



Experiences with and New Application Ideas for the Interspeech App

Sebastian Möller, Tilo Westermann

Quality and Usability Lab, Telekom Innovation Laboratories, TU Berlin

tilo.westermann@tu-berlin.de, sebastian.moeller@telekom.de

Abstract

For Interspeech 2013 in Lyon, we launched a first version of a mobile phone app, which should support conference participants prior to and during the conference. We anticipated that the app will be of use to conference organizers alike, and that it could be turned into an open-source tool for the spoken language technology community. In this contribution, we will demonstrate usage experiences collected during the 2013 and 2014 Interspeech conferences, and propose and implement ideas for carrying out speech research with such an app.

Index Terms: mobile phone app, Interspeech conference, community platform

1. Introduction

The Interspeech conference 2013 held in Lyon saw the introduction of a new conference app for smartphones, in both an iOS and an Android version. While conferences in the field of human-computer interaction (HCI) are providing apps for supporting conference attendees for some time (e.g. CHI, UIST), this was not the case for the Interspeech conference series so far. Based on an analysis of the functionality of existing conference apps and on an online survey with members of the scientific community in Germany, we defined a set of functions which either cover the information obtained in the printed conference program, or supplemented it by typical smartphone functions such as calendar entries, search, or full paper download [1]. Results indicated that the main focus for potential users is on obtaining the same information that a conference attendee otherwise gets via the conference program, and not functions which are available on the smartphone anyhow (such as social network integration). However, these desires expressed before the launch of the app were not 100% congruent with the ratings of perceived usefulness we obtained after the Lyon conference, cf. Table 1.

It can be seen that the variation of the mean is quite low with relatively high standard deviations, indicating different

Table 1. Ranking of available functions at the Lyon conference (1="most useful", 8="least useful")

Rank	Functionality (Android) (M ± SD)	Functionality (iOS) (M ± SD)
1	Paper download (2.72 ± 2.18)	Calendar integration (3.06 ± 2.13)
2	Search (2.75 ± 2.29)	Paper download (3.1 ± 2.07)
3	Authors (3.13 ± 1.99)	Search (3.37 ± 2.23)
4	Agenda (3.28 ± 2.62)	Floor plan (3.41 ± 2.13)
5	Floor plan (3.41 ± 2.19)	Agenda (3.46 ± 2.64)
6	Calendar integration (3.59 ± 2.15)	Authors (3.5 ± 2.16)
7	Paper recommendation (4.25 ± 2.05)	General information (3.56 ± 2.25)
8	General information (4.26 ± 2.21)	Paper recommendation (3.65 ± 2.38)

usage patterns and preferences. Independent-samples t-tests for all functions were conducted to compare usefulness ratings for Android and iOS versions of the app, but no significant difference was found. While rankings are similar for most functions, the calendar integration differs in that it takes the top rank for the iOS version, but the 6th for Android.

Based on the experiences in Lyon, the app has been reshaped and used for the Singapore conference in 2014. Figure 1a shows a screenshot of the original main menu next to Figure 1b with the resulting one. Both versions provide the following base functionality: *Agenda* for viewing the general conference program and listing sessions and papers. The paper detail view provides an abstract for the selected paper and the possibility to mark it as a favorite (and optionally adding it to the smartphone calendar), downloading the PDF, or to read the paper through a Text-To-Speech system, see below; *Authors* for an overview of participants and their papers; and a *Floor Plan* for spatial navigation through the conference (restricted to static images up to now). *Search* in Figure 1a was included in *Agenda* and *Authors* in Figure 1b, respectively. The *Information* tab with information about the conference and venue was replaced by a conversational agent (SARA) [2] providing information through a chat, see below.



Figure 1a: Screenshot of the original main menu (iOS version, Lyon '13); 1b: Screenshot of the updated main menu (Singapore '14).

2. Data Analysis

During the Lyon and Singapore conferences, we logged the following data points:

- **App session information:** time (i.e. start and end timestamps), screen flow (i.e. screens and duration per screen)
- **Device information:** platform, model, OS version, language setting, screen resolution, app version
- **User ratings:** an in-app questionnaire was launched at the last day of the conference soliciting ratings for overall quality, perceived usefulness, pragmatic and hedonic usability aspects, as well as free comments.

In the following paragraphs, we will report exemplary analyses of each of these categories; more details will be presented at the Show&Tell demonstration.

We reached a total of 735 users (406 iOS, 329 Android) in 2013, i.e. more than 50% of all Interspeech 2013 attendees. In 2014, the percentage of reached attendees was similar, with a total of 584 users (368 iOS, 216 Android). These users originate from a large variety of language backgrounds, as can be seen from the language settings. In 2013, we found 27

different language settings, the most popular being English (337 users), French (92), German (82), Japanese (53) and Chinese (42). This aligns well with the distribution of origins of conference attendees, suggesting that the app is equally accepted by all attendees, regardless of origin.

The distribution of downloads gives an indication of the app usage in relationship to the conference starting date (Aug. 25, 2013; Sep. 14, 2014). Figure 2 shows the number of downloads per day from 6 days prior to the conference to the end of the conference. Main triggers for download were a newsletter sent to participants (Aug. 23, 2013, Sep. 10, 2014) and a mention during the opening keynote (26th, 15th).

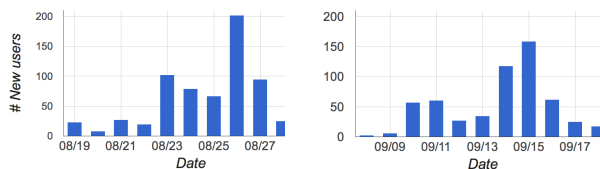


Figure 2: Number of downloads from 6 days prior to the conference to the conference end (left: 2013, right: 2014).

156 out of 735 (21%) users provided subjective ratings on the app in 2013. Their overall rating was positive for both versions, while ratings for the iOS version ($M=4.29$, $SE=.08$) were slightly better than for the Android version ($M=4.01$, $SE=.09$). This is also reflected in the ratings of the AttrakDiff Mini questionnaire, which reveals the Hedonic Quality – Identity dimensions with the best rating, followed by the Pragmatic Quality, and overall Attractiveness.

In addition to these scales, users provided wishes for new features. Most desired features include *organizational* ones (sending messages to conference attendees; feedback targeting the conference through the app; integration of sponsors), *social features* (“meet me” functionality; best paper voting; integration of coffee- and lunch-breaks; paper “like” button; option to contact the author of a paper), *user interface* optimizations, and *paper annotation* possibility.

3. Speech Technology Integration

The Singapore version of the app integrated two new speech technology features. First, we made it possible to read the abstract text of each paper using the native TTS integrated in the smartphone such as Siri for iOS. Second, the Human Language Technology Department, Institute for Infocomm Research, Singapore provided a dialogue system informing about Singapore, its touristic attractions and the conference, which could be launched using the *Ask Sara* tab (c.f. Figure 1b) in the app.

At the Show&Tell demonstration during Interspeech 2015, we’ll showcase how other TTS may be integrated using an API by example of a server-based TTS system (open-source MaryTTS: <http://mary.dfki.de>) which can be used under both Android and iOS operation systems.

4. New Research Applications

The app has already been utilized to carry out research in its natural domain of Human-Computer Interaction (a study on acceptance of a) paper recommendations and b) push notifications; both still to be published). The most recent study by Moore on potential future events relating to progress in the

speech technology domain [3] shows that speech technology applications on mobile devices are expected to take an important role in the near future. As initially planned, we foresee a number of applications scenarios for the app, which might facilitate and stimulate research on spoken language technology. Examples are:

- *Integration of different TTS systems:* For this purpose, we would like to agree on a standard interface facilitating research on mobile TTS usage. In case that a number of different TTS systems can be made available by the community, we propose to integrate a rating functionality so that feedback on perceived quality can be collected online from conference participants.
- *Integration of ASR systems:* Similarly, ASR developers could test their systems under realistic background noise conditions. Again, a standard interface has to be defined. We propose to record the speech data in the app and forward it to a common server, so that a speech database with realistic speech and background noise can be collected and made available to the community.
- *Role of speech in multimodal interaction:* As better speech input and output capabilities are integrated into the app as complements to touch-based input and graphical output, it would be interesting to analyze modality choice and preferences in the field, potentially also categorizing different usage contexts.
- *Collection of speech and noise data:* By integrating a recording functionality, participants of Interspeech could be solicited to record speech data with pre-defined or self-selected linguistic content through the app, potentially with different affective or other paralinguistic shaping. For this purpose, it would be helpful to trace the device and acoustic user interface characteristics, so that the recording characteristics could be estimated accordingly.
- *Content analyses:* Analyses of scientific topics and their interrelationship could be made in the long run, by building graphs of papers which are linked by common interests (e.g. paper presentations attended by the same person). The contents and their relationships could be rated by conference participants, e.g. through paper or poster ratings (potentially leading to attendees’ best poster prize).
- *Dissemination and scientific discussion:* The app could also be used for stimulating the discussion amongst attending and non-attending members of the community. The 2015 Interspeech features a first experiment with “virtual” (e.g. remote audio-visual) presentations, which could be integrated into the app in the future. The app could also stimulate the discussion on papers through in-app commenting and question-answering.

5. References

- [1] R. Schleicher, T. Westermann, J. Li, M. Lawitschka, B. Mateev, R. Reichmuth, and S. Möller, “Design of a Mobile App for Interspeech Conferences: Towards an Open Tool for the Spoken Language Community”, Proc. Interspeech 2013, 775-777, 2013.
- [2] A.I. Niculescu, H.Y. Kheng, L.F. D’Haro, K. Seokhwan, J. Ridong, R.E. Banchs, “Design and evaluation of a conversational agent for the touristic domain”, Proc. APSIPA 2014, 1-10, 2014.
- [3] R. K. Moore, “Progress and Prospects for Speech Technology: Results from Three Sexennial Surveys”, Proc. Interspeech 2011, 1533-1536, 2011.