Improving Predictive Text: Defining the Problem Space and Needfinding Plan

Clinton McCardell
cmccardell3@gatech.edu

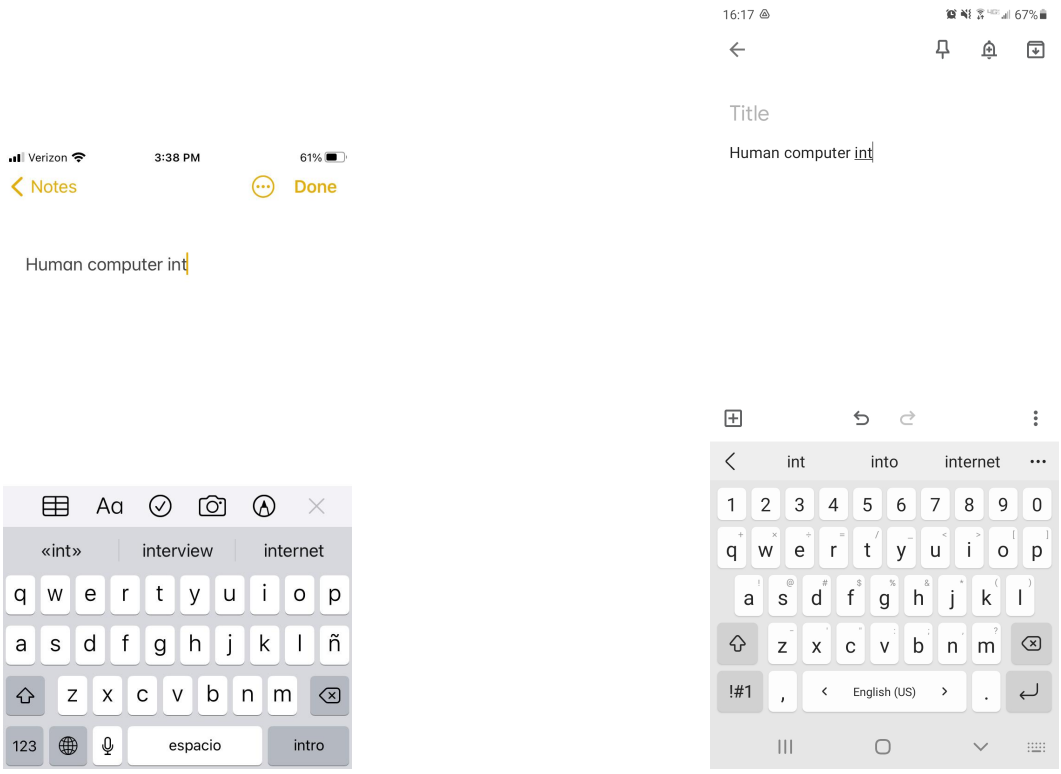
***Abstract*—**With the rising ubiquity of mobile devices such as smartphones and tablet computers, a greater percentage of text entry is taking place away from a standard keyboard. Because of this, the task of quickly and accurately entering text is becoming more burdensome for users who are faced with smaller screens, divided attention, and less tactile precision. Predictive text, the act of suggesting and correcting words as a user types, was introduced to improve interaction (Mackenzie, 2002), but suffers from various usability problems. We travel through one cycle of the design life cycle in order to improve predictive text and thus the user's ability to quickly and efficiently digitize their thoughts.

1 DEFINING THE PROBLEM SPACE

Though specifics such as coloration and location may be unique to each platform, some form of predictive text is present on virtually every mobile device. Depending on the sophistication of the predictive text functionality, predictive text can suggest words to users as they type, correct perceived misspellings and poor grammar, and even complete entire sentences. Predictive text draws it suggestions and corrections from a baked-in corpus as well as prior user interactions. This means that as a user uses their mobile device more frequently, predictive text should become more accurate and thus useful. Predictive text activates on a mobile device nearly every time the keyboard appears–whether the user is composing an email, writing a text message, or making a note or calendar entry (Figure 1).

Although the using predictive text can make user text entry more efficient (Arnold, Chauncey, and Gajos, 2020), many improvements could be made to increase usability. Though the purpose of needfinding is to identify these improvement spaces, some cursory ideas involve expanding the initial corpus to include a wider variety of discipline-specific words and acronyms and adding a command to revert a corrected word or phrase to the previously-written words (such as through the use of the backspace key). Also, because global communication

is becoming more prevalent, better suggestions and corrections during code-switching (quickly switching between languages in conversation) would also be a big improvement. During needfinding, we will gather and refine these improvements to make the task of quickly and effectively entering text on a mobile device more usable through improvements to predictive text.



(a) Apple iPhone 6

(b) Samsung Galaxy s20

Figure 1—A standard use case for predictive text while writing a note on a variety of mobile devices (Images by the author).

2 TARGET USER TYPES

Because mobile devices are so ubiquitous and are commonly used by the majority of people, we are interested in designing for people of all ages and genders. Mobile device use is also not constrained to people of certain socioeconomic classes or ethnicities, so we plan to include as wide of a diversity of people as possible in our work. Mobile device users have varying levels of device expertise, so we want to include users across the expertise spectrum as well. We want to improve predictive text for all skill levels–those who are novices and might need extra guidance, but also expert users who do not want or need additional

inputs or interface to hinder their workflows. Similar to level of expertise, we want to include users with a variety of motivations for using text entry on their mobile device. We are interested in the casual user who occasionally writes text messages or emails to their friends and family, but also those who rely on mobile devices to organize their lives or assist them in their professions. Though there might not exist a one-size-fits-all solution for all demographics and user types, it is crucial that feedback and thoughts are obtained from all kinds of users during the design process and especially before any changes are made to the existing interface or functionality.

3 NEEDFINDING PLAN 1: PARTICIPANT OBSERVATION

As a preliminary needfinding activity, we will perform a participant observation using a mobile device for text entry and interacting with our device's predictive text capabilities. In the following subsections, we outline rationale and methods for this observation, discuss the data that will be gathered, relate this needfinding activity to the data inventory, and outline the potential biases we might encounter during the participant observation.

3.1 Methods

During a participant observation, the researcher participates in the task themselves and observes how they react to said task. For the mobile text entry task, we will participate by entering text into our mobile device in a variety of contexts, namely composing a form email, taking a note with generic sample text, taking a note with more complex, discipline-specific text, and sending a form text message. We will retrieve this form text from example emails and notes that might accurately represent an average user's mobile text entry interaction. We will repeat this exercise in a variety of different settings–such as while sitting at a desk, walking, and conversing–to better emulate what a user might experience when interacting with their mobile device. During each text entry context/environment pair, we will first call up the application that will be used to enter the text, then write out our message, and finally note our thoughts about the task and our interactions with predictive text.

3.2 Data Gathering

The data that is gathered during a participant observation are the researcher's notes and thoughts that are recorded while or immediately after performing the

task. While writing out our sample texts in the variety of different contexts, we will also run a voice recorder to capture any spoken thoughts. A screen recorder could also be used to capture the interaction we have with the device and to review the suggestions and corrections that are made throughout the process. After performing the task, we will also record any final thoughts about the experience, specifically successes and failures during the task, thoughts about the interface, and potential opportunities we see for improvement.

3.3 Participant Observation and the Data Inventory

Of the seven components of our data inventory–data we focus on gathering about the task, the users, and the interaction during needfinding–participant observation helps answer two of them: what a user needs and the tasks to be performed. By participating in the mobile text entry task ourselves, we better understand and identify what the tasks are and what an average user needs to succeed at their tasks. By exploring various contexts and environments, we can see how these affect the difficulty and accuracy of mobile text input and see how predictive text bridges or fails to bridge the gap caused by the user's environment.

3.4 Combating Potential Biases

Though by participating ourselves in the mobile text entry task we learn about the task and how to achieve it, we must be careful not to let our own experiences cloud the rest of our needfinding moving forward. Specifically, we may experience confirmation bias–that the strengths and weaknesses we find are those which the average user will identify–as well as observer bias–we may act differently completing the task because we are metathinking about the experience. It is important to remember that we are not the user for whom we are designing, so any observations and data we collect serve only to springboard future questions and interactions with real users. By remembering that we are not the user, we can properly frame our experiences and collected data as cursory investigation and limit the impact of these biases.

4 NEEDFINDING PLAN 2: THINK-ALOUD AND POST-EVENT PROTOCOLS

As a secondary needfinding activity, we will perform various think-alouds and post-event protocols with real mobile device users. In the following subsections,

we outline rationale and methods for these protocols, discuss the questions that will be asked, relate this needfinding activity to the data inventory, and outline the potential biases we might encounter during the protocols.

4.1 Methods

While participating in a think-aloud or post-event protocol, users perform a task while thinking through the steps out loud so the researcher can understand their line of thinking. The researcher also asks questions either during or immediately after the task For the mobile text entry think-aloud, we will have two or three users perform the same text entry tasks as described in the Participant Observation (entering form emails, messages, and recording sample text notes). While the users are completing their tasks, we will ask a minimal number of questions, instead focusing on what the users' natural reactions are to the task. After each task, we will ask more in-depth questions about the tasks. During and after the tasks, we will record the thoughts of the users using an audio recording device. We will also use screen recording software on the mobile device to capture the text entry and specific suggestions or corrections the predictive text functionality made.

4.2 Questions to Ask

As stated, while the user is actually performing the text entry task, we will keep our questions to a minimum. However, we will ask clarifying questions about comments that the user makes or actions they perform. We will mainly rely on the users' out-loud thought processing during the task. Immediately after each task, we will ask each user a variety of questions about their perceived completion of the text entry task. For example, some questions we will ask include how the user would rate their overall experience completing the task, what parts or subtasks went well and with which parts did they struggle, which parts or subtasks were pleasurable or frustrating, and what needs for improvement did they see while performing the text entry. We will frame these questions in terms of the user's interaction both with mobile text entry in general and also specifically concerning predictive text. We will also ask follow-up questions specific to the answers given to the previous questions, such as why they gave a high/low rating, why the subtask was easy/hard, and why they see a need for improvement.

4.3 Think-alouds and the Data Inventory

Of the seven components of our data inventory, the think-aloud and post-event protocols address three of them, namely: user tasks, user subtasks, and what users need to successfully complete their tasks. By listening to real users think through actually performing the task we are studying, we can identify the subtasks that a user completes in order to accomplish the larger, overall task of mobile text entry. We can ask follow-up questions about these subtasks to understand them better and compare subtasks between users. Additionally, because we will be intimately studying the exact process a user takes to complete their task, we will gain insight into the things a user needs to successfully complete it. This may include components the interface already provides the user or functionality that is not given, but desired by the user.

4.4 Combating Potential Biases

Though in a think-aloud, we are observing and interacting with real users as they complete the assigned text entry task, biases are still present. For example, because the users know they are being observed and have to narrate their actions out loud, they may experience observer bias and alter their natural actions. Similarly, demand characteristics may also be present. Additionally, because we will be asking them questions about their experiences, there may be some courtesy or social desirability bias for users to give answers that please us as researchers. Because we are asking the users questions, there is always a possibility of question bias from the ordering and phrasing of the questions. Finally, there may be some recall bias because we are asking questions after the user has completed the task.

To minimize the effects of the observer bias, we will ask the users to perform the tasks as naturally as possible and limit our questioning during the tasks. For courtesy bias, we will remind the users to be as honest as possible and let them know that the answers to their questions will not affect the outcome of the research. Question bias will be minimized by grouping similar questions together and by wording our questions neutrally. Finally, because we are asking post-event questions immediately after completing the tasks, this will minimize recall bias.

5 NEEDFINDING PLAN 3: SURVEY

As a final needfinding activity, we will conduct a survey regarding mobile text entry and predictive text usage. In the following subsections, we outline rationale and methods for conducting the survey, discuss how potential participants will be selected, relate this needfinding activity to the data inventory, and outline the potential biases we might encounter.

5.1 Methods

The purpose of using a survey for needfinding is to gather lots of broad data about the user base quickly and efficiently while receiving many responses. To take advantage of the benefits of a survey, we will first include demographic questions about age, gender, ethnicity, and location to better familiarize ourselves with the user base. Then we will ask about their mobile device interactions, how frequently they perform text entry on their mobile device and for what purpose, where they are located when performing mobile text entry, and what level of expertise they have using their device for text entry. We will also define predictive text and ask about the respondent's familiarity with it. Finally, we will adapt SUS questions to determine how usable users consider mobile text entry and predictive text to be, and ask for suggestions for improvements.

5.2 Survey Participants

Because we would like to capture an accurate sample of the mobile text entry user base, we will use a variety of participant-finding activities with the goal of receiving at least 25 responses. We will recruit participants from fellow Georgia Tech students, co-workers, friends and family members, and social media followings. We believe that by including such a wide range of recruitment methods, we can diversify our sample and get more accurate overall responses.

5.3 Surveys and the Data Inventory

Of the seven components of our data inventory, a user survey addresses four of them, namely: determining who the user is, where the user is, the context in which the user performs the tasks, and the goals the user has for performing the task. Because we will begin our survey with demographic questions, we will get a good idea of who are user base is and where they live and work. By asking where, why, and with what frequency they perform mobile text entry, we will

better understand the external factors that will affect their precision and speed and help us properly determine the appropriate level of control the predictive text capabilities should have.

5.4 Combating Potential Biases

There are a variety of potential biases that should be minimized during a needfinding survey. For example, because we send the survey to many people and not all of them respond, we are prone to response bias where only those with polarizing experiences respond. To minimize this bias, we can make our survey not as widely available and ask general questions that everyone can answer. Recall bias can also be present in surveys as the respondent must draw upon past experiences to answer it. We can avoid this bias by asking questions about trends and general experience as a whole and not focus on isolated events the respondent must remember. Finally, because the survey is done completely via written text, question bias–including phrasing and ordering–could be prevalent. To minimize this bias, we will reuse SUS questions that have already been refined to be neutral and follow a wide-to-narrow progression. Additionally, before releasing our survey, we will review the questions and refine them to remove language that could bias responses.

6 REFERENCES

[1] Arnold, Kenneth C., Chauncey, Krysta Z., and Gajos, Krzysztof undefined (2020). "Predictive text encourages predictable writing". In: *Proceedings of the 25th International Conference on Intelligent User Interfaces*. DOI: 10.1145/3377325.3377523. URL: https://doi.org/10.1145/3377325.3377523.

[2] Mackenzie, I. Scott (2002). "KSPC (Keystrokes per Character) as a Characteristic of Text Entry Techniques". In: *Human Computer Interaction with Mobile Devices Lecture Notes in Computer Science*, pp. 195–210. DOI: 10.1007/3-540-45756-9_16. URL: http://www.yorku.ca/mack/hcimobile02.html.