March 01, 2019

Gesture-to-Speak

MingHung Hsieh

Gesture-to-Speak

Abstract:

This document describes an easy-to-use communication system that converts a received gesture into audible, spoken words and phrases. The communication system includes a mobile device having a gesture-to-speak manager application that associates gestures received through an interface of the mobile device to audio files containing pre-recorded phrases. The gesture-to-speak manager application directs the mobile device to play the audio files, as associated to the received gestures, through an audio output of the mobile device.

Keywords:

gesture recognition, speaking aid, speech aid, speech software, voice recorder, augmented and alternative (AAC) communication device, word card, speech-generation device

Background:

In some instances, a person may have limited speaking capabilities due to one of a number of physiological reasons, such as stroke aphasia, downs syndrome, or autism. In other instances, another person may have limited speaking capabilities due to his unfamiliarity with pronunciation of words and phrases of a language that is other than his native language. Today, in instances such as these, a user may utilize an assisted-speaking communication tool such as a voice recorder, a picture book with images, a word card, or an augmentative and alternative (AAC) communication device to help him communicate.

Such an assisted-speaking communication tool is often burdensome to the person. The AAC communication device, for example and in some instances, may be a dedicated computing device, which is expensive. In other instances, an application that enables a non-dedicated

computing device to perform AAC functions may limit flexibility by constraining the person to using icons or symbols from a symbol library[1]. Other assisted-speaking communication tools, such as the voice recorder, the picture book with images, and the word card each have limited capabilities to adapt to changes in communication needs that the person might experience. Assisted-speaking communication tools today are neither as efficient nor as effective as needed.

Detailed Description:

Figure 1 illustrates an example mobile device having gesture-to-speak capabilities.

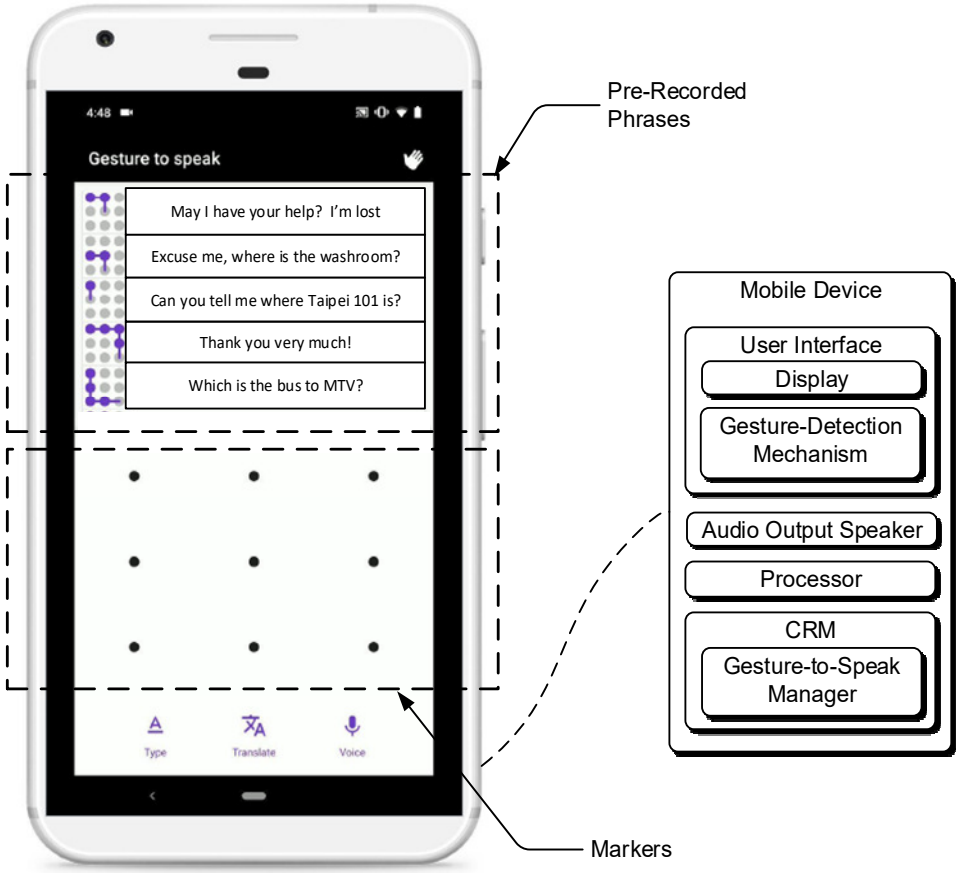

Figure 1

[1] https://ablenetinc.com/soundingboard

The example mobile device has an interface that includes a display having a gesture-detection mechanism. The example mobile device also includes an audio output speaker, a processor, and a computer-readable media (CRM) having a gesture-to-speak manager application. As illustrated, the processor of the mobile device is executing the gesture-to-speak manager application, directing the mobile device to operate in a gesture-to-speak active mode.

A first area of the display of the mobile device is presenting a textual representation of pre-recorded phrases. Adjacent to each pre-recorded phrase, the first area of the display presents a corresponding pattern that serves as a key for a user to select, through a gesture, one of the pre-recorded phrases. In some instances, the textual representation of the pre-recorded phrases may include textual representation of the pre-recorded phrases in multiple languages and/or scripts associated with the multiple languages.

A second area of the display of the mobile device is presenting an array of markers that combine with a touchscreen capability of the display (*e.g.*, a touchscreen capability based in a capacitive, a resistive, a reflective, or a grid-interruption technology) to form the gesture-detection mechanism. Through the user "tracing" a pattern through the presented array of markers in the second area of the display, the mobile device (*e.g.*, the processor executing the gesture-to-speak manager application) detects a gesture that corresponds to a pre-recorded phrase. In response to detecting the gesture that corresponds to the pre-recorded phrase, the mobile device plays, through the audio output speaker of the mobile device, the pre-recorded phrase.

Figure 2 illustrates an example technique performed by the mobile device in accordance with one or more aspects of gesture-to-speak. The technique includes the processor of the mobile device executing the gesture-to-speak manager application to operate in a gesture-to-speak configuration mode.

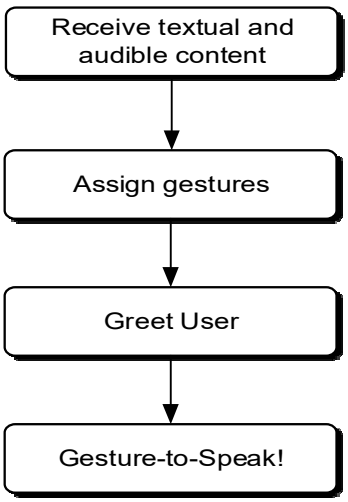

Figure 2

Receiving the textual and audible content may rely on a combination of techniques, as illustrated by Figure 3:

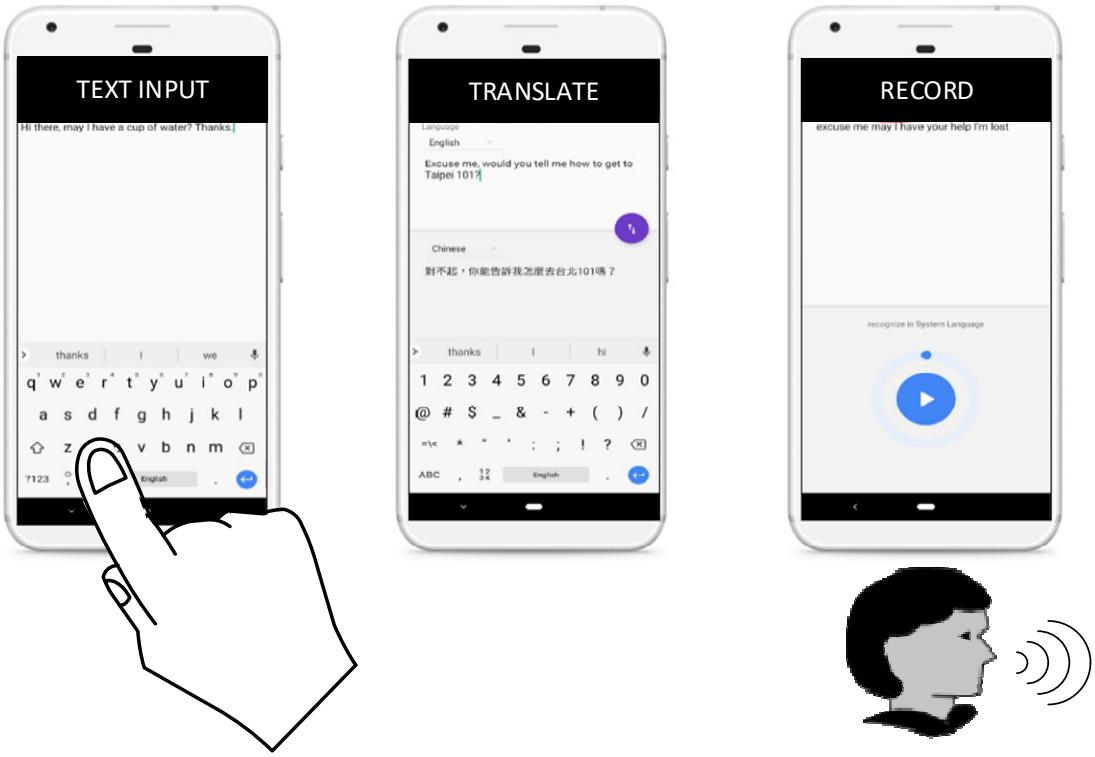

Figure 3

Figure 3 illustrates example techniques that include receiving the textual content (*e.g.*, content that is used by the mobile device to associate a textual representation to pre-recorded phrases) through keyed inputs from the user of the mobile device as well as translating the textual content via a translation application. Receiving the audible content, as illustrated by the example techniques of Figure 3, includes recording audible, spoken phrases from the user.

Assigning gestures to the received textual and audible content may rely on techniques as illustrated by Figure 4:

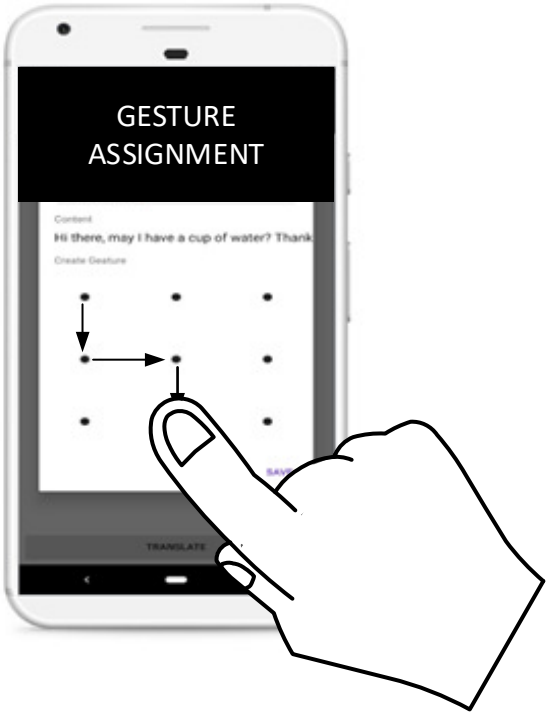

Figure 4

Figure 4 illustrates example techniques that include the mobile device presenting, to the user, an array of markers that the user can use to provide inputs that correspond to a pattern. The array of markers is spaced such that wide degrees of skill (or motor-functions) associated with the user providing the inputs are accommodated. In some instances, the user may use his finger to

trace the pattern using a sequence of the markers, while in other instances the user may use a stylus or other mechanism to trace the pattern. And yet, in other instances, the mobile device may present a suggested pattern for gesture assignment that the user may accept or decline.

Entering the gesture-to-speak active mode may rely on techniques as illustrated by Fig. 5:

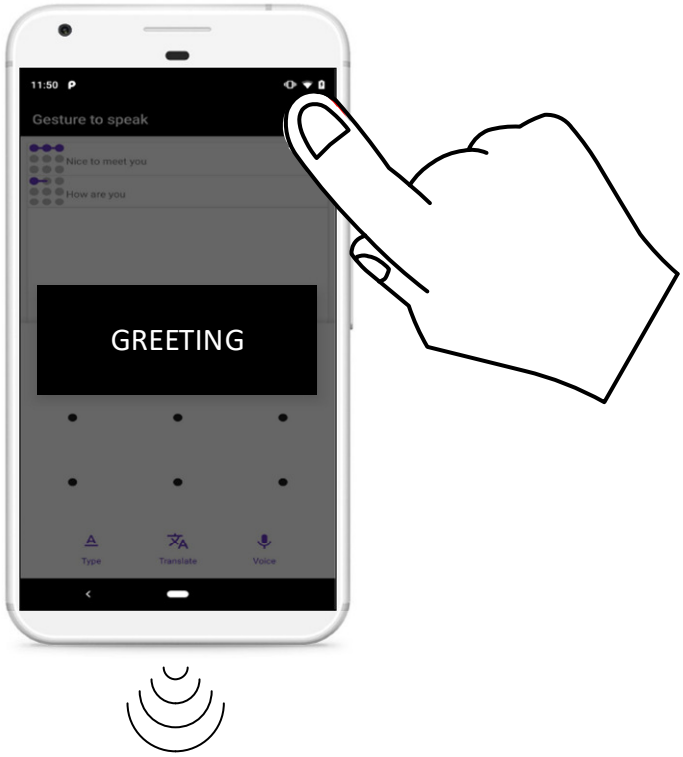

Figure 5

Figure 5 illustrates example techniques that include the user selecting an icon to initiate the gesture-to-speak active mode. In response, the mobile device provides a greeting to the user that may be textual or audible, and may offer basic instructions to the user for proceeding in the gesture-to-speak active mode.

The aforementioned systems and techniques are not limited to the examples described herein, and are extendable to encompass additional permutations and variations. As a first example, and in contrast to the gesture-to-speak manager application being local to the device, a

cloud-based service provider may offer portions or functionalities of the gesture-to-speak manager application. As a second example, the mobile device may receive textual content or audible content from another service provider as opposed to the user. And, as a third example, the gesture-detection mechanism of the mobile device may be a visual gesture-detection mechanism or a radar-based gesture-detection mechanism.

The aforementioned example systems and techniques, directed to the gesture-to-speak manager application that is part of the mobile device, demonstrate improvements over communication tools used for assisted-speaking today. These improvements include overcoming the need for a dedicated and expensive communication tool (such as the AAC communication device) as well as introducing configuration capabilities to meet the user's particular communication needs.