further, we also must look to burgeoning communities dedicated to conversational user interface research specifically (as represented by the CUI series of conferences [1, 2])

In HCI, it is well-known that guidelines, principles, heuristics, etc. are intended to ensure critical factors—usability, cognitive load, usefulness, desirability and accessibility which significantly influence user experience [31]. Implementation of guidelines or heuristics e.g., Jakob Nielsen’s 10 Usability Heuristics for user interface design, account for these factors. First documented in 1990 by Nielsen and Molich, such classic guidelines have perhaps naturally found themselves repurposed for VUI design [19]. However, questions have been raised about the suitability of applying these to voice [40].

IVR systems in the past have also had their own tradition of design guidelines [10, 11]. Killam and Autry [11] note that despite IVRs increased popularity at the time, designers struggled with the design and evaluation of IVRs due to lack of publicly available standardised design guidelines [11, 18]. The design guidelines which were available were proprietary and companies would not disclose them [18]. Consequently, human factors practitioners started developing their own design guidelines to support designers who wanted to perform evaluation of IVRs [10, 11]. Some of these IVR design guidelines have laid the foundation for codifications of VUI design knowledge developed by academic researchers [46].

Wei and Landay [46] developed what they call “usability heuristics” for VUI design after reviewing design guidelines for GUI and telephony systems, and design guides by Amazon and Google [3, 23]. Some of the heuristics they developed revolve around conversation design, error handling guidance for users (e.g., discoverability and what a voice assistant can help the user accomplish), and length of prompts or speech. Murad et al. [14] used a different approach to compile what they term as “design guidelines” for voice interfaces. They first reviewed literature to identify VUI usability issues, and mapped them to existing codifications of GUI design knowledge associated with privacy and social context. Next, they developed a mapping system between GUI design knowledge and a combination of work by aforementioned Wei and Landay [46] and Suhm [9]. This mapping system helped reveal the overlapping notions between codifications of GUI and VUI design knowledge which, it is argued, would support the shift from GUI to VUI, for designers [14].

2.3.1 Industry and practitioner initiatives for VUI design guidelines

As a result of increasing popularity of voice interfaces, voice industry has generated its own approaches to codifying design knowledge in terms of principles, guidelines and heuristics that tend to be developed in parallel to more academic approaches outlined above. For instance, practitioners have published books [12, 38] and blogposts [8, 19, 25] which reveal the VUI design process and provide ‘formal’ articulations of VUI design knowledge. For instance, Pearl shares what she calls “VUI design principles” centered around topics like conversation design, user expectations, confirmations, error handling, etc., in [12]. Additionally, instead of listing dos and don’ts, she shares examples of human-agent interactions to demonstrate how to apply these principles and to distinguish between good design and bad design. Similarly, Cohen et al. [38] present 6 “core VUI design principles” derived from analysis of lab and field studies. These principles reflect an understanding of human factors, linguistics, speech technologies and voice interface design e.g., “minimize the cognitive load”, “accommodate

conversational expectations”, “gracefully recover from errors” and “ensure high accuracy” [38]. Unlike Pearl, the Cohen et al. speak of VUIs in a broad sense—they make no distinction between IVR’s and voice assistants; hence, the principles are more generic than those provided by Pearl [12].

Therefore, it can be argued that multiple efforts are being made by companies, practitioners, and academic researchers to develop what they variously term as design principles, guidelines or heuristics—i.e., practitioner-oriented codifications of design knowledge—for voice interfaces. Some of these codifications overlap or are conceptually similar. Indeed, there are initiatives which attempt to bring together a common, shared set of guidelines for VUIs. For example, the Multi-Agent Design Guide [4] published by Amazon to facilitate support for multiple voice services offered by different companies (e.g., Accenture, BBC, etc.) on the same device. We take this as a key example of what we mean by ‘harmonisation’—a concern addressed during our interviews with practitioners.

3 METHODOLOGY

We were interested in learning how VUI experts implemented codifications of design knowledge, such as design guidelines, while creating VUIs. Our conversations tended to centre on currently popular voice assistant platforms like Amazon Echo, Siri, or Google Home—platforms our participants were experienced with. Furthermore, we were intrigued by what practitioners’ opinions were on harmonising design guidelines given that this impulse seems to have been a driving force for prior UI codifications. So, we did a preliminary survey to recruit participants and gather background information for our main research—an interview study.

Prior to the study, we submitted an ethics application to the Computer Science Research Ethics Committee at University of Nottingham. We also ensured data collection and storage of survey and interview data was GDPR compliant. Personal information such as names and email address were only collected in our data if participants expressed their interest to interview, otherwise respondents could share information anonymously within the survey. Given we were talking to industry practitioners, we emphasised the need to avoid disclosing any confidential information or information bound by non-disclosure agreements.

3.1 Participant recruitment through preliminary survey

We decided it was essential to recruit participants who had professional experience with building *voice*-based conversational interfaces, although this necessarily meant that many of them had broader ‘conversational design’ experience too (e.g., chatbots). To this end, we initially created a survey in part to better understand practitioners’ experiences designing and building voice assistants, but also to recruit interview participants for developing a more in-depth appreciation of this. We shared our survey on social media platforms such as Twitter and LinkedIn. With the help of VUI practitioner networks and members of the CUI research community on Twitter and other social media we were able to advertise our survey and obtained a total of 64 respondents.

The survey had 13 questions designed to elicit detail about practitioners’ experiences. Firstly, we asked participants if they had any professional experience working with VUIs. By asking about their ‘professional’ experience, we wanted to make a distinction between work that is conducted in a more ‘professional’ capacity (whether in academia or industry) versus work done as a part of passion projects or coursework. The survey terminated for those who responded with ‘no’ to this question. Secondly, we wanted to know how many years of experience the participant had. These questions were important to help funnel participants who had significant experience working with voice interfaces towards the interview stage. Next, we asked participants if they had used codifications of design knowledge like design guidelines, heuristics, and so on in the course of this work, such as those published by Amazon, Fjord, Google [3, 22, 23], etc. We also asked about technologies they worked with (e.g., Alexa, Google Assistant). For these two questions, we enabled multi-selection to allow participants to choose more than one option for the answer and also added ‘Other’ as an option to encourage them to tell us more about any other design guidelines or platforms they may have built voice applications with. Additionally, we added open-ended questions to the survey to learn about their perceptions on design guidelines for VUIs. We collected the contact information of those who were interested to schedule an interview.

Although we had 64 respondents who attempted the survey, only 42 (i.e., 66%) of those respondents reported significant professional experience—i.e., ‘real world’ familiarity with designing VUIs for actual deployment. As a result, only those 42 survey respondents could answer all the questions to complete the survey. Initially, 16 out of 42 survey respondents (i.e., 38% of our sample) expressed their interest to interview, while 9 respondents actually agreed to interview. Filtering respondents who had significant experience largely played a role in the small sample size for our interview study. In addition, the duration of the study i.e., 3 months, and voice design as a field being seen as quite new (in spite of historic developments in, say, IVR) may have been other contributing factors.

Table 1 below summarises our interviewees’ experience with VUIs and their background. It shows all interviewees had at least 2 years of experience in the voice industry. The 2 interviewees with over 10 years of experience also had experience with IVR systems.

Table 1: Summary of interviewees’ professional experience with VUIs

Participant ID	Years of Experience	Education	Role	Area of expertise	Platforms
P1	5-10	Postgraduate degree in Creative Writing	Conversation Designer	Conversation design, UX and UX writing	Alexa, Google Assistant
P2	2-3	Undergraduate degree in Creative Writing	UX Writer	Design VUI documentation, content strategy	Alexa
P3	3-5	Undergraduate degree in Computer Science	Developer	Conversational interface design and development	Google Assistant, Google Glass, Alexa
P4	10+	Undergraduate degree in Economics	Managing Director	VUI design, development and testing, IVR systems	Nuance
P5	3-5	Undergraduate degree in Marketing	Director	Product design (created prototyping tools for voice interfaces)	Alexa, Siri, Google Assistant
P6	1-2	Postgraduate degree in Cognitive Linguistics	Conversation Designer	Conversation design, Linguistics	Alexa, Jovo Framework, Google Assistant
P7	2-3	Undergraduate degree in English Literature	Conversation Designer	VUI design, Language	Alexa, Google Assistant
P8	10+	Undergraduate degree in English Language	Entrepreneur	VUI and conversation design, IVR systems	IVR platforms, Alexa, Google Assistant, Cortana
P9	2-3	Postgraduate degree in Speech Communication	VUI Architect	VUI architecture, conversation design, IVR	Hello Magenta, Google Assistant, Alexa

3.2 Interview study

We first describe how we used the findings from the survey to shape the interview study and formulate the questions we asked our interviewees. We then briefly describe the inductive and deductive approach used to perform thematic analysis on our interview data.

3.2.1 Interview study: Using preliminary survey responses to inform our interview questions

As stated previously, the preliminary survey was useful to gather background information needed for the interview study. Through the survey responses we learned about the platforms practitioners often worked with, their perceptions on codifications of design knowledge for VUIs, and limitations of existing VUI design knowledge.

From our survey responses, we discovered Alexa (90.5%) and Google Assistant (59.5%) platforms were more popular amongst the respondents. Similarly, our respondents were more familiar with codifications of design knowledge for these two platforms compared to other options listed in the survey: Alexa Design Guide, 83.3%, and Actions on Google Design Guide, 64.2% (percentages are calculated by considering the 42 respondents with ‘professional’ experience; as respondents were allowed to select multiple platforms and design guides, the percentages do not sum up to 100%). Upon learning this, we closely examined these design guides prior to the interviews to expand our understanding of VUI design for these

platforms. Some of our respondents also reported they worked with ‘other’ platforms such as Nuance VoCon², IVRs, Amazon Lex³, and Amelia IPSoft⁴. Several respondents have also created or worked with proprietary VUI design guidelines. Few of them indicated they had referred to Grice’s Maxim [27] or Pearl’s book [12].

The survey further helped us scope how practitioners respond to the notion of codifying design knowledge and thus informed how we conducted interviews. Most of the survey responses conveyed that design knowledge for voice interfaces are “necessary” or “critical”. Moreover, respondents deemed useful to “standardise” or create “consistent experiences”. Additionally, they suggested design guides “played an important role” in their learning process and are “helpful in everyday work”.

Some of our respondents reported problems with existing voice design guides. One shared that design guides are “lacking examples”, and there is “a lot of need” to make design guides “really applicable”. Another respondent complained design guides “oversimplify the process” and are “aimed at beginners”. Some criticised existing design guides as “mostly surface level”, “quite linear” and lacking “deep psycholinguistic foundations”. Other respondents were more conciliatory about this, expressing their sense that since VUI systems are “emerging tech”, so design knowledge for them is in “early stages” but is “evolving rapidly”. Making note of this while analysing our survey responses, we realised it might be important to explore these limitations of voice design guides further during the interviews.

Despite the current challenges with codifications of VUI design knowledge, most of our respondents seemed to feel positively about it. One of the respondents said “creation of design guidelines facilitates much-needed discussion within the VUI community”. Another respondent emphasised that “VUI design guidelines need to be established” and “continually updated [...] as voice assistants improve”. Therefore, our survey responses supported the notion that practitioner communities *could* benefit from HCI research that builds further codifications of voice design knowledge in some way (although we would suggest this is predicated upon HCI and CUI research developing deeper understandings of practitioners’ actual practices so as to coherently ground and situate design knowledge).

The survey identified for us specific areas to explore in more detail during the interviews. We organised our interview questions around the following key topics for discussion with participants: their professional experience; descriptions of their VUI design process, and how design guidelines fitted into that; how they performed evaluation of VUI designs; their views on design guidelines; their views on ideas about ‘unified’ or ‘standardised’ guidelines; descriptions of instances where design guidelines were limiting or problematic; and, areas of VUI design they thought design guidelines might address.

3.3 Interview study: Approach and analysis

We followed a semi-structured interview format mainly to delve deeper and better understand our interviewees’ perspectives on codifications of VUI design knowledge. We wanted to follow-up on what interviewees mentioned in the survey. And, if during the interviews they shared something that was unusual yet relevant, inquiring further to better understand the context.

We tried to be sensitive towards the terminology practitioners used to speak of design guidelines. Of course, a key part of our interviews revolved around understanding what ‘design guidelines’ meant for participants. In some cases, practitioners seemed to adopt a deliberate use of language, whereas if we noticed them using varying terminology, we asked them to explain what they meant by this (e.g., if they saw any difference in terms and therefore distinguished between codifications of design knowledge). This enabled us to unpack assumptions and support a shared understanding of the topic being discussed.

Once all interviews were completed, we cleaned up the transcripts by revisiting the interview recordings and rectifying any errors observed. This helped us familiarise ourselves with the data collected and prepare it for thematic analysis [34, 44]. Our deductive approach involved developing an initial set of codes based on questions asked during the interviews. The initial list of codes included ‘design process’, ‘VUI design evaluation’, ‘views on design guidelines’, ‘design guidelines

2 Nuance VoCon - https://www.nuance.com/content/dam/nuance/en_us/collateral/mobile/automotive/data-sheet/ds-vocon-hybrid-speech-recognition-data-sheet-en-us.pdf

3 Amazon Lex - <https://aws.amazon.com/lex/>

4 Amerlia IPSoft - <https://amelia.ai/>

limitations’, and ‘unification of design guidelines’. As we conducted semi-structured interviews, there was a possibility that several other themes would emerge. So, our inductive approach involved re-reading the interview data, and looking beyond the initial codes we developed. So, returning to our discussion of terminology above, for instance, where a few participants spoke about, say, ‘best practices’ and ‘heuristics’ in *addition* to ‘design guidelines’ while others used ‘best practices’ and ‘design guidelines’ interchangeably, this led us to develop a new theme around ‘inconsistent terminology’.

After following an iterative process of re-reading and coding, we eventually combined the codes under three main themes: terminology, VUI design processes and considerations, and harmonisation of design guidelines. These themes of course sit in a relationship to the kind of questions we asked, and it is clear that there is an interaction between the way our interviews were framed and the kinds of themes that emerged from the data—in that sense our analysis is limited by this fact. Further, our themes should be read as preliminary, as befits the limited scope of this initial study into industry practitioners’ orientations to VUI design.

4 UNDERSTANDING VUI PRACTITIONERS’ ORIENTATIONS TO CODIFICATIONS OF DESIGN KNOWLEDGE

Our analysis of the interviews led us focus on three core issues impacting practitioners’ relationship with codified design knowledge. Firstly, our interviews reveal a wide range of terminological issues, and what concepts and language design practitioners use to talk about VUI design. Secondly, we explore practitioners’ accounts of how design guidelines feature in VUI design processes. Finally, we discuss practitioners’ responses to familiar questions of ‘harmonisation’ of design guidelines as coherent, codified bodies of knowledge.

4.1 Trouble with terminology

One might perhaps naively think that VUI design practitioners have some kind of common language that enables them to talk about VUI design. To this end we asked practitioners to discuss with us the ways they talked about design guidelines (a term we introduced with reference to e.g., platform ‘guidelines’ but used as way of initiating discussion of codified design knowledge for VUIs in general). We found that practitioners firstly produced many different related terms beyond simply ‘guidelines’, including ‘heuristics’ and ‘best practices’, and secondly some made definite distinctions between such terms while others did not do so, using them seemingly interchangeably.

For those that made distinctions, some interviewees offered different definitions of best practices, design guidelines and heuristics. In one view, our interviewees explained that ‘best practice’ held connotations of there being only one way to go about something—that a ‘best practice’ tells you what interaction with VUIs *should* ‘look’ like to end users and the specifics of it make it restrictive in nature (P8). This is thus introducing a ‘moral’ flavour to notions of ‘best practice’ (e.g., as P8 describes below). On the other hand, interviewees suggested terms like ‘design guidelines’ could in contrast to ‘best practices’ be considered more like general attributes or characteristics one should try to accomplish in one’s VUI design, i.e., that ‘design guidelines’ express how a VUI should ‘act’ and ‘feel’ like:

“You’ll see things that are labelled on your best practices like um, ‘don’t give anybody more than three choices at its time.’ Whereas, guidelines are ‘make it clear by structuring information in a logical order and not presenting it too quickly’” (P8).

Another participant, P1, made a different distinction between heuristics and design guidelines. They described design guidelines are a “checklist of dos and don’ts” and heuristics as “a set of values”:

“Heuristics side is [...] these high-level things like ‘this system should be able to handle context’ which includes everything from pronoun context to, situational context” (P1).

Regardless of the specific terms, then, we can see a general point being made here by interviewees in terms of a tension between *guidelines-as-prescriptions* and *guidelines-as-values*. As we will see, there was a general reaction that prescription for VUI design is perhaps differently problematic than for its existing GUI equivalents due to the nature of the design material, i.e., language.

Some interviewees did not seem to differentiate as much between these terms. For example, we observed P7 using the terms 'design guidelines' and 'design practice' interchangeably and so we clarified if they were making a distinction. They replied "I don't know that I'm using it super precisely there, so I don't know that I'm using it differently".

Finally, some interviewees sat within some 'middle ground' between the two extremes described previously. For instance, P3 positioned VUI design guidelines as related to best practices in a more complex way:

"The earliest GUI design guidelines there were things like you know, for dropdown menus, 'make sure your menus are between 5 and 7 items on it', no more, no less, because here's what the research leads us to believe why that would be good. I think our design guidelines for voice these days are much the same. [...] Based on what we know, here are some good guidelines. [...] So, I like the fact that our current guidelines are more a set of best practices than hard bricks that we're laying down to follow" (P3).

In this P3 firstly sets up GUI design guidelines as an analogy to then suggest that guidelines for VUI may be more or less 'strict'. Overall, we can conclude that the understanding of the terms 'design principles', 'guidelines', 'heuristics' and 'best practices' varies quite strongly between practitioners and there is little to suggest the presence of a set of agreements (tacit or not) over how such language is to be deployed.

4.2 Doing VUI design with guidelines

We talked to interviewees about their VUI design process experiences and the role of design guidelines in them. In this section we report firstly on their accounts of this process, then we turn to examining how interviewees described guidelines meshing with this process (if they could). Then we describe various 'considerations' which were not explicitly described as 'guidelines' but nevertheless acted as a vernacular set of stable concerns for design across instances.

Interviewees described beginning by determining the scope of project by collaborating with stakeholders and clients, identifying user needs to develop user personas, establishing if voice is the appropriate medium, and defining use cases and a persona for the VUI.

Next, in response to identified needs, interviewees described typically starting prototyping work by writing a 'script' or a dialogue to depict the interaction between the user and their agent [20]. Interviewees transposed the concept of the 'happy path'—the 'optimal' or 'best' interactional route through the design that satisfies various stakeholder and user requirements—to that of the VUI, however notably introducing the notion of a 'script' into this process, i.e., the simplest conversational path the user might take to accomplish their goal [26]. This 'script' was then subject to further elaboration via prototyping processes, eventually working towards an object more compatible with the requirements of VUI development e.g., a "bulky flow chart" or dialogue in a tree form [43] as P7 outlines nicely here:

"So, we do the sample scripts, if I have time, I annotate them with kind of references to design guidelines. Once we approve the scripts or start it, usually I'm also recording myself kinda talking through the dialogue. I have a mic; you know right here and I'll sometimes I'll record myself with Alexa having that dialogue for me. [...] Once we had a point where we feel good, the client feels good, they've heard a few recordings, they have a good sense for how at least the happy path is going to sound, that's when I go in and start to do really bulky flow chart. And I started to say, OK, these prompts that we're providing give rise to these conversational pathways." (P7)

Finally, practitioners would evaluate VUI designs, but the approaches varied: having "users read through [...] scripts" (P1), employing "expert reviews" (P9) or a "friendly user test" with 'good' users (P4), testing on "mockups" and "Wizard of Oz testing" (P9), log file evaluation (P4).

Accounts from our interviewees had a collective coherence in that they all described a surprisingly similar design process for their VUI work. We observed that interviewees tended to articulate process in line with structures of design thinking to explain the VUI design process they followed. Their accounts also followed the form of 'standard' narratives about user-centred design projects (e.g., see ISO 9241-210). We thus noticed that designers tended to articulate their process in ways similar to GUI design process discover, define, ideate, design and evaluate. This does not tell us that VUI design is 'similar' to classical UI design, merely that the approach taken is consonant with it—whether this is the 'right'

approach is a different question, as this could indicate simply applying old context to new without questioning if the approach needs fundamentally reengineering.

4.2.1 How did VUI guidelines feature in accounts of design process?

Given interviewees' rich accounts of their design processes, perhaps surprisingly we found that most of our interviewees struggled to articulate the how design guidelines featured in their design process at all. Instead, they tended to talk about such guidelines in terms of broad design resources that may or may not be called upon as 'tools to hand' at a given moment in their design practice.

Building upon this, we noted a correlation between how the VUI designers we interviewed reported using (for instance) platform-specific VUI guidelines, and how much experience they stated having in the field. Those newer to VUI design tended to describe using guidelines as an authoritative resource, arguing, for instance, that various platform guidelines from Google or Amazon offered a firm "foundation" (P7) for establishing their design practice. Yet, the development of design competence [16, 17] played an important role in VUI design practice. One more experienced participant expressed it this way:

"In those early days, I like read a lot of those design guides and kind of treated them as like a set of rules. And those things about like how to word a prompt in a way that's easy for people to listen to or what are certain types of error messages that you should always include things like that. Now I would say that I have a much deeper understanding of design and how it's like a user experience process. I actually don't like the concept of design guides anymore because I find them very limiting and I think there is no best way to do anything at all. [...] It's about weighing those pros and cons, forming hypothesis and testing it." (P1).

Further, as competency with VUI design developed, we found that interviewees reported developing their own design guidelines and curated 'collections' from more formal guidelines. This was especially the case when dealing with cross-platform VUI design where specific platform guides may no longer serve this higher-level purpose. These home-grown guidelines would be rooted in practitioners' experiences within VUI design practice.

4.2.2 Design considerations

We found interviewees also revealed there are several factors which influence the design choices they make that act in a somewhat similar way to 'design guidelines' but nevertheless were not referred to as such. We've called these 'considerations' to indicate their relevance across instances of VUI design practice (since it would be presumptive for us to label them 'guidelines' given our interviewees did not do so themselves) and their quality more as 'questions' to pose.

Interviewees shared that before they would begin designing for voice, it was important to consider if voice was the appropriate modality for the application. This is important as not every application required voice and instead other forms of conversation design could be applied (e.g., chatbot). P7 describes having to make this decision, sometimes having to say "hey, that's terrible idea, let's not do that voice, not a great voice experience', or vice versa, you know, like 'hey, that would be a fantastic voice experience'". VUI designers interviewed also reported weighing if voice is to be the only mode or multimodal integration is required. For the latter, there are several other aspects to consider, including "people mixing modes, switching between modes, signaling which mode they prefer versus just using a mode that's faster" (P3).

Additionally, we found interviewees described several considerations to weigh as they scripted interactions, including how the language of the VUI is 'built' i.e., its 'components', how content is delivered, and how temporality (e.g., length of speech) shapes design (P8). Interviewees argued that the components of language such as syntax, grammatical rules, semantics and lexical morphology dramatically influence the user experience of voice interfaces (P6) and therefore provided a go-to set of reflective considerations in the course of VUI design. For example, P3 mentioned plurals: "you have one item in your cart or you have two items in your cart" and also gave an example of lists, comparing "your choices are bananas, oranges, tomatoes" with "or tomatoes". P3 here wanted to avoid the impression of a "bulleted list" and instead offer interaction that was more "natural". Interviewees also raised considerations of avoiding long speech, which they felt often results in users missing out on key information and leads to poor user experience.

4.3 Harmonisation of codified VUI design knowledge

As we have noted earlier, many existing codifications of VUI design knowledge seem to overlap conceptually. Further, it is argued that there are benefits from cohering around common or shared set of (e.g.) guidelines in practice [11, 18]—practices such as interface consistency have been important in enhancing effective use. But our interviewees presented various responses this idea of ‘harmonisation’ for VUI design. While interviewees did acknowledge that guidelines are broadly similar, they felt there were other considerations at play. Some suggested in spite of the potential benefits of having a consolidated set of guidelines, it might be too early to do so due to design immaturity. It was sometimes argued by our interviewees that harmonisation, associated with the development of shared approaches or more formal standards will likely grow in potential with greater understanding of the field of VUI design and VUI technologies.

However, some interviewees seemed more doubtful about the notion of harmonisation. They noted it would be challenging because the impetus towards a shared set of common guidelines potentially worked against what they saw as the constantly evolving nature of language:

“Standards really have come out of our software and before that equipment. So, you know standards [are] really important because otherwise we build infrastructure systems that don't work with each other. [...] But language is not a infrastructure system, and it isn't a technical system; it wasn't invented in the same way. [...] Language continues to be very emergent and flexible, and we interact with voice today, differently than we did 25 years ago and that was different than 25 years before that.” [P8]

Instead, interviewees saw harmonised design knowledge for VUIs being potentially relevant more for narrow, particular technical characteristics of VUI design, such as audio, volume, clarity and comprehension, while drawing parallels with standards associated with visual features (e.g., colour, contrasts, etc.).

There are also platform-related challenges to harmonisation. P2, for instance, argued that domain specific differences for different voice assistants constrained design along different line; P2 suggested that Amazon would potentially offer better shopping experience than Google, so Amazon’s guidelines would be tailored to serve that domain, whereas Google’s would be better suited to respond to queries. Lastly, some interviewees reminded that modality-based guidelines are also platform-specific. So, guidelines around switching between modes, displaying data on multimodal interfaces would differ based on the platform being used.

5 DISCUSSION & CONCLUSIONS

Voice user interface design practitioners shared insights into the present state and role of codifications of VUI design knowledge in practice which we believe are not currently represented in HCI research. Here we bring together some of the key points our study raises, and the implications it surfaces for further research.

A key question has been about how practitioners consider and potentially use existing codified design knowledge. Our study shows a somewhat mixed picture. For starters, there is considerable lack of common shared language or terminology for consistently describing and distinguishing principles, guidelines, heuristics and so on. Thus, a source of terminological trouble might be caused by a) reliance on platform design guidelines and b) vendor-introduced brand distinctions. So, we found some participants (e.g., P1, P2)—in attempting to describe what they thought of the difference between e.g., “design guidelines” and “heuristics”—recalled Amazon’s (platform) language describing these as “tenets”. The author of a blogpost published on Amazon’s website refers to these design patterns as “tenets” too [6]. In contrast, in the Alexa design guidelines, what the participant recalled as “tenets”, are listed as “design patterns” [3]. Furthermore, IBM uses ‘design guidelines’ and ‘principles’ to refer to the design guidance they published [28]. Thus, we can build a case about this confusion over varying terminology being driven by platform and vendor. HCI research should reflect on the role of platform and vendor when producing design knowledge for VUI practitioners—presently we think the importance of these practical issues tend to be underplayed.

Some participants argued that present codifications like guidelines basically do not fit their work practices, which is why they reject them (or pay less attention to them). For instance, they pointed out that guidelines listed as “dos and don’ts” “simplifies the work” but is “harmful to the user experience” (P1). The reasoning here is that participants argued this

encourages a tendency (especially when new to VUI design) to follow them as rules, rather than trying and testing if the design guidelines actually apply to a given use case and is a success with the users. It is less a case of ‘missing’ guidelines and more that practitioners have to work out *how* to fit guidelines *into* their practical work. Further, by their own accounts, practitioners described various different ‘modes’ of practice with respect to applying VUI guidelines, from considering guidelines as a kind of knowledge-bootstrapping process or pedagogical tool, through to a ‘pick and mix’ approach, as well as a ‘roll your own’ approach. Practitioner attitudes towards guidelines also mesh with the development of their own particular journey towards VUI design competence, consonant with UX competence development more broadly [16, 17]. Practitioners may well start their journey in voice design by adopting design guidelines but over time become more agile and adaptive, can better judge what actually works, and subsequently rely more on the needs of a given design process than guidelines themselves.

Thus, we argue that appreciating the texture of these various different modes and their relationship to competence development is likely critically important if CHI is going to continue its investment in conversational user interface research, and in seeking the development of design guidelines, heuristics, and so on, which aspire to industry adoption and use. If we create such VUI design knowledge, what practitioner career point are they aimed at? Further, what style or way of working might they be tailored towards? So, on one hand, our research encourages industry practitioners, especially those new to VUI design, to think more critically about existing platform-oriented design knowledge and not treat it as a ‘one-stop shop’ for designing voice interfaces. On the other it calls for further investigation by HCI researchers into VUI designers’ actual practices, so that better descriptions of these different ‘modes’ may be established.

The second point we return to is matters of harmonising VUI design guidelines, i.e., ways in which guidelines etc. as codifications of design knowledge—often initially tied to specific platforms—may be brought together in some way, whether as a formal standard or something looser. Maybe surprisingly, practitioners had mixed feelings about the potential value of this concept in spite of industry initiatives to do so [4]. This is further complexified by the nature of how they reported use of existing guidelines in practice and how VUI design processes unfold, as we described above. Specifically, we suggest that the VUI design process itself along with various informal design ‘considerations’ jointly shape how and in what ways design guidelines are ‘used’ (if at all). For instance, platform-relevant matters seemed consistently to loom large across the design processes engaged in and any harmonisation of design knowledge needs to navigate this. This means that there are potential bumps along the road for HCI research into VUIs if it is to develop guidelines that knit together multiple approaches in an attempt to produce ‘unified’ design techniques for voice user interfaces. The tendency in academic research is towards generalisation and abstraction, which seems resonant with the idea of ‘harmonisation’ of codified design knowledge—this approach is valued due to the emphasis on knowledge accumulation and progress. But this urge might run counter to the needs of actual practitioners, so we are left questioning *who* attempts at harmonisation might ultimately be serving? When is it appropriate and when may harmonisations between codified design knowledges be counter-productive for practice?

We also need to discuss the language codified design knowledge for VUIs: of guidelines, heuristics, best practices, principles, and the like. It is an open question as to whether the fluidity of this language really ‘matters’. We saw, for instance, some practitioners switching between them seamlessly, while others had strong distinctions to make. Instead, we suggest HCI research into VUI design moves beyond language to consider *motivations* that drive terminological troubles—for instance, as we pointed out, there was a distinction to be made between more ‘prescriptive’ guidelines and those inculcating certain ‘values’. This poses a question regarding the style and type of guidelines being generated in HCI research and in which part of the spectrum from ‘prescription’ to ‘values’ they sit. Are we making specific must-follow rules or suggesting particular values to be embodied?

Finally, there are implications for further research here too. Based on our survey findings, it appears there is a need for codification of voice design guidance and as a growing field there is much rapid movement in this area. Our interview study has been limited to a relatively small number of experienced VUI designers in industry and drawn out some initial themes. Whilst we have not looked directly at VUI designers’ practices, we have seen accounts of those practices in interview and note the prominent gaps that further research should examine, such as further details on the objects of collaboration and how VUI designers work with developers and other designers. As part of this, there are likely many aspects of how codified design knowledge is considered by designers and enacted *in practice* that we have missed. As Murad et al. state “For VUI-

specific guidelines or heuristics to be effective, they must be *adoptable* by designers” (our emphasis) [14]. HCI and CUI research into conversational user interfaces in general thus needs to understand VUI design processes ‘in the wild’ to adequately determine what type of object we are or should be building, and whether, say, ‘design guidelines’ are even an appropriate vehicle for those aims. Further work should seek to investigate practitioners’ work practices with the design of VUIs to help calibrate ‘adoptability’. From our interviewees’ struggle to pinpoint in detail how codifications like design guidelines were embedded in their practice, we suggest that study of an ethnographic nature examining work practice may have a greater potential to unpack how guidelines are worked with / into practices, that moves beyond more ‘scenic’ accounts of their process [5].

ACKNOWLEDGEMENTS

We thank the reviewers for their valuable feedback and suggestions. This work was supported by the Engineering and Physical Sciences Research Council [EP/T022493/1].

Data access statement: In line with consent restrictions in place for interviews performed for this paper including sensitivities towards protecting identities and confidential information (e.g., non-disclosure agreements, commercially sensitive information), anonymised transcripts are available at the following resource: <http://dx.doi.org/10.17639/nott.7174>

REFERENCES

- [1] 2019. *Proceedings of the 1st International Conference on Conversational User Interfaces*. Association for Computing Machinery, New York, NY, USA.
- [2] 2020. *Proceedings of the 2nd Conference on Conversational User Interfaces*. Association for Computing Machinery, New York, NY, USA.
- [3] Amazon. Alexa Design Guide - Get Started with the Guide. Retrieved February 8, 2021 from <https://developer.amazon.com/en-US/docs/alexa/alexa-design/get-started.html>
- [4] Amazon. Multi-Agent Design Guide. Retrieved February 25, 2021 from <https://developer.amazon.com/en-US/alexa/voice-interoperability/design-guide#multi-agent-design-guide>
- [5] Andrew Crabtree, Mark Rouncefield, Peter Tolmie. 2012. Précis. In *Doing Design Ethnography*. Springer London, London, UK, 1-5. DOI:https://doi.org/10.1007/978-1-4471-2726-0_1
- [6] Anna Van Brookhoven. 2019. New Alexa Design Guide: Create Engaging Alexa Skills Using Situational Design. (May 9, 2019). Retrieved August 10, 2021 from <https://developer.amazon.com/blogs/alexa/post/ee0e00c9-37cd-46ac-8695-06552e0885b0/new-alexa-design-guide-create-engaging-alexa-skills-using-situational-design>
- [7] Apple. Human Interface Guidelines. Retrieved February 8, 2021 from <https://developer.apple.com/design/human-interface-guidelines/>
- [8] Arun George. 2017. Designing Voice Experience. (September 28, 2017). Retrieved February 11, 2021 from <https://uxdesign.cc/voice-user-experience-design-and-prototyping-for-mere-mortals-ef080c843640>
- [9] Bernhard Suhm. 2003. Towards best practices for speech user interface design. In *Eighth European Conference on Speech Communication and Technology*, September 1-4, 2003, Geneva, Switzerland. 2217-2220.
- [10] Bernhard Suhm. 2008. Ivr Usability Engineering Using Guidelines And Analyses Of End-to-End Calls. In *Human Factors and Voice Interactive Systems. Signals and Communication Technology*. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-68439-0_1
- [11] Bill Killam, and Marguerite Atry. Human Factors Guidelines for Interactive Voice Response Systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Volume 44, Issue 3*, July 2000. 391-394, doi:10.1177/154193120004400308.
- [12] Cathy Pearl. 2016. Designing Voice User Interfaces: Principles of Conversational Experiences (1st, ed.). O'Reilly Media, Inc., Sebastopol, California, CA.
- [13] Christine Murad and Cosmin Munteanu. 2019. "I don't know what you're talking about, HALexa": the case for voice user interface guidelines. In *Proceedings of the 1st International Conference on Conversational User Interfaces (CUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 9, 1-3. DOI:<https://doi.org/10.1145/3342775.3342795>
- [14] Christine Murad, Cosmin Munteanu, Benjamin R. Cowan and Leigh Clark. Revolution or evolution? Speech interaction and HCI design guidelines. *IEEE Pervasive Computing*, 18, 2 (2019), 33-45. DOI: 10.1109/MPRV.2019.2906991.
- [15] Christine Murad, Cosmin Munteanu, Leigh Clark, and Benjamin R. Cowan. 2018. Design guidelines for hands-free speech interaction. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, 269-276. DOI:<https://doi.org/10.1145/3236112.3236149>
- [16] Colin M. Gray, Austin L. Toombs, and Shad Gross. 2015. Flow of Competence in UX Design Practice. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3285-3294. DOI:<https://doi.org/10.1145/2702123.2702579>
- [17] Colin M. Gray. 2014. Evolution of design competence in UX practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1645-1654. DOI:<https://doi.org/10.1145/2556288.2557264>
- [18] Daryle Gardner-Bonneau. 1999. Guidelines for Speech-Enabled IVR Application Design. *Human Factors and Voice Interactive Systems*. The Springer International Series in Engineering and Computer Science, vol 498. Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-2980-1_7
- [19] Deepika Mittal. 2018. How Nielsen's 10 usability heuristics apply to Voice UI (April 2, 2018). Retrieved February 11, 2021 from <https://uxdesign.cc/10-usability-heuristics-for-voice-user-interface-design-69ad9ea4f166>
- [20] Designit. 2018. Scriptwriting for Voice UI. City. (December 11, 2018). Retrieved February 25, 2021 from <https://medium.designit.com/scriptwriting-for-voice-ui-e1035fffd2a5>
- [21] Diana Deibel and Rebecca Evanhoe. 2021. *Conversations with Things: UX Design for Chat and Voice*. Rosenfeld Media, New York, NY, USA.
- [22] Fjord. A Guide to Voice Interfaces: Six Principles for Designing for Voice UI. Retrieved December 25, 2021 from <https://voiceui.fjordnet.com/>

- [23] Google. Conversation design. Retrieved February 8, 2021 from <https://designguidelines.withgoogle.com/conversation/conversation-design/welcome.html>
- [24] Google. Material Design Introduction. Retrieved February 8, 2021 from <https://material.io/design/introduction>
- [25] Grace Chen. 2019. A Concrete Guide to Designing Voice User Interfaces. (May 11, 2019). Retrieved February 11, 2021 from <https://medium.com/voice-tech-podcast/a-concrete-guide-to-designing-a-voice-ui-bd186777b704>
- [26] Guillaume Privat. 2018. Fundamental Elements of VUI Design. (December 10, 2018). Retrieved February 23, 2021 from <https://blog.prototypr.io/fundamental-elements-of-vui-design-8630077a7009>
- [27] Herbert P. Grice. 1975. Logic and conversation. *Syntax and semantics Vol. 3: Speech acts*. Academic Press, New York.
- [28] IBM. Talk meets technology: Conversation design guidelines. Retrieved February 8, 2021 from <https://conversational-ux.mybluemix.net/design/conversational-ux/>
- [29] Interaction Design Foundation. Design Principles. Retrieved February 8, 2021 from <https://www.interaction-design.org/literature/topics/design-principles>
- [30] Interaction Design Foundation. How to Conduct a Heuristic Evaluation for Usability in HCI and Information Visualization. Retrieved February 8, 2021 from <https://www.interaction-design.org/literature/article/how-to-conduct-a-heuristic-evaluation-for-usability-in-hci-and-information-visualization#:~:text=A%20heuristic%20is%20a%20simple,these%20%E2%80%9Crules%20of%20thumb%E2%80%9D>
- [31] Interaction Design Foundation. The 7 Factors that Influence User Experience. Retrieved February 8, 2021 from <https://www.interaction-design.org/literature/article/the-7-factors-that-influence-user-experience>
- [32] Interaction Design Foundation. Voice User Interfaces. Retrieved December 20, 2021 from <https://www.interaction-design.org/literature/topics/voice-user-interfaces>
- [33] Jakob Nielsen. 1994. 10 Usability Heuristics for User Interface Design. Nielsen Norman Group, City. (April 24, 1994). Retrieved February 8, 2021 from <https://www.nngroup.com/articles/ten-usability-heuristics/>
- [34] Jon Swain. A Hybrid Approach to Thematic Analysis in Qualitative Research: Using a Practical Example. In *SAGE Research Methods Cases*. <https://www.doi.org/10.4135/9781526435477>
- [35] Katherine K. Fu, Maria C. Yang and Kristin L. Wood. 2015. Design Principles: The Foundation of Design. In *Proceedings of the ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 7: 27th International Conference on Design Theory and Methodology*. Boston, Massachusetts, USA. August 2–5, 2015. V007T06A034. ASME. <https://doi.org/10.1115/DETC2015-46157>
- [36] Kristina Höök, Peter Dalsgaard, Stuart Reeves, Jeffrey Bardzell, Jonas Löwgren, Erik Stolterman and Yvonne Rogers. Knowledge Production in Interaction Design. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, April 2015, Seoul, Republic of Korea. Association for Computing Machinery, New York, NY, USA, 2429–2432. DOI:<https://doi.org/10.1145/2702613.2702653>
- [37] Kyoko Sugisaki and Andreas Bleiker. 2020. Usability guidelines and evaluation criteria for conversational user interfaces: a heuristic and linguistic approach. In *Proceedings of the Conference on Mensch und Computer (MuC '20)*. Association for Computing Machinery, New York, NY, USA, 309–319. DOI:<https://doi.org/10.1145/3404983.3405505>
- [38] Michael Harris Cohen, James P. Giangola and Jennifer Balogh. 2004. Voice User Interface Design. Addison-Wesley, Boston.
- [39] Michael McTear, Zoraida Callejas, and David Griol. 2016. Speech Input and Output. *The Conversational Interface*. Springer, Switzerland. DOI: https://doi.org/10.1007/978-3-319-32967-3_5
- [40] Nicole Yankelovich, Gina-Anne Levow and Matt Marx. 1995. Designing SpeechActs: Issues in speech user interfaces. In *Proceedings of the SIGCHI conference on Human factors in Computing Systems (CHI '95)*. ACM Press/Addison-Wesley Publishing Co., USA, 369–376. DOI:<https://doi.org/10.1145/223904.223952>
- [41] Robert J. Moore and Raphael Arar. 2019. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. Association for Computing Machinery, New York, NY, USA. DOI:<https://doi.org/10.1145/3304087>
- [42] Susan Weinschenk and Dean T. Barker. 2000. Designing effective speech interfaces. John Wiley & Sons, Inc., USA.
- [43] Teresa Castle-Green, Stuart Reeves, Joel E. Fischer, and Boriana Koleva. 2020. Decision Trees as Sociotechnical Objects in Chatbot Design. In *Proceedings of the 2nd Conference on Conversational User Interfaces (CUI '20)*. Association for Computing Machinery, New York, NY, USA, Article 27, 1–3. DOI:<https://doi.org/10.1145/3405755.3406133>
- [44] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. In *Qualitative Research in Psychology, Volume 3 Issue 2*. 77–101. DOI: 10.1191/1478088706qp063oa
- [45] Yvonne Rogers, Helen Sharp, and Jenny Preece. 2011. *Interaction Design: Beyond Human - Computer Interaction* (4th ed.). John Wiley & Sons.
- [46] Zhuxiaona Wei and James A Landay. 2018. Evaluating speech-based smart devices using new usability heuristics. *IEEE Pervasive Computing*, 17, 2 (2018), 84–96. DOI: 10.1109/MPRV.2018.022511249.