Enhancing Non-Visual Emoji Accessibility: Inclusive Design with Blind and Low-Vision Co-Researchers

by Mingran Yang

Submitted to OCAD University in partial fulfillment of the requirements for the degree of Master of Design in Inclusive Design
Toronto, Ontario, Canada, 2025

CREATIVE COMMONS COPYRIGHT NOTICE

ACKNOWLEDGMENTS

Dr. Peter Coppin, Principal Advisor

I would like to express my deepest gratitude to my principal advisor for your invaluable guidance and support throughout my research journey in the Inclusive Design (INCD) program. This project has been both challenging and rewarding, and your mentorship played a crucial role in helping me navigate unfamiliar territory. Thank you for guiding me with your knowledge in the fields of perceptual-cognitive research, blindness and low vision, cross-sensory design, and semiotics. Thank you for providing all the resources and instructions. Your encouragement, thoughtful feedback, and unwavering patience gave me the confidence to persevere and ultimately complete my Major Research Project. Thank you for being a constant source of support and inspiration.

David Barter, Secondary Advisor

Thank you for witnessing the development of my emoji project from Lab 1 through to the final semester. Your feedback after each class has been incredibly valuable and deeply appreciated. And thank you for all your help with refining the wording and clarity of my writing. I am also deeply grateful for the support, guidance, and encouragement you provided, which made it possible for me to carry out my research.

Dr. Richard Hunt, Secondary Advisor

I have always been passionate about continuing my studies in graphic design, and it has been a pleasure to have your guidance throughout my master's journey. I will never forget the two books I borrowed from your office library. Thank you for supporting and guiding me. Your insights were instrumental in helping me find direction and establish a clear timeline for my Major Research Project

TABLE OF CONTENTS

ABSTRACT

Emojis present accessibility challenges for blind and low-vision (BLV) individuals, particularly when navigating through screen reader interfaces. Interpretations of emojis have been extensively researched, highlighting the influence of cultural differences on individual comprehension, the need to proactively avert potential misunderstandings, to explore the denotations and connotations of emojis, and the use of emojis in online communication. This paper explores the challenges of emoji interaction for blind and low-vision (BLV) users within the field of Human-Computer Interaction (HCI), focusing on the absence of visual cues and the resulting ambiguity in emoji affordance when using screen readers. This study examines communication barriers in digital environments, highlighting the need for user-centered design within Information and Communication Technologies (ICTs). It proposes strategies to improve emoji accessibility and non-visual communication experiences for BLV users, drawing on insights from semi-structured interviews, co-design sessions, and user testing.

Keywords: Emoji Accessibility; Blind and Low Vision (BLV); Screen Readers; User-Centered Design; Inclusive Design; Accessible Communication; Non-Visual Interaction; Human-Computer Interaction (HCI)

1. INTRODUCTION

1.1 Background

Emojis have become a core part of modern digital communication, enriching text with emotion, tone, and visual nuance. However, their interpretation heavily relies on visual cues, which creates significant accessibility challenges for blind and low-vision (BLV) users who depend on screen readers. However, for these users, these visual elements are often translated into standardized textual descriptions, leading to a loss of emotional nuance, intent, and cultural meaning. As a result, BLV users may experience gaps in shared understanding, missed social signals, and increased cognitive load when navigating emoji-laden content.

Previous Research in inclusive design and human-computer interaction highlights the importance of preserving the multi-modal affordances of digital symbols to ensure equitable access. Zong et al. (2022) demonstrate how to enhance screen reader users' experiences with rich, interactive elements to support blind users' access to complex visual information. Yet, there is a lack of user-centred research exploring how BLV individuals interpret and interact with emojis—and what alternative modes (e.g., sonification, tactile cues, or improved metadata) could better support their experience. Compared to virtual chat environments, in-person communication offers richer access to nonverbal cues such as facial expressions and body language, which are often lost or diminished in digital interactions. This gap becomes particularly evident in platforms such as Zoom or Microsoft Teams, where emojis are frequently used as quick emotional feedback—such as reacting to a message in the chat or briefly appearing on screen as a reaction to a speaker (e.g., thumbs-up, clapping, laughing)—yet these visual cues are often inaccessible or only partially conveyed to blind and low-vision users through screen readers. For blind and low-vision users, this creates an uneven communicative landscape—where others receive real-time social cues visually, however they experience delays, ambiguity, or even total absence of that information.

This research project explores the accessibility challenges BLV users face when interacting with emojis through screen readers. It investigates the semantic and emotional gaps in current screen reader interpretations, how users navigate these limitations, and what affordances are lost in translation. The purpose is to identify the current challenges in emoji communication for BLV users and to develop alternative design strategies for a more inclusive experience. Through interviews and co-design sessions with BLV participants, the research will explore how emoji interpretation differs from that of sighted users, identify affordance mismatches in non-visual contexts, and propose alternative design ideas, such as sound cues, that can better preserve emojis' expressive functions.

1.2 Research question and objectives

This MRP seeks to investigate how blind and low-vision (BLV) users experience, interpret, and navigate emoji-based communication through screen readers. It aims to identify the limitations of current accessibility practices and explore alternative, non-visual modalities that can better preserve the emotional, social, and cultural dimensions of emoji use. By focusing on the screen reader experience of BLV individuals, the study contributes to more inclusive design strategies for digital communication tools. This investigation began with the following research question:

- How do blind and low-vision users interpret the meaning of emojis in screen reader environments?
- What is the affordance of emojis in the context of the experience of sighted individuals and screen reader users?
- What are the differences in emoji interpretation between sighted and BLV users, and how do these affect digital communication clarity and emotional nuance?

1.3 Limitations and scope

The limitations of this study were scoped to be achievable within the limited timeline and resources of a Master's-level major research project. These constraints affected both the theoretical breadth and the number of participants involved in interviews and co-design sessions. Emojis are highly context-dependent and culturally variable. This study could not address the full breadth of symbolic interpretation across different communities or languages. The number of blind and low-vision (BLV) participants was small, which may limit the generalizability of the findings. Additionally, participants may not represent the full spectrum of screen reader users in terms of age, sex, tech familiarity, or cultural background. The user feedback on the experimental sound cues indicated that the addition of sound could be distracting or perceived as "noise." The study was limited to further develop an optimized multisensory solution due to time frame constraints.

.

2. DEFINITIONS

2.1 Blind & Low-vision

This research project focuses on individuals who are blind or have low vision (BLV). The term 'blind and low-vision' (BLV) is used as an umbrella term for anyone who identifies as blind or has low vision. Rather than defining blind and low-vision individuals solely by visual impairment, it is essential to

recognize their identity as members of a diverse sensory culture, where "blindness is not a lack of vision, but a different way of being in the world" (Kleege, 2018). Visual impairment encompasses various levels of vision loss, including blindness in one or both eyes, as well as conditions such as macular degeneration, cataracts, glaucoma, and diabetic retinopathy (World Health Organization, 2019). Additionally, there are distinctions between congenital blindness and acquired blindness. Most BLV individuals retain some degree of vision, with complete blindness being relatively rare. However, there remains a gap in understanding how both blind and low vision individuals engage with emojis through screen readers.

In keeping with the principles of inclusive design, design can be used to create equal and accessible opportunities for BLV individuals. Dr. Kenneth Jernigan, former president of the National Federation of the Blind (NFB) from 1968 to 1986, defined blindness as follows: 'One is blind to the extent that the individual must devise alternative techniques to efficiently accomplish tasks that would otherwise be performed with normal vision. A person may be considered blind when the need for alternative techniques substantially alters their daily living' (Jernigan, 2005).

This definition aligns with the criteria used by the Canadian National Institute of the Blind (CNIB), which states: 'You do not need to be legally blind to access services from the CNIB Foundation, Vision Loss Rehabilitation Canada, or other CNIB organizations. As soon as your eyesight begins to affect your daily life, you're eligible' (CNIB, 2022).

2.1.1 Spatial-topological synchrony

3. RATIONALE

The hypothesis is that among sighted people, the meaning of emojis affords something other than text descriptions alone. This research acknowledges the critical role emojis play in communication and responds by proposing design recommendations to enhance interactive and communicative experiences for BLV users. These recommendations are grounded in observations of participants' lived experiences with text-described emojis and developed through co-designing new emoji interactions with them. Sergeant (2019) notes that emojis are powerful tools for conveying emotion, nuance, and social context, much like gestures, facial expressions, and tone in face-to-face interactions. Emojis have evolved to carry rich cultural and subcultural meanings that are not reflected in their official descriptions. Previous researchers have identified significant mismatches between intended meanings and received interpretations (Miller et al., 2016). This evolution has led to a growing knowledge gap between sighted users and screen reader users—one that not only limits access to important contextual

information but also leaves many users unaware that such information is missing or inaccessible through current screen reader technologies.

Emojis have clearly evolved into a distinct form of visual “language” (Francesco 2016). Yet, the specific factors driving this evolution remain insufficiently understood. Their meanings are often shaped by users' cultural backgrounds and social contexts, which can give rise to new metaphors and interpretations. As Francesco (2016) notes, individuals may interpret the same emoji in different ways. For instance, the 🙏 emoji—commonly interpreted by sighted users as a symbol for prayer—is sometimes labelled as a "high-five" in accessibility metadata, leading to potential confusion for screen reader users (Guntuku, 2019). This kind of interpretation highlights the “conceptual ambiguity” (Coppin, 2014) afforded by emojis. This can become a problem because when blind people who rely on screen readers only receive audio descriptions of emojis provided by accessibility experts (or AI), they are receiving a single interpretation that does not convey the conceptual ambiguity that the emoji is affording for sighted computer users (Barter and Coppin, 2022). This leads me to further speculate that BLV individuals may face additional barriers and challenges when using text emoji in online communication. What is the affordance of emojis in the context of the experience of sighted individuals and screen reader users? What are the differences in emoji interpretation between sighted and BLV users, and how do these affect digital communication clarity and emotional nuance? How can alternative sensory modalities (e.g., sound cues) be designed to improve the accessibility and interpretability of emojis for blind and low-vision users?

4. RELATED WORK

4.1 Cultural Influence

Several studies have investigated the diversity of emoji and their distinct interpretations across cultures. These studies suggest that emojis in online communication tend to be understood in significant, often vastly different, ways depending on the interpreting user’s cultural background, language, and social experiences. Even individuals from similar backgrounds may interpret the same emoji differently due to subtle socio-geographic variations (Francesco, 2016).

Emojis evolved in response to significant communicative needs in online communication, and have become extremely popular and diverse as a result (Grosz et al., 2023). The official labels and descriptions of emojis are determined by the Unicode Consortium—a private, non-profit organization primarily composed of representatives from major technology companies—which holds exclusive authority over which emojis are encoded, how they are named, and what they represent (Miltner, 2021). However, these standardized descriptions often diverge from culturally embedded interpretations,

resulting in potential misunderstandings, particularly in cross-cultural communication. As Haralambous and Dürst (2019) note, “As all Unicode characters, emojis have names such as grinning face, rocket, or zombie. Nevertheless, users are not necessarily aware of names: they choose emojis only according to their shape, and thus attach their own meaning to each emoji.” This disconnect between the official labelling and user perception underscores the drift that can occur when visual symbols are interpreted outside their standardized framework.

The interpretation of emojis is influenced by cultural connotations, social context, individual experiences, and even irony, leading to potential miscommunication. Abercrombie’s (1968) concept of "paralinguistic" gestures offers a useful framework for understanding emojis as non-verbal cues that require contextual interpretation. By treating emojis as a paralinguistic modality, their meaning can be better understood within the context of the conversation.

4.2 Paralinguistic Cues

Emojis contribute to written communication by conveying emotions that are typically expressed through prosodic features in spoken language—such as rhythm, tone, pitch, and stress. As Alnuzaili et al. (2024) observe, "the observed construction is encoded with different placements of emojis and emojis are displaying diverse manifestations to exhibit prosodic features in Computer-Mediated Communication (CMC) such as intonation, rhythm, duration, pause and stress." These features allow emojis to express a range of emotional states, including "aggression, love, hate, sadness and joy." In face-to-face or verbal conversations, people rely on vocal cues such as intonation, pauses, and emphasis to convey feelings such as sarcasm, excitement, or sadness. However, in written online communication, these vocal cues are absent, which might lead to misunderstandings or a lack of emotional context. Previous scholars have noted that emojis can carry the emotional weight in written communication that is typically conveyed through prosody in spoken language (Wagner, 2016; Zappavigna & Zhao, 2017). As Cole (2015) explains, “Prosody is often described in terms of intonation and rhythm – the musical qualities of speech. Intonation and rhythm (or, more generally, timing) are suprasegmental aspects of speech because they define patterns that are largely independent of the segmental makeup (i.e., the consonant and vowel phones) of a given word or phrase.” Similarly, Wagner (2016) and Zappavigna and Zhao (2017) emphasize that prosody in spoken language reveals emotional states and speaker attitudes—for example, indicating whether someone is happy, sad, critical, or condescending. In digital communication, where prosody is absent, emojis often fill this gap. As Gesselman, Ta, and Garcia (2020) note, “the presence of emojis significantly influenced recipients’ perceptions of emotional tone, often shifting otherwise neutral messages into positively or negatively

valenced interpretations." For instance, when a "smiley face" emoji accompanies a neutral message such as "Could you send this file? 😊", it alters the perceived tone, making it seem friendlier or more polite. In this way, emojis function as substitutes for the emotional and tonal cues naturally conveyed through speech.

There are more factors to consider when trying to determine the influence of emoji affordances. The interpretation of emojis is influenced by cultural connotations, social context, individual experiences, and even irony, leading to potential miscommunication. Abercrombie's (1968) concept of "paralinguistic" gestures offers a useful framework for understanding emojis as non-verbal cues that require contextual interpretation. By treating emojis as a paralinguistic modality, their meaning can be better understood within the context of the conversation. In earlier research, Trager (1958) gave the term 'paralinguistic' a detailed definition: "If the interaction includes speech, the voices may be loud or soft, harsh or warm, slow or fast. Also audible may be sniffs, clicks, coughs, laughs, hesitation sounds such as 'uh,' inhalations, and sighs. Vocal actions of this sort have been termed 'paralinguistic'". The presence of paralinguistic supports human conversation by conveying non-verbal expressions, making one's emotions and attitudes perceptible to others. These perceptions help structure speech, guide the flow of personal interactions, receive instant feedback, etc. In digital communications, emojis have limited expressive power absent of a visual graphic format, for example when conveyed by screen readers, they are presented as a robotic voice that communicates only words and not the aforementioned rich non-linguistic vocal actions of paralinguistic cues and gestures.

4.3 Screen Reader

A screen reader is an assistive technology that allows individuals with little or no functional vision to interact with digital content by converting on-screen information into speech or Braille output (Leporini, 2022). This technology plays a vital role in promoting digital accessibility, enabling users with visual impairments to navigate websites, read documents, and engage with various applications, thus fostering greater independence and inclusivity. However, screen reader users often encounter specific challenges when navigating emoji lists, which can be time-consuming and confusing (Zhang, 2021). These difficulties are further exacerbated by inconsistencies in emoji descriptions across different operating systems, such as iOS and Android, leading to fragmented and sometimes misleading user experiences. As Tigwell (2020) and Griggio (2024) highlight, a major accessibility concern lies in the disconnect between the visual design of emojis and the textual descriptions provided by screen readers. Beyond these technical limitations, research has shown that the interpretation of emojis is highly variable across cultural contexts, shaped by users' linguistic

backgrounds, cultural norms, and social experiences (Francesco, 2016). Even individuals from similar regions may interpret the same emoji differently due to subtle socio-geographic nuances.

Screen readers also include emoji accessibility features for BLV users, though this functionality highlights several key challenges. For instance, Zhang (2021) notes that when an emoji appears multiple times on a screen, screen readers produce a repetitive audio experience, making it difficult for users to comprehend the content. Tigwell (2020) and Griggio (2024) further explain that discrepancies between visual emoji designs and their textual descriptions lead to varying interpretations by different screen reading technologies, causing accessibility issues. However, users still face challenges when navigating long emoji lists sequentially, which is time-consuming (Zhang, 2021). Tigwell's research (2020) demonstrates that BLV users struggle with the inconsistency between the visual design of emojis and their textual descriptions. Griggio (2024) expands on this by suggesting that different screen reading technologies may generate varying interpretations of the same emoji, further complicating accessibility.

Shifting focus from cultural misinterpretation of emojis to the barriers faced by visually impaired individuals, prior research highlights numerous challenges in online communication. Screen readers present significant challenges—not only due to the absence of alternative modalities but also because of medium-specific affordance limitations and a fundamental misunderstanding of the implications behind accessibility standards that enforce strictly textual representations. Building on Barter and Coppin's (2022) critique of text-based accessibility for diagrams, the researcher suggests that the current accessibility standards for screen readers, which depend on text-based descriptions of emojis, often fail to convey the nuances embedded in their visual design. As Barter and Coppin (2022) illustrate, "the current accessibility standards, which prioritize text-based descriptions, often fail to convey the nuanced meanings of iconographic emojis, leading to a loss of affordance for screen reader users." These limitations hinder effective communication between blind and low-vision users and their sighted counterparts, especially in contexts that rely heavily on visual cues like emojis. Although some efforts have been made to improve emoji accessibility, such as tactile elements or sound control technologies (Yuri, 2017; Zhang, 2021), challenges remain. Zhang et al. (2021) introduced "Voicemoji," a web application that enables visually impaired users to input emojis using voice commands. This approach leverages speech input to facilitate emoji entry, addressing the difficulties that screen reader users often face when navigating complex emoji lists. However, while such innovations offer promising solutions, they also highlight the need for broader systemic changes in how emojis are represented and accessed across different platforms and assistive technologies.

5. METHODOLOGY

The research participant group consists of five individuals with varying degrees of visual impairment and different experiences with vision loss, ensuring a range of perspectives. All participants are experienced users of assistive technologies and use screen readers for online communication on a daily basis. and were selected based on their access to compatible devices and software. Through interviews and screen reader testing sessions, I explore how participants navigate emoji interpretations and the challenges they face, aiming to inform the development of more inclusive emoji communication methods.

This study employed a qualitative, user-centred research design grounded in inclusive design and human-computer interaction (HCI) principles, with a specific focus on the experiences of blind and low-vision (BLV) users navigating emoji-based communication through screen readers. The study explores the full range of emoji affordances, both visual and non-visual, identifying where breakdowns occur and how these gaps might be addressed through alternative interaction design strategies. The researcher introduced additional sound cues to enhance interpretation for visually impaired users. Non-linguistic sonification, which uses alternative sensory channels to preserve iconic qualities, may help recover iconic information lost in text-based translation (Coppin, 2016). The study critically examines the limitations of existing screen reader descriptions, particularly the absence of conceptual ambiguity that sighted users access visually (Coppin, 2014). I developed my methodology by integrating insights from studies such as Image-Based Tactile Emojis: Improved Interpretation of Message Intention and Subtle Nuance for Visually Impaired Individuals (Park & Takagi, 2017) and Voicemoji: Emoji Entry Using Voice for Visually Impaired People (Lee et al., 2019), both of which emphasize multi-sensory and cross-modal strategies to improve emoji accessibility.

Drawing on these insights, the researcher conducted semi-structured interviews with five participants—three blind and two with low vision—who use screen readers on a daily basis. Each session lasted approximately 45–60 minutes and was conducted in person. Participants were first asked to describe their general experiences with screen readers, followed by a guided walkthrough of how they encounter and interpret emojis in messaging apps. Participants were asked about their frequent use of emojis and which ones they found confusing. We then invited them to interpret a curated set of emojis that go beyond basic objects or simple emotions like laughter, anger, or sadness. Instead, we focused on more complex or ambiguous emojis—such as 🥱 Yawning Face and 🤣 Rolling on the Floor Laughing—and asked participants to describe the meanings they associate with each and how they imagine these emojis might look visually. To see whether those emoji labels aligned with their

expectations or the emoji's intended expression. These sessions revealed that the affordances of emojis differ significantly between sighted users and screen reader users. Emojis, when experienced aurally, often lose the visual cues that support nuanced or context-specific meanings. Participants reported that descriptions were frequently unclear, inconsistent, or reductive, which contributes to a disconnect between visual and audio interpretations. The research highlights how the conceptual ambiguity (Coppin, 2014). that makes emojis expressive for sighted users becomes a barrier to communication when reduced to static, text-based labels in screen readers.

6. QUALITATIVE FINDINGS

6.1 Data Analysis

Throughout the inductive research, we found that sighted individuals primarily interpret the meaning of emojis through their visual appearance, which often makes emojis seem conceptually specific (Coppin, 2014). This visual clarity generally ensures that both the sender and receiver experience the same emoji, even in the face of technological inconsistencies, such as variations in emoji design across platforms such as iOS and Android, or issues introduced by screen readers. However, Tigwell (2020) notes that visually impaired users encounter challenges when there are discrepancies between an emoji's visual design and its textual description. Similarly, Griggio (2024) observes that different screen reading technologies can alter how emojis are presented and described, further complicating their interpretation. Nevertheless, even for sighted individuals, emojis can be conceptually ambiguous (Coppin, 2014), as their interpretation often depends on a user's cultural background, language, and social experiences (Francesco, 2016). This amplifies the impact of emoji ambiguity across different user experiences.

In the experience of screen reader users, the interpretation of emojis differs significantly from that of sighted individuals. Emojis become conceptually certain (Coppin, 2014) as BLV individuals must understand emojis based on the screen readers' audio descriptions/emoji official labels. Emojis themselves, when serving as engaging graphical elements to enhance text-based online communication, are inherently conceptually ambiguous (Coppin, 2014) when perceived as such due to their interpretive variability and contextual dependence. The theory is based on findings on the affordances of emojis for sighted people in online communication and further investigating these affordances in relation to the screen reader user experience.

In this study, researcher focus on scenarios in which the affordances of emojis — particularly those tied to their visual properties — are not perceivable when using screen readers. In other words, the affordances that arise from visual features are lost in translation, making the perceptual certainty of

emojis ambiguous, as the feelings and tones they convey to sighted users effectively disappear. While the official label represents the existence of emoji, it only carries the conceptual meaning. The textual description in the emoji labels is certain, therefore, we conclude that the information the screen reader gives about emojis—the text description—provides conceptual certainty and is perceptually ambiguous (Coppin, 2014), which is in contrast to the affordances of emojis for sighted people.

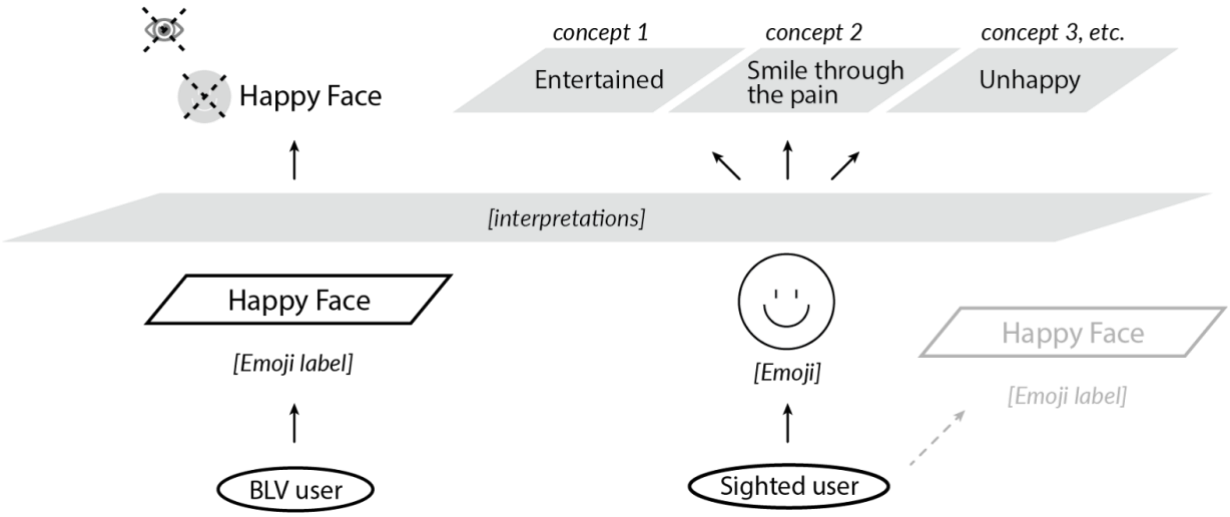

Figure 1: Comparative interpretation pathways of emoji for sighted and BLV users

In this model, the left side illustrates the interpretation process for BLV users from bottom to top. Since BLV users can only perceive the emoji's official label when interacting with emojis through a screen reader, their interpretations are limited to a single textual cue. However, the right side of the model which represents the interpretation process for sighted users, depicts a different scenario. Sighted users interpret emojis visually at first sight, drawing on contextual factors such as communication context, tone, familiarity with the sender, and cultural background. Although the emoji label technically exists for sighted users too, it is not easily accessible and is therefore rarely noticed or relied upon. Thus, the interpretation is not limited, sighted users often generate multiple interpretations that depend on individuals.

Previous research in the field of emoji and screen readers has been most invested in the variety of screen reading technologies that lead to different emoji descriptions (Griggio 2024) and understanding the gap between visual design and text description (Tigwell 2020). In exploring the disparity in emoji comprehension between sighted and visually impaired individuals, we found that the underlying

causes are largely rooted in differing affordances — and, in some cases, a relative absence of affordances. The following excerpts from the interviews illustrate these findings:

Participant 1 (P1) described a miscommunication when sending a grocery shopping request to his friend online through emojis as a type of shorthand. The miscommunication reveals how connotations of many emojis, when perceived visually, differ from the linguistic definitions of their labels. Fruit emojis, such as eggplant, peach, banana, and cherry emojis were interpreted as having meanings other than what could be interpreted via the text-label descriptions of the emojis. While the sexual meanings of these emojis are not conveyed through their descriptions, the disparity in knowledge of colloquial meanings between visually impaired and sighted individuals disrupts the "feedback loops" that typically support mutual understanding in communication. In online emoji-text interactions, this disruption can affect both the sender's intent and the receiver's interpretation, though the ultimate influence — whether clarifying or confusing — remains uncertain.

++pwc scratchpad

++

I don't know why people … [use emojis] …because they have so many meetings, I actually don't put fruit in my posts anymore …[because] so I got into trouble … I had no idea that certain vegetables mean, like certain things. Like various body parts …, if you're trying to have a conversation and you're literally saying, …, we're gonna go shopping. What do you need? And you would put down all the emojis. And next thing you know, someone's like, I hate you. And you're like, why? Well, like when you like, if you're gonna use the like they said, the eggplant, the peach, the banana. Even cherries. It's like everything has an alternative meaning. Yes, meaning that you know, in an entirely different subculture, … people, need to open up a thesaurus because they have no idea what some of the alternate words mean. I was sending it to my friend who was at the grocery store. And I said, hey, can you grab a couple of things? … It was a female …, so the response was, you know, understandable 'cause I already know who I'm sending it to. Right. But … let's say you have, like, a personal shopper, right? And you say, oh, also, can you get these and they reply back …[“What are you talking about?”] …Now, did you know that there was a double meaning behind those things? …It was also the way I typed it out because I was doing it on the base of searching. So I'm swiping through and … some emojis come up before others. Yeah. So why don't you use emojis instead of just typing the fruit? — Participant 1

Another challenge revealed for screen reader emoji users is the confusion caused by colour variations. For instance, heart emojis of different colours, such as red, blue, or purple, often lack clear or distinct descriptions of meanings outside the colour itself, making it difficult for users to interpret their specific meanings or emotional nuances. This ambiguity can lead to misunderstandings, as the colour meanings don't show in the audio description.

Well, if you send a heart, a Purple Heart, to someone, I can't remember what it means, but it has a meaning, and that person knows what that means, and it may be completely off base for what you want to say. — Participant 2

I find the emojis because there's so many of them, yeah, having gone through what some don't actually mean, because some of them have specific meanings, like the different colours of hearts and things. But A, I can't see the different colours when I'm doing it. And B, there's just so many of them. So I find that the world of emojis is just overwhelming. A simple smile or clap or a thumbs up is good enough for me. Cross side winking face with I like the green mancing and then Winky is very straightforward Winky things. Yeah, usually you could. You could put that with fun or sarcasm. I usually use the upside down face one. Yeah. What does that mean? It's like I don't understand. What the hell was description? Upside down, please just says upside down face. — Participant 3

6.2 Design Reflections from Lab Interviews with Visually Impaired Users

During the induction process, the study participants expressed concerns with my approach to introducing sound cues designed for screen readers to provide more context for emojis. This intervention failed because it neglected the importance of establishing a shared understanding between researcher and participants regarding emoji affordances. Without building a technical description of what an emoji affords, a hard rejection of this induction process followed. The experience highlights that the sighted researcher failed to anticipate the experiences of visually impaired screen reader users, and it underscores how unexamined assumptions about achieving shared understanding between researcher and participants can critically undermine accessibility efforts within inclusive design methodologies.

Emojis are represented as simple text labels by screen readers to describe the emoji that was used in conceptually specific terms, rather than the perceptually specific properties that sighted people perceive with their eyes (Coppin, 2014). The hypothesis is that adding sound cues to the screen reader in place of or in addition to the emoji text labels could enhance the user experience as it might provide sufficient cross-sensory redundancy to equate to near universal understanding between BLV and sighted individuals on emoji.

This concept mirrors accessible traffic lights, for which auditory cues guide pedestrians: pressing the button during a red light triggers a distinct 'wait' sound (like slow beeps), while the white/green light produces a different 'go' signal (such as rapid chirps). This audio feedback—with clear contrasts between the contexts of prohibition and permission to act—helps blind and low-vision users safely navigate crosswalks by distinguishing when to stop and when to cross.

This example was introduced for discussion by participants in interviews, resulting in the following response:

Where, when you're doing an emoji, yeah, it's like … another language, where you're trying to develop sounds that are going to match that, and not everybody's on the same speaking level. So it becomes a personal thing. Yes, whereas … sounds at the traffic lights as … you've used as a demonstration. They're unique. They've been set up specifically for that purpose. — Participant 2

In the natural lab setting, when we first introduced the idea of adding sounds to screen readers to enhance emoji accessibility, most participants described the sounds as mere noise. This reaction is perhaps unsurprising—asking users to imagine an unfamiliar concept is akin to challenging a group of 19th-century horse-drawn carriage drivers to envision an automobile before its invention in 1886. Not all participants agreed on the sound cue strategy. For instance, participant 4 noted:

Yeah, for me, I don't see any points. Having this kind of noise and like sounds, when I hear, see, like an emoji comes… why do we need that? As far as like, there's an indication that the emoji comes here, and then, like, I know … what kind of image it is, then I think it's fine for me. Can be distracting for me if there is a certain noise. And sometimes people, when texting put a lot of images, and then it's a lot of sounds comes in each time, and it's like too much … sounds and over, over, … I don't know why we need like sounds when it comes to emoji." — Participant 4

Participant 3 maintains a neutral perspective:

I'm always willing to try new things. But yeah, news to me, I'm still, I'm still working out the emojis, and to throw a third … component in I've, got the text or the gist of the text, you've got the emojis in place of words … or feelings or definitions, and then put music in as well. It's like, … sitting with my husband in the living room in the morning when the TVs go and he's old school and he's trying to talk to me at the same time. It's like, it's just too much information, or … depends on the tone of the music. It's … just that it sets a certain tone, then it can interfere with the delivery of the message. If like, somebody sent this emoji in a way that … doesn't fit in this context, if the sound is like, too uplifting to cheer for, [not sure if I want to hear that sound when I'm reading my text. — Participant 3

The sharing on sound cue strategy received feedback from participants with mixed reactions. Participant 4 expressed strong resistance and questioned the necessity of this design idea, describing them as potentially distracting and overwhelming, especially in text-heavy conversations. In contrast, Participant 3 maintained a neutral perspective, showing openness to experimentation, yet also expressing concerns about sensory overload and potential mismatched sound tones and message intent. A shared understanding of how emoji affordance functions across visual and non-visual experiences was missed by researcher to share before the co-design session. This led to the co-design session ending.

7. DISCUSSION

7.1 Effective Emojis and User Experience

All participants reported that they primarily use simple emojis to convey basic emotions. In response to the question, "Which emoji do you feel is most effectively described by its label, and why?", participants indicated that emojis representing clear emotional expressions or tangible objects were the most effective. Examples cited include:

– 😂 Face with tears of joy ("the floor laughing")
– 🙂 Smiling face
– 🥲 Smiling face with tear ("smiling face crying eyes" / "big tears")
– 🥱 Yawning face ("I'm bored")
– 😘 Face blowing a kiss ("kisses kissing")
– 😠 Angry face
– 🎂 Birthday cake ("peace birthday cake")
– 🎄 Christmas Tree

Interview findings reveal that users' interpretations of emojis in online communication are often shaped by their familiarity with the recipient. BLV individuals, in particular, tend to avoid using ambiguous emojis in order to minimize the risk of miscommunication or being misunderstood. Emojis that convey less common or more nuanced emotions, such as the 🙃Upside-Down Face—often lack clear meaning on what they represent when only presented through screen reader descriptions. This lack of clarity can lead to confusion and place the burden of interpretation entirely on the user, increasing the likelihood of communicative gaps.

When it comes to screen reader users' user experience, screen readers visualize each emoji based on how many times it appears in a message and read it multiple times. This repetition of audio labels could bring cognitive overload and frustration for users.

7.2 Ambiguity and Accessibility in Emoji Representation

The interviews revealed that many emojis carry complex or layered emotional meanings, such as expressing both happiness and sadness simultaneously. This confusion may stem from the discrepancy between the visual appearance of certain emojis and their textual descriptions and colour variations. Participants also noted that a single feeling could often be represented by multiple emojis with little difference in description, leading to ambiguity in interpretation, such as 😊 (smiling face); 😀 (Grinning Face); 😄 (Grinning Face with Smiling Eyes); 😊 (Smiling Face with Smiling Eyes) etc. Additionally, the meanings or descriptions provided by screen readers were frequently described as strange, unclear, or disconnected from the visual graphics of the emojis. This disconnect creates a gap in understanding, particularly for screen reader users who rely solely on audio descriptions.

When a screen reader reads a text message containing an emoji, it conveys only the textual label without any emotional tone or expression. As a result, screen reader users interpret the emotional meaning of the emoji independently, based solely on the label. This creates a gap in understanding between sighted individuals—who rely on visual cues to interpret emojis—and screen reader users, who receive no such visual or tonal information.

Emojis with colloquial or sexual connotations often lack labels that indicate these meanings, which limits users' ability to understand their intended use or social context. As Participant 1 experience illustrates, emojis such as the eggplant, peach, or banana may have implicit meanings that differ significantly from their literal definitions. Screen reader users rely on textual descriptions to understand emojis, but these hidden connotations remain inaccessible because they are not shown in the emoji label. This creates a disconnection between senders and receivers using screen readers, highlighting a critical challenge in inclusive digital communication.

7.3 Shared Understanding in Emoji Accessibility for BLV Users

Shifting focus from cultural misinterpretation of emojis to the barriers faced by BLV individuals, prior research highlights numerous challenges in online communication. The use of screen readers presents unique obstacles when communicating with sighted individuals. Although some efforts have been made to improve emoji accessibility, such as tactile elements or sound control technologies (Yuri, 2017; Zhang, 2021), there remains a significant gap in designing tools that address the needs of BLV users.

In an effort to reinforce the intended message in emoji-text-based conversation, this research introduced sound cues as an additional non-visual modality. The study investigates the challenges of improving emoji accessibility for blind and low-vision (BLV) users by critically examining how such non-visual affordances—particularly auditory cues—are interpreted and experienced. Prior research

has identified a paralinguistic substrate in the use of emoticons within text-based online communication platforms (Tantawi & Rosson, 2019), suggesting that affective and contextual nuances can be conveyed beyond literal textual content.

This inductive research situates itself within the broader discourse on inclusive design failures, where assumptions made by sighted researcher can result in misaligned interventions. The study highlights the gap between perceptual affordances (what emojis look like for sighted users) and conceptual descriptions (what they mean when heard through screen readers), arguing that the lack of shared understanding between designers and users can undermine well-intentioned innovations.

7.4 Emoji Affordance Findings

Sighted individuals primarily interpret emojis through their visual appearance; although they are visually clear, part of their affordance lies in their conceptual ambiguity (Coppin, 2014). Drawing on the "paralinguistic" gestures concept, emojis function as non-verbal cues that rely heavily on contextual interpretation (Abercrombie, 1968). Interpretation remains fluid and context-dependent, shaped by cultural background, language, and social norms (Francesco, 2016). As Coppin et al. (2016) explain, pictorial representations operate on the analogue side of the analogue–digital spectrum, capable of being "processed and interpreted for identification under multiple conceptual categories." This flexibility contributes to the conceptual ambiguity (Coppin, 2014) of emojis, whose meaning often depends on contextual and cultural interpretation rather than fixed symbolic representative attributes. In contrast, screen reader users access emojis through standardized textual labels, which offer conceptual specificity but strip away the perceptual richness and interpretive nuance present inherent in visual representations (Coppin, 2014). Without visual cues such as colour, facial expression and stylistic variation, there is a semiotic gap between the sender's intended meaning and the recipient's received interpretation.

Dimension	Sighted Experience	Screen Reader Experience
Perceptual	High certainty (visual form)	High ambiguity (text-only/unknown tone)
Conceptual	High ambiguity (cultural &language &social context influence)	High certainty (audio speaks label)

Coppin's notion of diagrams as hybrids, visual arrangements that combine pictorial and symbolic elements—parallels the structure of many emojis, which visually encode culturally embedded symbolic meanings. While emojis' visual form suggests pictorial qualities, their meaning is often shaped by symbolic conventions (e.g., 🍆 as a sexual innuendo or 🙏 as a gesture of gratitude or prayer). The breakdown of hybrid affordance when emojis are rendered as plain text by screen readers, stripping away their visual nuance and resulting in a loss of connotative meaning for blind and low-vision users (Coppin, 2014). This shift highlights a fundamental asymmetry: while some level of conceptual ambiguity can be beneficial—enabling flexible interpretation and user autonomy—excessive perceptual ambiguity (Coppin, 2014). As experienced by blind and low-vision users, the loss of iconic features in translation can hinder accurate interpretation and disrupt shared intentionality in online communication. Transforming visual symbols into fixed textual descriptors often introduces a cognitive gap, wherein the perceived meaning fails to align with the sender's intended message. This highlights the importance of distinguishing between useful ambiguity, which broadens expressive potential, and harmful ambiguity, which impedes comprehension and reduces information and possible actions (Barter & Coppin, 2022) — particularly in the communicative space between visual and non-visual users.

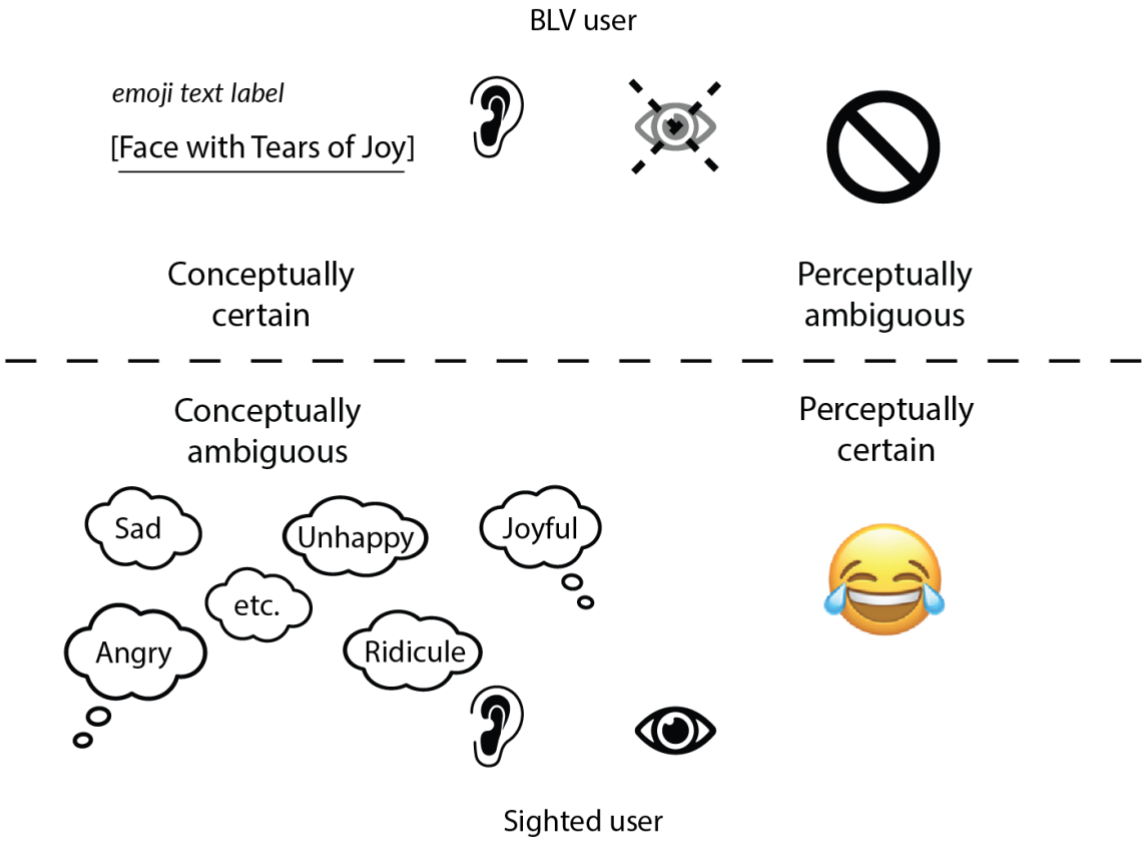

Figure 2: Model of the emoji perceptual & conceptual affordance in sighted and BLV user experience

For sighted users, the affordance of emojis is perceptually specific. Everyone sees the same visual shape, enabling a shared recognition of the symbol. However, this specificity at the perceptual level leads to conceptual ambiguity (Coppin, 2014), as interpretations of the same emoji can vary widely depending on context, culture, and personal experience. For instance, the 😂emoji may be visually identical for all users, but its meaning can shift between genuine amusement, sarcasm, or even emotional deflection. Sighted users exercise interpretive agency by drawing on contextual cues to determine whether the emoji conveys a positive, neutral, or negative tone within a given conversational setting. BLV users only know the official description of this emoji 😂- [Face with Tears of Joy]. The shared accessible information is limited which leads to conceptual specificity, but is perceptually ambiguous (Coppin, 2014).

An important observation to note is that conceptual ambiguity (Coppin, 2014) to sighted users is not necessarily a limitation; rather, it is a defining and playful feature of emojis, enabling them to convey nuances and sentiments that written text alone often cannot express. Emoji accessibility is not solely

dependent on technical translation (e.g., providing alt text) but requires rethinking how affordances are presented across sensory modalities. However, for screen reader users, ambiguity arises from perceptual unknowns and unpredictability — leading to multiple misunderstandings.

7.5 Future Work

While this study provided valuable insights into how blind and low-vision (BLV) users experience emojis through screen readers, there remain several directions for future exploration. Future research could expand the participant pool to include a broader range of users—such as those with varying levels of generation, tech literacy, cultural backgrounds, or non-English language preferences—would help uncover additional nuances in emoji interpretation and reflection on sound cues.

Future research could also further investigate the design of alternative representations, with particular emphasis on the refinement sound cues strategy to better understand their potential. And testing to enhance the user experience of BLV individuals. Collaborating with developers of mainstream screen readers and communication platforms (e.g., Zoom, Microsoft Teams, WhatsApp) will be essential to test the feasibility of these solutions and advance.

Finally, through community-led design or participatory emoji proposals, future research could explore the development of an inclusive emoji archive—documenting emojis whose label descriptions differ significantly from their visual appearance to BLV users and emojis that often create substantial gaps in understanding for BLV users compared to sighted individuals. To develop a future set of emoji dictionaries that are more accessible, representative, and semantically rich for all users, regardless of their visual ability.

8. CONCLUSION

This study argues the current accessibility standard for screen readers in emoji interacting experience - fails to convey the visual affordance of emoji communication for BLV users. This paper has demonstrated that the affordances of emojis are perceived differently by BLV users and sighted users in different pathways. Official emoji labelling constrains their perceptual richness compared to their visual appearance, which limits the perception of BLV users. Compared to the experience of sighted users, emojis are often interpreted beyond their standardized textual descriptions. These findings highlight the urgency for screen reader designers to reconsider the accessibility standards for emoji and develop more inclusive representations that account for non-visual modes of interpretation.

9. REFERENCES

Abercrombie, D. (1968). Paralanguage. British Journal of Disorders of Communication, 3(1), 55–59. https://doi.org/10.3109/13682826809011441

Alnuzaili, E. S., Amin, M. W., Alghamdi, S. S., Malik, N. A., Alhaj, A. A., & Ali, A. (2024). Emojis as graphic equivalents of prosodic features in natural speech: Evidence from computer-mediated discourse of WhatsApp and Facebook. Cogent Arts & Humanities, 11(1), 2391646. https://doi.org/10.1080/23311983.2024.2391646

Barbieri, F., Ronzano, F., & Saggion, H. (2016). What does this emoji mean? A vector space skip-gram model for Twitter emojis. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016) (pp. 3967–3972). European Language Resources Association (ELRA).

Barter, D., & Coppin, P. (2022). A diagram must never be ten thousand words: Text-based (sentential) approaches to diagrams accessibility limit users' potential for normative agency. In S. Linker (Ed.), Diagrammatic Representation and Inference: Diagrams 2022 (Lecture Notes in Computer Science, Vol. 13495, pp. 86–102). Springer. https://doi.org/10.1007/978-3-031-15146-0_7

Canadian National Institute for the Blind. (2022). Breakthroughs: Annual report 2021–2022. https://www.cnib.ca/en/2021-2022-annual-report

Choi, Y., Hyun, K. H., & Lee, J.-H. (2017). Image-based tactile emojis: Improved interpretation of message intention and subtle nuance for visually impaired individuals. International Journal of Human–Computer Interaction, 33(11), 899–911. https://doi.org/10.1080/07370024.2017.1324305

Barter, C., & Coppin, P. (2022). Translating iconographic emoji for screen reader users: A study on affordance loss and sonification proposals. In Proceedings of the 19th International Web for All Conference (W4A '22) (pp. 1–9). Association for Computing Machinery. https://doi.org/10.1145/3493612.3520450

Coppin, P. W. (2014). Perceptual-cognitive properties of pictures, diagrams, and sentences: Toward a science of visual information design (Doctoral dissertation, University of Toronto). https://hdl.handle.net/1807/67215

Coppin, P., Li, A., & Carnevale, M. (2016). Iconic properties are lost when translating visual graphics to text for accessibility. Manuscript submitted for publication. http://openresearch.ocadu.ca/id/eprint/1035/

Cole, J. (2015). Prosody in context: A review. Language, Cognition and Neuroscience, 30(1–2), 1–31. https://doi.org/10.1080/23273798.2014.963130

Grosz, P. G., Greenberg, G., De Leon, C., & Kaiser, E. (2023). A semantics of face emoji in discourse. Linguistics and Philosophy, 46(4), 905–957. https://doi.org/10.1007/s10988-022-09369-8

Gesselman, A. N., Ta, V. P., & Garcia, J. R. (2019). Worth a thousand interpersonal words: Emoji as affective signals for relationship-oriented digital communication. PLOS ONE, 14(8), e0221297. https://doi.org/10.1371/journal.pone.0221297

Guntuku, S. C., Li, M., Tay, L., & Ungar, L. H. (2019). Studying cultural differences in emoji usage across the East and the West. Proceedings of the International AAAI Conference on Web and Social Media, 13(1), 226–235. https://doi.org/10.1609/icwsm.v13i01.3224

Haralambous, Y., & Dürst, M. (2019). Unicode from a linguistic point of view. In Y. Haralambous (Ed.), Grapholinguistics in the 21st Century: Proceedings (pp. 1–20). Fluxus Editions.

Kleege, G. (2018). More than meets the eye: What blindness brings to art. Oxford University Press.

Leporini, B. (2022). Screen readers. In Assistive technologies and accessibility (pp. 1–20). ResearchGate.

Lee, S., Takagi, H., & Asakawa, C. (2019). Voicemoji: Emoji Entry Using Voice for Visually Impaired People. Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19), 648–650. https://doi.org/10.1145/3308561.3354603

Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L., & Hecht, B. (2016). “Blissfully happy” or “ready to fight”: Varying interpretations of emoji. Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM), 2016, 259–268. https://ojs.aaai.org/index.php/ICWSM/article/view/14633

Miltner, K. M. (2021). “One part politics, one part technology, one part history”: Racial representation in the Unicode 7.0 emoji set. New Media & Society, 23(3), 515–534. https://doi.org/10.1177/1461444819899624

Omvig, J. H. (2005). The blindness revolution: Jernigan in his own words. Information Age Publishing.

Park, S., & Takagi, H. (2017). Image-Based Tactile Emojis: Improved Interpretation of Message Intention and Subtle Nuance for Visually Impaired Individuals. Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '17), 67–76. https://doi.org/10.1145/3132525.3132540

Seargeant, P. (2019). The emoji revolution: How technology is shaping the future of communication. Cambridge University Press.

Tigwell, G. W., Gorman, B. M., & Menzies, R. (2020). Emoji accessibility for visually impaired people. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Article 140, 1–13. Association for Computing Machinery. https://doi.org/10.1145/3313831.3376267

Tantawi, Y., & Rosson, M. B. (2019). The paralinguistic function of emojis in Twitter communication. In Proceedings of the Social Media Corpora for the Humanities (CMC-Corpora2019).

Trager, G. L. 1958. Paralanguage: A first approximation. Studies in Linguistics 13:1 12.

World Health Organization. (2019). World report on vision. https://www.who.int/publications/i/item/9789241516570

Wagner, M. (2016). How to be kind with prosody. In J. Barnes, A. Brugos, S. Shattuck-Hufnagel, & N. Veilleux (Eds.), Proceedings of Speech Prosody 2016 (pp. 250–1253). Urbana, IL: Speech Prosody Special Interest Group (SProSIG).

Zhang, M., Wang, R., Xu, X., Li, Q., Sharif, A., Wobbrock, J.O. (2021). Voicemoji: Emoji Entry Using Voice for Visually Impaired People. Association for Computing Machinery. https://doi.org/10.1145/3411764.3445338

Zappavigna, M., & Zhao, S. (2017). Selfies in 'mommyblogging': An emerging visual genre. Discourse, Context & Media, 20, 239247.