An Enriched Emoji Picker to Improve Accessibility in Mobile Communications

Maria Teresa Paratore[1[0000-0002-9089-8445]], Maria Claudia Buzzi[2[0000-0001-7818-0601]], Marina Buzzi[2[0000-0003-1725-9433]] and Barbara Leporini[1[0000-0003-2469-9648]]

[1] ISTI – CNR, via Moruzzi 1, 56124 Pisa, Italy
[2] IIT – CNR, via Moruzzi 1, 56124 Pisa, Italy

mariateresa.paratore@isti.cnr.it, claudia.buzzi@iit.cnr.it, marina.buzzi@iit.cnr.it, barbara.leporini@isti.cnr.it

Abstract.

We present an emoji picker designed to enrich emojis selection on mobile devices using audio cues. The aim is to make emojis selection more intuitive by better identify their meanings. Unlike the typical emoji input components currently in use (known as “pickers"), in our component each emotion-related item is represented by both an emoji and a non-verbal vocal cue, and it is displayed according to a two-dimensional model suggesting the pleasantness and intensity of the emotion itself. The component was embedded in an Android app in order to exploit touchscreen interaction together with audio cues to ease the selection process by using more than one channel (visual and auditory). Since the component adds non-visual information that drives the emoji selection, it may be particularly useful for users with visual impairments. In order to investigate the feasibility of the approach and the acceptability/usability of the emoji picker component, a preliminary remote evaluation test involving both sighted and visually impaired users was performed. Analysis of the data collected through the evaluation test shows that all the participants, whether sighted or visually impaired, rated the usability of our picker as good, and also evaluated positively the model adopted to add semantic value to emojis.

Keywords: Emojis, Accessibility, Inclusive design, Visual impairments, Audio affect bursts, Circumplex model of affect, mobile communication

1 Introduction

Computer-mediated communication, i.e., communication delivered through networks of computers (in short, CMC), is now pervasive and affects our daily social interactions, especially in text-based formats, both synchronous (instant-messaging applications) and asynchronous (email, social media, blogs) [3]. This phenomenon has dramatically grown worldwide during the recent COVID-19 pandemic, when nationally imposed lockdowns and forced self-isolation have led to a massive adoption of social media and messaging applications to replace face-to-face interactions. Not only business meetings and school lessons [29, 30] have been held in virtual mode, but also social relations with friends and relatives have been carried out via electronic devices on a regular basis.

In these exceptional circumstances, CMC has fully shown its potential, becoming the safest (and sometimes the only possible) communication channel among people located in different places. Thanks to CMC, co-workers have been able to exchange and comment documents from different locations and store copies on their personal devices, while friends could chat through virtual groups via messaging applications such as WhatsApp. Although the term "Mobile Mediated Communication" may be found in the literature to identify interactions that occur via smartphones or tablets, for simplicity's sake we will adopt the term "CMC" in these cases as well.

1.1 Adding Empathy to CMC: Emoticons and Emojis

Despite the aforementioned advantages, text-based CMC presents a major shortcoming compared to face-to-face communication; the lack of non-verbal cues, such as facial expressions or different tones of voice, creates a sort of psychological distance and makes it less "empathic" and more prone to misunderstandings [6, 8].

Since the time when emails became a widely adopted communication medium, users have strived to add socioemotional cues to their messages. Examples of such a behavior are the use of capitalized words to simulate a loud tone of voice, or typing nonverbal interjections such as "ehm", "oh-oh", etc.

The way users manage timing in asynchronous CMC may also convey socioemotional hints to messages; a reply may in fact be perceived differently depending on its promptness, suggesting a different degree of intimacy/liking or dominance/submissiveness with respect to the sender [3, 9]. Emoticons are one of the earliest expedients adopted to enrich text-based communication, ever since the spread of emails and SMS text-messaging, and are still in use nowadays.

An emoticon is formed by a sequence of text characters (typically punctuation and symbols) that represent facial expressions when viewed sideways (e.g. :-) to describe a smiling face). In the literature, many studies confirm the function of emoticons as effective substitutes of non-verbal cues. Derks et al. [6] examined the role of emoticons in synchronous textual CMC (chat sessions) with respect to the social context (socio-emotional vs task-oriented). Their findings show that emoticons are mainly used in socioemotional contexts (i.e., with friends), in both positive and negative situations, in accordance with what happens in face-to-face communication. Moreover, Walther and D'Addario [8] found that when communicating via email (i.e., asynchronously), emoticons add a significant bias to the interpretation of messages when associated with a negative emotional valence (e.g. :-("sad face", >:-("angry face").

Emojis are single-character pictographs used to add non-verbal emotional cues, and can be considered an evolution of emoticons [7]. Unlike emoticons, which are not subjected to any standard or supervision, they are standardized according to an encoding maintained by the Unicode Consortium [11].

At the time of writing, the Unicode standard (release 13.1) [12] accounted for 3521 emojis, according to the classification shown in Table 1.

Table 1. The number of emojis according to official Unicode classification release 13.1 [12].

Smileys & Emotions	People & Body	Com-po-nents	Animals & Nature	Food & Drink	Travel & Places	Activi-ties	Objects	Sym-bols	Flags	Total
156	2049	9	140	129	215	84	250	220	269	3521

Emojis are more expressive than emoticons, since they provide a graphical rendering. However, their rendering depends on platforms and applications, which may cause misunderstandings. Miller et al. [23] explored whether emoji renderings or differences across platforms give rise to diverse interpretations of emojis. To this aim, they distributed an online survey to solicit people's interpretations of a sample of the most popular emojis, each rendered for multiple platforms. The survey was completed by 304 users and the variance in interpretation of the emojis, quantifying which emojis were most (and least) likely to be misinterpreted, which were analyzed both in terms of sentiment and semantics. Results showed that in many cases, when two people consider the same emoji rendering, they may interpret both the sentiment and semantic meaning differently. Disagreement about both sentiment and semantic interpretations across different platforms was also highlighted.

According to Cramer et al. [10], the contextual meaning of a particular emoji or sequence of emojis can be quite difficult to untangle, since the linguistic function for both sender and recipient is flexible and open to interpretation. Emojis can depict facial expressions, but also pictorial representations of objects, symbols and actions, and their intended meanings can go far beyond their definition according to the Unicode standard. They suggest that when interpreting emojis, it is important to detect the sender's intent, before 'translating' [10]. Semantic ambiguities may arise from cultural, interpersonal and social differences between sender and receiver [7, 13, 14]. The variety of emojis usages and interpretations will depend on specific social practices and norms, e.g. a private long-standing group of friends may send messages containing emojis with meanings not interpretable without an intimate knowledge of their shared history. Volkel et al. [7] assert that emojis may be misunderstood since their intended meaning often remains ambiguous. They present the results of an online large survey (N=646) showing that personality traits influence the choice of emojis, and that an emoji's choice and interpretation varies greatly between users, even if it is presented in a defined message context. They also infer that emojis interpretations differ depending on whether users are in a public (Twitter, blog posts, etc.) or private context (text chat).

Barbieri et al. [13] explore the meaning and usage of emojis across a number of languages and dialects, also comparing the language-specific models of each emoji. Results suggest that although the overall semantics of the most frequent emojis are similar, some emojis are interpreted in a different way from language to language, and this could be related to socio-geographical differences. Lu et al. [14] analyze how smartphone users adopt emojis based on a very large data set collected from a popular emoji keyboard, extracting a complete month of emoji use by 3.88 million active users from 212 countries and regions, in order to compare user behaviors and preferences across countries and cultures. Their findings show that categories and frequencies of the emojis used provide rich signals for the identification and the understanding of cultural differences among smartphone users.

Emojis’ ambiguities are also confirmed in more recent works. Alismail and Zhang [24] conducted a study focused on the perception, interpretation, and liking of a facial emoji-based Likert scale as a means of evaluation assessment in online surveys; their research highlights different experiences of usage among participants.

Herring and Dainas [25] performed an online survey to assess how English-speaking social media users interpreted the pragmatic functions of emojis in examples adapted from public Facebook comments. Their findings show that while female and male interpretations of emojis are generally similar, there are differences in the appreciation and understanding of their usage depending on age, older males being the least likely to use and appreciate emojis, the opposite of the younger females.

1.2 Emojis and Accessibility for Users with Visual Impairments

Tigwell et al. [4] conducted a survey to find out the problems faced by people with visual impairments in relation to emojis. Their findings highlight the issue that assistive technologies, instead of facilitating the inclusion of emojis in CMC, may complicate it, thus being an obstacle to social inclusion.

Screen readers translate emojis into textual descriptions. When a screen reader user receives a message, they perceive an emoji as extra text embedded in the current conversation, which may become very annoying, especially if it contains many emojis. Moreover, the text that is read aloud by the screen reader for an emoji may be a source of misinterpretation itself, since the associated description may not correspond to the emotion that the sender meant to convey.

Searching for an emoji to embed into a message as it is composed emerges as one of the most challenging tasks. Fig. 1 shows two screenshots of the emoji picker provided by the Twitter client for Android; the picker comprises a panel containing some controls and a grid of selectable items. The left screenshot shows how emojis are presented according to the frequency of use, while the right one shows the grid for the 'smileys and people' category. In both cases, choosing an emoji via an assistive technology may be cumbersome, given the number of items the user has to navigate. Moreover, default descriptions associated with pictographs may be misleading, if not cryptic, and as a result, it may be difficult to understand whether a chosen emoji actually conveys the desired meaning.

A good example of such ambiguity is given by the emoji whose Unicode name is "Face with Look of Triumph", which was originally meant to convey positive emotions associated with pride and personal satisfaction [4]. In spite of its Unicode definition and original meaning, the emoji is currently used on many platforms to suggest negative feelings, such as frustration, anger or contempt. As a result, a recipient might be puzzled on how to interpret this character in a message; should they intend it as a "cool" expression 😎 (described as a “smiling face with sunglasses” according to the Unicode specifications), or rather as an "angry" one 😠 (described as an “angry face”)?

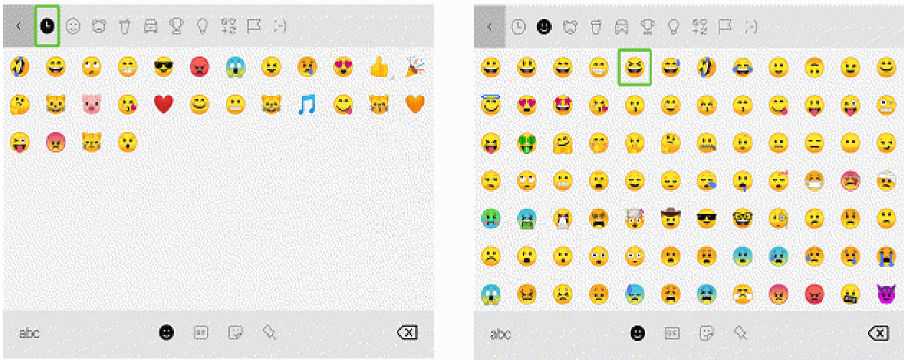

Fig. 1. Two screenshots of the emoji picker from the Twitter client for Android.

The redundancy of symbols used to express essentially the same feeling is another problem. Table 2 shows a number of emojis currently in use to express happiness. Different graphic representations for the same emotion are meant to convey intensity or qualitative differences, but they are not very simple to differentiate and can be confusing for users with visual impairments.

Table 2. Windows 10 rendering, Unicode encoding and Unicode standard textual description of "smiley" emojis

Windows 10 Rendering	Encoding	Description
	U+1F642	slightly smiling face
	U+1F60A	smiling face with smiling eyes
	U+263A	smiling face
	U+1F603	grinning face with big eyes
	U+1F600	grinning face
	U+1F601	beaming face with smiling eyes

2 Theoretical Background

We have highlighted that emoticons, and later emojis, originated from the need to enrich textual CMC with emotional cues. In this section, we will provide an overview of psychological models related to emotions.

2.1 Basic Emotions Theories and Models

Many different theories exist that identify basic emotions [1]. Plutchick [2] devised eight primary emotions as an evolutionary development of the human being; anger, disgust, fear, sadness, anticipation, joy, surprise, trust. Primary emotions are strictly related to the survival of the human species. Plutchick later developed a circular graphic scheme called the 'wheel of emotions', a two-dimensional model that relates different emotions and attitudes to each other by means of spatial distribution and colors. The model, shown in Fig. 2, highlights how primary emotions are pairwise opposites; joy vs. sadness, anger vs. fear, trust vs. disgust, and surprise vs. anticipation.

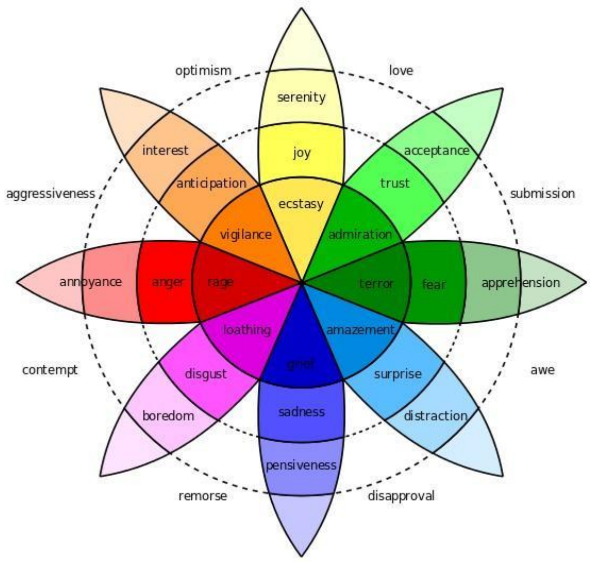

Fig. 2. The Plutchick Wheel of Emotions.

Ekman & Friesen [28] identified a different set of basic emotions related to non-verbal communication. Their studies on facial expressions led to the definition of six "modal" emotions (anger, disgust, fear, sadness, surprise, happiness), associated with well-defined facial expressions, which constitute the "Ekman faces"; a validated set of grayscale photographs that portrait different actors' faces expressing the modal emotions, plus a neutral expression. The number and quality of Ekman's basic emotions are subject to debate [18], although the Ekman faces are widely adopted in research as a valid tool to assess emotions, primarily due to the fact that Ekman faces refer to cross-cultural non-verbal stimuli.

According to Russell et al. [1, 2, 18], a state of mind can always be described in terms of a linear combination of two neuropsychological states, the valence (the level of pleasure/displeasure) and the arousal, or alertness. This model is well-known as the

"Circumplex Model of Affect", and introduces the concept of "core affect" as a consciously recognizable feeling in a certain moment. Based on Russell's theory, core affects are placed in a two-dimensional space, where the X-axis represents the pleasantness and the Y-axis the arousal experienced in a certain state of mind. Yik et al. [2] identify 12 such core emotional components, evenly distributed as 12 sectors on a circular surface. Their model is referred to as the 12-Point Affect Circumplex, and is shown in Fig. 4 (Sect 3.1).

2.2 Theories of Emotions and Auditory Perception

Ekman & Frieser's theories have mainly been criticized [15, 18] because many emotions are not referable to a characteristic facial expression; moreover the same facial expressions can be related to different emotions (e.g., a smile may be used to convey happiness as well as commiseration or sarcasm).

Alongside the analysis of emotions based on facial expressions, research has been conducted in the field of auditory affective processing, focused on the role of nonverbal interjections in manifestations of emotional feelings [19, 20]. The results of these studies essentially highlight that human beings can infer emotions from nonverbal vocal cues exactly as occurs for the Ekman faces (especially in a cross-cultural way). Belin et al. [15], in particular, propose an auditive alternative to the Ekman faces, in which facial expressions are substituted by "affect bursts" (i.e., eight short non-verbal vocal interjections). According to their study, eight basic emotions (anger, disgust, fear, pain, sadness, surprise, happiness, and sensual pleasure) can be associated with the same number of nonverbal bursts. Participants (15 of them male, 15 female), were asked to evaluate an initial set of 198 vocalizations reproduced by actors, each related to a basic emotion. At the end of the evaluation process, a validated set of audio samples was obtained, which constitutes the so-called Montreal Affective Voices (MAV). MAV is composed of 90 audio samples, i.e., the registered vocalizations of ten actors (5 males and 5 females), each portraying the eight basic emotions, plus a "neutral" expression.

Cowen et al. [16] adopted a different approach to classify emotions in terms of vocal bursts. These authors refined the two-dimensional model based on valence and arousal, and theorized a multi-dimensional semantic space where each emotion is obtained as a combination of emotion categories. Categories do not belong to a discrete set; on the contrary, they blend into each other with smooth gradients that determine a continuous semantic difference. In order to infer the dimensionality of the space, a large-scale, multi-stage experiment was conducted, at the end of which 24 semantic dimensions were identified: adoration, amusement, anger, awe, confusion, contempt, contentment, desire, disappointment, disgust, distress, ecstasy, elation, embarrassment, fear, interest, pain, realization, relief, sadness, surprise (negative), surprise (positive), sympathy, triumph. As in [15], in this study as well a possible limitation may arise from cultural bias. Actually, while vocalizations were gathered from participants from four different countries, only US-English speaking users participated in the task of recognizing emotions. The result of the experiment is available within an online interactive map [17], in which the 2032 vocal bursts analyzed in the study are clustered according to the identified semantic components.

3 Designing a Novel Emoji Picker Based on "Real" Emotions

As we previously highlighted, standardized Unicode emojis bear intrinsic semantic ambiguities and accessibility issues due to their graphical nature, which may hinder the quality of social relationships via CMC. This is even more true if the sender and/or the recipient of a message is a person with visual impairments. Picking one or more emojis from the standard set while typing a message may prove to be an overwhelming task, given the enormous number of items available. In order to address these shortcomings, we designed an emoji picker that guides the user through the choice of an emoji thanks to auditory cues and an arrangement of the pictographs on the screen oriented by their emotional components. In order to test the emoji picker, we integrated it in an Android app; one of the authors, who is totally blind, actively collaborated on the interface design, and on the prototype evaluation throughout all phases of development. Moreover, intermediate versions of the emoji picker were periodically published as beta tests on the Google Play Store, so that other users could give us feedback. Beta testers were recruited through specialized email groups and comprised both sighted users and users with visual impairments who relied on Android's screen reader, namely TalkBack.

3.1 Arranging Emotions on the Screen

In order to make the browsing functionality more intuitive, we associated the action of sliding one's finger across the screen with a coherent change in the emojis' semantic value. This solution was straightforward once we decided to adopt Russell's circumplex model, in a similar way as proposed in [5]. The graphical user interface (GUI) of the emoji picker was modelled upon a circular shape, similar to a wheel divided into sectors, where each sector corresponds to an emotion. Thanks to the underlying semantic structure of Russell's model, users can adjust the characteristics of the emotion they want to convey by sliding the finger on the screen of the mobile device. Fig. 3 shows how the movement of the finger on the GUI is related to the emoji that is going to be selected.

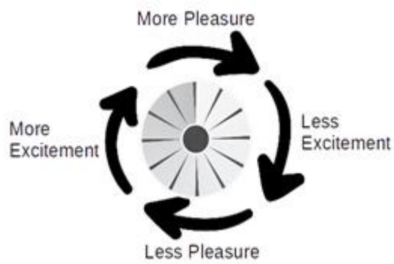

Fig. 3. Sliding the finger across the GUI helps find emojis according to degree of pleasure and excitement.

Our design strategy aimed to create a compromise between a fair range of emotions and a narrow space, since our emoji picker was expected to be mostly used on medium-sized smartphones. Thus, one of our main design requirements was to achieve a good compromise between expressiveness and usability. Prototypes of wheel-shaped GUIs existing in the literature [32, 33] rely on the Plutchick model and the Geneva Emotion Wheel [34], which also account for the intensity of emotions. In our case, due to the small size of the screen, implementing a wheel of reduced size with concentric sectors would result in very poor user interaction, especially when selecting emojis close to the center of the wheel. A very small movement of the finger in that area would in fact result in hearing a confusing rapid succession of different vocal bursts. If the interaction were performed by a person with visual impairment, especially via screen reader, the emoji picker would actually become not usable.

Concerning the number of emotions to be considered, in spite of the many categories existing nowadays, the most frequently used emojis in messaging apps and social media are those related to feelings and emotions, in accordance with the primary function of these special characters [21]. Unicode statistics [35] also show that only a very limited subset of emojis classified as "smileys and emotions" appears within the first ten positions of the most frequently used. This encouraged us to adopt the 12-Point Circumplex Model of Affect (12-PAC) by Yik, Steiger and Russell [12, 22], which presents an adequate granularity of emotions for our purpose.

Fig. 4 shows the 12-PAC model and our rendering of it as a wheel with 12 sectors associated to emojis. In order to associate each core emotion with a pictogram, we referred to the aforementioned statistics and our personal experiences with social media, instant messaging apps and emails. While some associations are immediate (e.g., Scared, Disgusted), others may be ambiguous (Excited, Satisfied), but we were confident that audio hints would have conveyed clearer information.

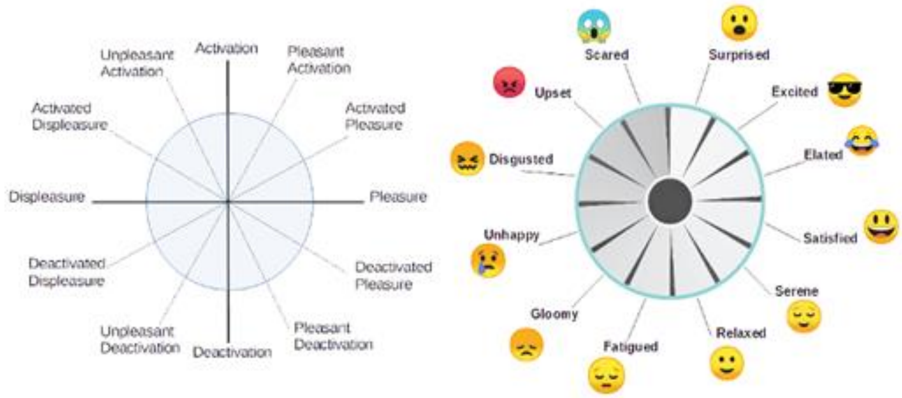

Fig. 4. The 12-PAC model and how it is rendered in our emoji picker.

3.2 Associating Affective Vocal Bursts

We adopted non-verbal emotional hints from the MAV set [15], integrated with vocal bursts provided by [16], in order to reduce semantic ambiguity and further improve accessibility. For each emotion sector in our wheel, a validated auditory, non-verbal counterpart was associated.

Our circumplex-based model maps the following emotional states; surprised, excited, elated, satisfied, serene, relaxed, fatigued, gloomy, unhappy, disgusted, upset, scared. MAV model [15], on the other hand, only accounts for eight emotions, i.e., anger, disgust, fear, pain, sadness, surprise, happiness, and sensual pleasure. We hence chose two sets of eight MAV audio files, one per gender, and inferred the four missing emotions from [16] using the online interactive map provided [17]. Given the wide range of nuances available, we used the interactive map to identify the missing vocal cues according to the percentage composition of emotional components.

Table 3 shows the four missing emotional bursts (one per gender) chosen from the map (Excited, Elated, Fatigued, Gloomy) and their emotional components.

Table 3. The vocal bursts chosen from the model proposed by Cowen et al. [16, 17]

Gender	Emotional Burst	Composition
M	Excited	84% Awe, 8% Surprise (positive), 8% Sympathy
M	Elated	53% Elation, 25% Triumph, 8% Contempt, 8% Awe, 8% Ecstasy, 8% Pride
M	Fatigued	35% Pain, 17% Distress, 8% Confusion, 8% Desire, 8% Elation, 8% Embarrassment, 8% Fear, 8% Neutral
M	Gloomy	58% Sadness, 18% Pain, 8% Contentment, 8% Disappointment, 8% Distress
F	Excited	50% Elation, 26% Triumph, 8% Contempt, 8% Ecstasy, 8% Pride
F	Elated	67% Elation, 17% Triumph, 8% Confusion, 8% Contentment

F	Fatigued	34% Distress,17% Confusion, 17% Embarrass-ment, 8% Contempt, 8% Desire,8% Disappoint-ment, 8% Sadness
F	Gloomy	29% Distress, 21% Embarrassment, 8% Confu-sion, 10% Disappointment, 8% Disgust, 8% Fear, 8% Pain, 4% Contempt, 4% Elation

4 Testing the emoji picker

As previously mentioned, in order to test the emoji picker, we integrated it in an Android app. The app's distribution was provided worldwide in English and Italian. In order to address aspects related to accessibility and the rendering of audio hints, we actively involved our co-author, who is blind, in the design phase along with two volunteers with visual impairments recruited via a specialized Facebook group. Their suggestions and feedback were important in defining the interactions of our emoji picker with Android's assistive technologies.

After a first phase of internal testing within our research group, the Android app with the emoji picker was released for evaluation as an "open test app" on the Google Play Store. Participants to the evaluation phase were recruited via social media and mailing lists; invitations to participate were posted on Facebook and email groups focused on accessible technologies for users with a visual impairment. Overall, 30 testers were recruited, 12 of whom were visually impaired and regular TalkBack users.

The test was divided into two phases; in the first part, the users were presented with the consent form to participate in the study and then were introduced to the emoji picker with a detailed description of how it works. At the same time, they were asked to customize the emoji picker by choosing a gender for the voice that will utter the affect bursts. Finally, the emoji picker was displayed and users were invited to practice with it (see Fig.5) in order to become more familiar with the different interaction modalities.

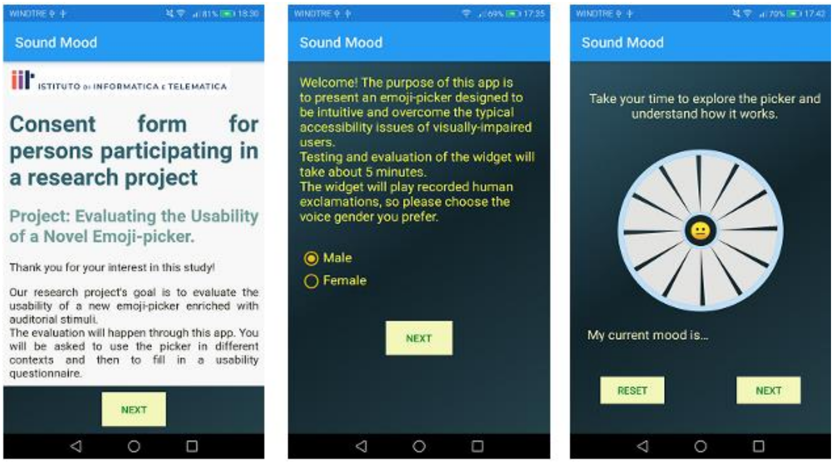

Fig. 5. Screenshots of the app: first part of the evaluation process.

Once users had completed the phase of familiarization with the emoji picker, they were asked to complete three sentences by adding an emoji as if they were using an instant messaging app. Each sentence was conceived to fit with an emoji from a different quadrant as shown in Fig.6, so that test users were encouraged to explore most of the wheel to complete the test. In the case that an assistive technology was in use, each sentence was read aloud by the screen reader as soon as it was presented to the user; while dragging with the finger on the GUI, only emotional audio bursts were heard to announce the underlying emojis. As the user chose an emotion, the focus returned to the sentence and the screen reader read again the whole sentence and reproduced the emoji in terms of its pictogram's description.

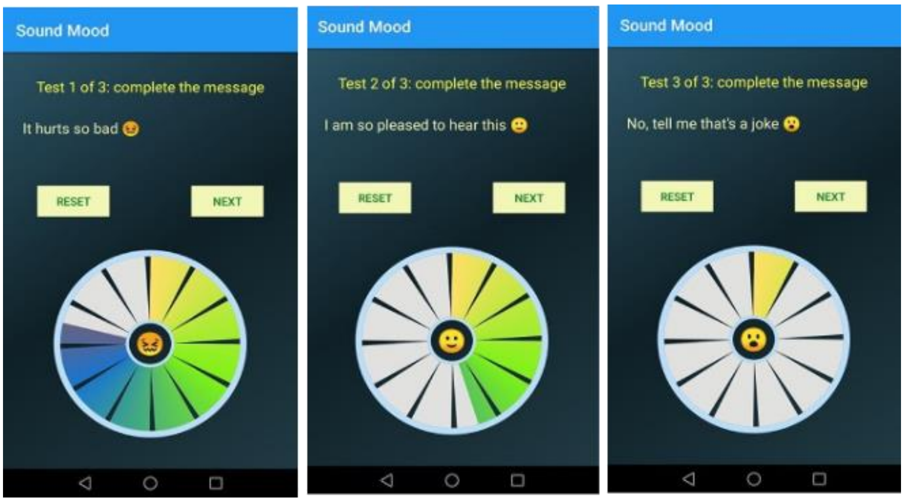

Fig. 6. Screenshots of the app: second part of the evaluation process (the simulated chat messages).

After the test, users were invited to fill out an evaluation questionnaire, which was anonymized via a unique identifier. The goal of the questionnaire was to assess the usability of the emoji picker and the validity of our model, hence we adopted the ten-question System Usability Scale (SUS) [23, 24], to which three questions were added, related to the model:

- The position of the emotions on the screen helped me find what I wanted to pick (Q1)
- The audio cues (exclamations) helped me find what I wanted to pick (Q2)
- The audio cues (exclamations) gave a good description of the emotions (Q3)

All the questions used 5-point Likert items, answered in the same way as the SUS, from Strongly Disagree (1) to Strongly Agree (5). To complete the questionnaire, testers were finally asked to provide their gender, age, level of education and if they had visual impairments.

5 Results

Once all the questionnaires were completed, we checked data for inconsistencies and consequently eliminated six questionnaires in which the answers showed obvious inconsistencies. SUS scores [26, 27, 29] were then evaluated for each of the 24 remaining participants (ten of whom with visual impairments).

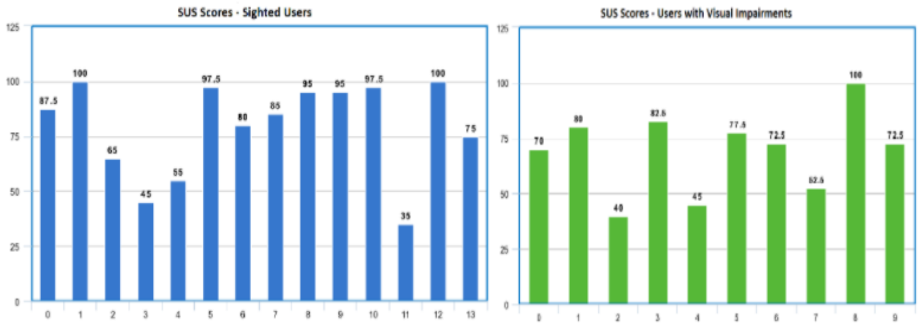

Figure 7: SUS scores collected for each participant in our study.

Table 4. Means and standard deviations for SUS for sighted and visually impaired participants

Evaluation Group	M	SD
Sighted (N=14)	79.46	21.6
Visually impaired (N=10)	69.25	18.41

SUS values for both evaluation groups are described in Figure 7. In the sighted group there are some high values (8 higher than 80/100) but also three values below 60/100.

In the visually impaired group there are only two values higher than 80/100 and three values below 60/100; anyway, the overall analysis of the results shows a substantial equivalence between sighted and visually impaired participants in their SUS scores ($t = -1.21$, $df = 22$, n.s.).

The ratings on the three questions about the model were very skewed towards the "Strongly Agree" end of the scale, so non-parametric statistics, as shown in Table 5, were appropriate.

Table 5. Median and interquartile range (IQR) for ratings on the three questions related to the model for sighted and visually impaired participants, with one sample Wilcoxon Signed Rank test results

Evaluation Group	Median	IQR	W	Probability
Sighted (N=14)				
Q1: Position of emotions	4.5	1.0	81.5	0.009
Q2: Auditory cues helped	5.0	1.0	99.0	0.002
Q3: Auditory cues good description	4.0	1.0	67.0	0.021
Visually impaired (N=10)				
Q1: Position of emotions	4.0	1.0	42.0	0.018
Q2: Auditory cues helped	4.0	1.0	37.5	0.068
Q3: Auditory cues good description	4.0	2.0	48.0	0.031

Comparing the ratings with the midpoint of the scale (neither agree or disagree) both sighted and visually impaired participants rated all three aspects significantly above the midpoint (although Q2 on whether auditory cues helped just failed to reach the 0.05 level of significance for visually impaired participants, but there was a strong trend towards significance). Thus both sighted and visually impaired participants were positive about these three aspects of the emoji picker.

There were no significant differences in the ratings on the three aspects between sighted and visually impaired participants (Mann-Whitney U test, Q1: $U = 77.5$, n.s.; Q2: $U = 92.0$, n.s.; Q3: $U = 56.0$, n.s.).

In relation to the participants' personal characteristics there was a significant difference between men and women in their ratings of Q1, but not the other two questions (see Table 6). There were no significant differences due to age, but the number of participants in each age group was probably not sufficient for a robust analysis.

Table 6. Median and interquartile range (IQR) for ratings on the three questions related to the model for men and women participants, with Mann-Whitney U test results

	Men (N=16)		Women (N=7)			
	Median	IQR	Median	IQR	U	Probability
Q1: Position of emotions	4.0	1.0	5.0	0.0	24.0	0.033
Q2: Auditory cues helped	4.0	2.0	5.0	1.0	36.5	n.s.
Q3: Auditory cues good description	4.0	1.0	4.0	1.0	47.5	n.s.

6 Discussion and Conclusions

We have presented an emoji picker for mood and emotions based on a model inferred from theories of emotions. The aim of the emoji picker is to provide users with spatial and auditory hints related to the semantics of emotions, in order to overcome the typical problems of ambiguity and misinterpretation posed by the choice of an emoji. At the same time, our picker is aimed at overcoming the accessibility issues met by users with visual impairments when they wish to express feelings through emojis. The emoji picker relies on 12 basic emotions (the 12-PAC emotion model) and on a validated set of emotional vocal hints (the Montreal Affective Voices, plus four audio cues selected from among those identified by Cowen et al.). The 12-PAC model enabled us to arrange the emojis based on a semantic structure, while the purpose of the audio cues was to provide immediate, non-ambiguous descriptions of the emotions, especially helpful when accessing the emoji picker via a screen reader. While exploring the wheel by touch, audio feedback that reproduces a non-verbal description of emotions is triggered and immediately perceived, so that users do not need to listen to a long and possibly ambiguous description spoken by a voice synthesizer.

An evaluation study was conducted via an Android app, which was completed by 24 participants, ten of whom have visual impairments. The app presented the emoji picker, proposed a series of short tests and finally administered a questionnaire to assess the SUS score and an evaluation of the model. Values obtained for the SUS score are encouraging, since the results collected showed that both sighted and visually impaired participants rated the emoji picker's usability as good. The spatial and auditory model's evaluation was also positive for both categories of users, even though the ratings suggest that improvements must be considered, for users with visual impairments, in finding more effective techniques to convey audio hints. Future efforts in the improvement of the emoji picker will be mainly devoted to finding effective solutions to express the emotions' intensity and add more customization options (e.g., personalized associations between emotions and vocal bursts).

In order to gather more test users, we are planning to develop a prototype of our tool for the iOS platform. This will enable us to make a comparison between different assistive technologies (i.e., TalkBack vs VoiceOver).

Finally, we are aware that ambiguities in the interpretation of emojis may persist on the recipient's side, but the design of a software component such as a "recipient's translating layer" was beyond the scope of this study. However, this topic will be a subject of our research in the near future.

7 Acknowledgments

We wish to thank the following email groups devoted to accessibility and users with visual impairments, who showed particular interest in our study and helped us find participants: tech-vi@groups.io, accessible@googlemail.com. We wish to also thank the Blind & Visually Impaired Android Users on Facebook.

References

1. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology, 39, 1161-1178 (1980)
2. Yik, M., Russell, J. A., & Steiger, J. H.: A 12-point circumplex structure of core affect. Emotion, 11(4), 705–731 (2011)
3. Walther, J., & Tidwell, L.C.: Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. J. Organ. Comput., 5, 355-378 (1995)
4. Tigwell, G.W., Gorman, B.M., & Menzies, R.: Emoji Accessibility for Visually Impaired People. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (2020)
5. Toet, A., & Erp, J.V.: The EmojiGrid as a Tool to Assess Experienced and Perceived Emotions. Psychology, 1, 469-481 (2019)
6. Derks, D., Bos, A.E., & Grumbkow, J.V.: Emoticons in Computer-Mediated Communication: Social Motives and Social Context. Cyberpsychology & behavior: the impact of the Internet, multimedia and virtual reality on behavior and society, 11 1, 99-101 (2008)
7. Völkel, S.T., Buschek, D., Pranjic, J., & Hußmann, H.: Understanding Emoji Interpretation through User Personality and Message Context. Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services (2019)
8. Walther, J., & D'Addario, K.P.: The Impacts of Emoticons on Message Interpretation in Computer-Mediated Communication. Social Science Computer Review, 19, 324 – 347 (2001)
9. Liebman, N., & Gergle, D.: It's (Not) Simply a Matter of Time: The Relationship Between CMC Cues and Interpersonal Affinity. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (2016)
10. Cramer, H., Juan, P.D., & Tetreault, J.: Sender-intended functions of emojis in US messaging. Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (2016)
11. Unicode Inc. 2020. Unicode® Full Emoji List, v13.1, https://www.unicode.org/reports/tr51/ Accessed: 2020-10-12
12. http://unicode.org/emoji/charts/full-emoji-list.html
13. Barbieri, F., Kruszewski, G., Ronzano, F., & Saggion, H.: How Cosmopolitan Are Emojis? Exploring Emojis Usage and Meaning over Different Languages with Distributional Semantics. Proceedings of the 24th ACM international conference on Multimedia (2016)

14. Lu, X., Ai, W., Liu, X., Li, Q., Wang, N., Huang, G., & Mei, Q.: Learning from the ubiquitous language: an empirical analysis of emoji usage of smartphone users. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. (2016)
15. Belin, P., Fillion-Bilodeau, S., & Gosselin, F.: The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. Behavior Research Methods, 40, 531-539 (2008)
16. Cowen, A.S., Elfenbein, H.A., Laukka, P., & Keltner, D.: Mapping 24 emotions conveyed by brief human vocalization. The American psychologist. (2018)
17. https://s3-us-west-1.amazonaws.com/vocs/map.html#
18. Posner, J.D., Russell, J., & Peterson, B.: The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and psychopathology, 17 3, 715-34. (2005)
19. Banse, R., & Scherer, K.: Acoustic profiles in vocal emotion expression. Journal of personality and social psychology, 70 3, 614-36. (1996)
20. Schröder, M.: Experimental study of affect bursts. Speech Commun. 40, 99-116. (2003).
21. Updated Emoji Statistics: https://www.emojipedia.org/stats/
22. Zhong, K., Qiao, T., & Zhang, L.: A Study of Emotional Communication of Emoticon Based on Russell's Circumplex Model of Affect. HCI. (2019)
23. Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I.L., Terveen, L., & Hecht, B.J.: "Blissfully Happy" or "Ready toFight": Varying Interpretations of Emoji. *ICWSM*. (2016)
24. Alismail, S., & Zhang, H.: Exploring and Understanding Participants' Perceptions of Facial Emoji Likert Scales in Online Surveys. ACM Transactions on Social Computing, 3, 1 - 12. (2020).
25. Herring, S., & Dainas, A.: Gender and Age Influences on Interpretation of Emoji Functions. *ACM Transactions on Social Computing, 3*, 1 - 26. (2020)
26. Brooke, J.: "SUS: a "quick and dirty" usability scale". In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (eds.). Usability Evaluation in Industry. London: Taylor and Francis. (1986)
27. Brooke, J B.: SUS — a retrospective. Journal of Usability Studies, Vol. 8, Issue 2, February 2013 pp. 29–40. (2013)
28. Ekman, P. & Friesen, W. V. Constants Across Cultures in the Face and Emotion. *Journal of Personality and Social Psychology, 17(2)*, 124-129. (1971)
29. Bangor, A., Kortum, P.T., & Miller, J.T. Determining what individual SUS scores mean: adding an adjective rating scale. Journal of Usability Studies archive, 4, 114-123. (2009)
30. Sirait, E.R., & Zellatifanny, C.M. An Empirical Study: Computer-Mediated Communication and Collaboration among Government Employees during Flexible Working Arrangements. 2020 International Conference on Information Technology Systems and Innovation (ICITSI), 95-100 (2020)
31. Mirnaalini. V. Determining the effect of motivation on computer mediated communication during COVID-19. International Journal of Indian Psychology, 8(2), 86-92. (2020)
32. Runge, N., Hellmeier, M., Wenig, D., & Malaka, R. Tag your emotions: a novel mobile user interface for annotating images with emotions. Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct. (2016)
33. Warpechowski, K., Orzeszek, D., & Nielek, R. Tagging emotions using a wheel user interface. Proceedings of the 13th Biannual Conference of the Italian SIGCHI Chapter: Designing the next interaction. (2019)
34. Sacharin, V., Schlegel, K., & Scherer, K. Geneva Emotion Wheel Rating Study. https://archive-ouverte.unige.ch/unige:97849 (2012).
35. Frequency of emoji use: https://home.unicode.org/emoji/emoji-frequency/