Human Languages in HCI: Beyond User Interface Localization

Diego Moreira da Rosa[1(✉)], Leandro Soares Guedes[2], Monica Landoni[2], and Milene Silveira[1]

[1] Pontifícia Universidade Católica Do Rio Grande Do Sul, Porto Alegre, RS, Brazil
diego.rosa81@edu.pucrs.br, milene.silveira@pucrs.br
[2] Università della Svizzera italiana, Lugano, TI, Switzerland
{leandro.soares.guedes,monica.landoni}@usi.ch

Abstract. This paper investigates the role of language in interaction design and computer-mediated communication. It starts by distinguishing between *human languages* and *computer languages*. Human language, as the primary code used in text and voice-based communication, is critical in the majority of communication processes. We also analyze two methods of computer-mediated communication. First, interaction design based on semiotic engineering is investigated; followed by user-to-user computer-mediated communication. This paper provides insights into the design of user interfaces, cross-cultural communication, and social media by discussing the complexities of language in human-computer interaction and computer-mediated communication. The findings highlight the significance of language as the main component of effective communication involving humans and computers in various contexts.

Keywords: Language · Multilingualism · Communication Models · Computer-mediated Communication · Human-computer interaction

1 Introduction

Adapting software systems to international users has been a key issue for the software industry since the emergence of personal computers by the end of the '70 s. For many years, the processes of software internationalization (i18n) and localization (L10n) were the focus of development teams to attend the demands of users in many different countries [6]. Until the popularization of the Internet, most communication occurred between the system and the user, and once the software was installed and configured, a single language was used throughout this interaction. By the turn of the millennium, the emergence of Web 2.0, combined with a wider access to technology and Internet connection in developing countries, radically impacted how computer systems were designed and used [16]. Beyond the system-to-user communication, it is the computer-mediated user-to-user communication that gains importance as users from all over the world start

M. Kurosu et al. (Eds.): HCII 2023, LNCS 14054, pp. 564–574, 2023.
https://doi.org/10.1007/978-3-031-48038-6_36

interacting through social platforms. All these events have taken the multilingualism of human-computer interaction (HCI) to a higher level adding extra challenges to the software globalization process (g11n) [6].

In general, the number of languages or locales supported by a certain application or website was a good indicator of the level of globalization of the system [20,22]. Having a system localized to many different languages guarantees broader access to users from different countries. This can be seen as the multilingualism of the interface, aimed at dealing with the language diversity of the user base. In the context of the social web, in which users are not only consumers but also creators of content, the multilingualism of the individual gains importance, i.e., designers should take into consideration that every user has a linguistic background, may present some degree of bi- or multilingualism, and will have their own language preferences. To address these nuances, researchers in the field of HCI need to expand their focus beyond interface localization and adopt a broader perspective encompassing the multiple forms of multilingualism in interaction design.

This paper aims to provide a theoretical review of human languages within the context of HCI, establishing a foundation for future research within a comprehensive conceptual framework. The presented model of communication explores its relation to HCI, with particular emphasis on the role of human language in message encoding and decoding. Drawing upon theories of semiotic engineering, computer-mediated communication, and other relevant disciplines, we elucidate the diverse forms of communication that occur in HCI, supporting our analysis with real-world application examples. Finally, we discuss how designers, systems, and users negotiate language and engage in various forms of communication.

2 Methodology

To analyze the roles of human language in the many forms of communication involved in HCI, we adopted a three-step methodology as described below:

- **Review of definitions and types of language:** bibliographic research on the definition of language and the classification of the many types of language. Identify the definitions and types that are most relevant to the study of human-computer interaction.
- **Review basic models of communication:** revisit theories of communication and apply them in the context of HCI. Establish basic terminology and analyze the role of human language as the main code in the communication process.
- **Analysis of language use in the many forms of computer-mediated communication:** identify the many forms of communication in HCI and analyze how language is negotiated and used for message en-/decoding in each of them. Through exploratory research, present examples of current user interfaces to illustrate different scenarios.

3 Definition and Types of Languages

Communication and language are intricate concepts that have been extensively studied in fields such as Communication, Semiotics, and Linguistics. The definition and description of language can vary depending on the context and the specific interests or requirements of the study. Lyons suggests that most definitions view language as a system of symbols designed for communication among groups of human beings within a particular community [14]. Languages convey not only ideas but also express emotions, desires, identity, art and poetry [9,11,17,21]. Furthermore, languages can take various forms, including spoken, written, and hand symbols that belong to a convention system. This work focuses on written languages and their usage in human-computer interaction.

It is important to note that human language differs from communication systems used by other animals, such as whale vocalizations. Additionally, language is not the sole code employed in human communication and HCI. Nonverbal communication, such as body language, plays a significant role in human interaction [7], and iconography is essential in user interface design [3].

In the realm of computer science, a wide range of *computer languages* exists, which are formal languages used to communicate with computers. Among them, *programming languages* are employed to communicate instructions to machines, typically computers. In the field of Information Technology, the term *natural language* or *human language* is sometimes used to refer to a language employed in human communication, distinguishing it from various forms of computer languages (hence terms like *natural language processing*). Notwithstanding, some artificially constructed languages, such as Esperanto and Interlingua, effectively facilitate human communication and do not fit strictly within the definition of *natural language*.

In the domain of Human-Computer Interaction, professionals may use the terms *interface language* (referring to the language used by the system interface) and *user language(s)* (referring to the language(s) preferred or spoken by the user). In this study, the term *human language* is employed to denote language in its traditional definition as a system of human communication. This choice helps to avoid ambiguity with other types of language utilized in the field of Computing and its subareas.

4 A Model of Communication

When exploring language use within the context of HCI, it is valuable to establish a model of communication. Over the years, numerous models have been proposed, highlighting the multifaceted aspects of human communication [15]. Generally, a simple linear model that incorporates fundamental concepts is suitable for describing many forms of computer-mediated communication. Figure 1 depicts a straightforward model of communication based on the works of Shannon [19], Jakobson [12], and Schramm [18].

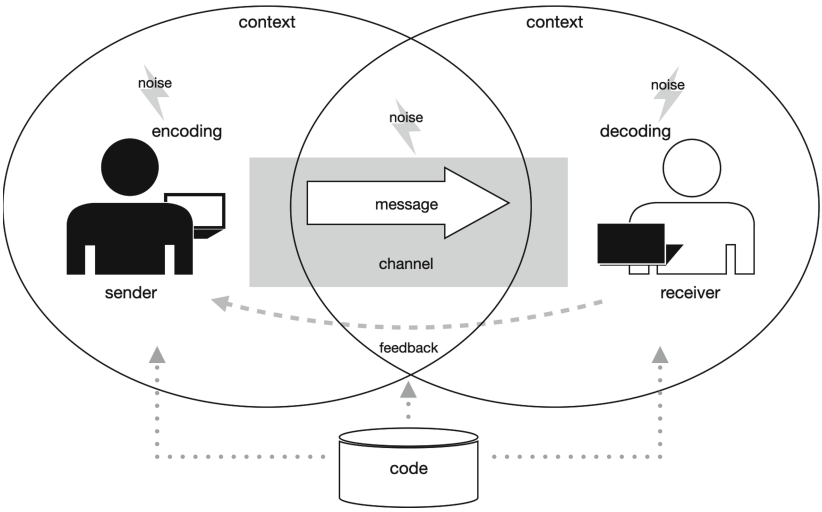

Fig. 1. Model of communication based on Shannon-Weaver's and Schramm's models.

This model incorporates essential elements that aid in describing the communication process and establishing a shared terminology:

1. **The sender**, also referred to as the *source*, is the individual who formulates, encodes, and transmits the message to the receiver.
2. **The receiver**, or *destination*, receives, decodes, and responds to the message sent by the sender.
3. **The message** comprises various types of signs (verbal, written, gestural, etc.) that convey information.
4. **Encoding** is the process of converting the message into a signal that can be transmitted through a channel.
5. **The communication channel** represents the medium, whether physical or logical, that enables the transfer of the message from the sender to the receiver.
6. **Decoding** is the process of translating the signal back into the message (the reverse of encoding).
7. **Feedback** is the response or reaction of the sender to the received message.
8. **The code** refers to any form of sign system, including human languages, utilized for encoding/decoding the message.
9. **The context** encompasses all the circumstances of the communication, including environmental, social, and personal aspects of each individual involved in the process.
10. **Noise** represents any influence on effective communication that has the potential to interfere with the interpretation of the message. Various types of noise can affect the communication channel (environmental) or the encoding/decoding process (psychological, semantic, etc.).

In the presented model, messages are encoded and decoded using a code, which is a sign system based on social convention. In text and voice-based communication, human languages serve as the primary code employed by the interlocutors. Therefore, language plays a key role in most communication processes.

Within the domain of human interaction, there are several types of communication, each in a different context [1]. According to the number of participants, communication can be classified in *intrapersonal* (communication with oneself), *interpersonal* (communication between two persons or in a small intimate group), *group communication*, and *mass communication*. *Organizational communication* describes the communication that takes place in larger, more permanent groups and contributes to the functioning of an organization. Finally, *computer-mediated communication* is the type of human communication that relies on electronic devices and software systems as a medium or channel of communication.

5 HCI, Computer Mediated-Communication, and Language

In the context of HCI, two forms of computer-mediated communication deserve a deeper analysis. The first one is the interaction design process, as described by semiotic engineering. The second one is the communication between users of social media. These two models present particularities regarding language usage in the communication process. In this work, we aim not to analyze the specifics of every existing communication model but to explore basic models that help understand the role of language in HCI.

5.1 Designer-to-User Communication (interaction Design)

De Souza originally proposed Semiotic Engineering as a semiotic approach to the design of user interface languages [4]. According to semiotic engineering, HCI can be viewed as a computer-mediated communication process between designers and users at interaction time [5] (see Fig. 2). In this process, the system itself is the message sent from designers to users. The system's user interface acts as the agent of the designer in communication and is capable of sending and receiving other messages. In other words, systems are metacommunication artifacts that should be engineered according to explicit semiotic principles [4].

The designer relies on a set of sign systems to design the user interface. These systems include the human language(s) in which the text and verbal-based signs of the interface will be produced. Other sign systems include iconography, color symbolism, and UI conventions. These sign systems must be totally or at least partially shared between the design team and the user for the metamessage to be understood. In cases the user base is expected to be representative of many different cultures, with considerably different mindsets, the user interface may have to be adapted in a process known as cross-cultural design.

In this type of designer-user communication, the interface language (the main language to be used at interaction time) is usually chosen by the user from a

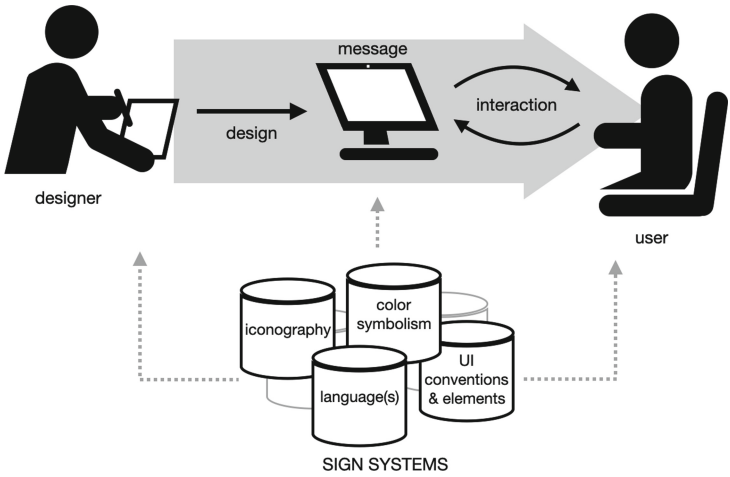

Fig. 2. Interaction design process viewed as designer-user communication according to semiotic engineering.

predefined set of languages or locales. In general, a locale is a combination of language and region/country and defines the linguistic preferences of the user interface. The process of adapting a software interface to different languages and regional peculiarities (including the *translation* of text strings) is called *localization* [6]. In order to operationalize the localization of software systems without the need of severe code changes, design teams adopt the *internationalization* design strategy.

Clearly, the selection of languages that the design team offers to the users is part of the metamessage. The website of Decathlon Switzerland[1] for example, is localized in German and French but lacks localized versions in the other two official languages of the country: Italian and Romansh (see Fig. 3). It would be as if the designer said to the user:

> "I understand that these are the languages that you might be interested in. In case you belong to the Italian- or Romansh-speaking communities of Switzerland, I assume that you can at least understand one of these two languages. Now you can select one of them and continue with your interaction."

Restricting the interface options to only two languages usually means that all associated services (account management, customer support, etc.) will be available only in those languages. Even though around 15% percent of the Swiss population will not have access to the retailer services in their native language, other aspects must also be considered, such as the elevated costs of the localization of an e-commerce platform.

[1] Sporting goods retailer: https://www.decathlon.ch/de/splashpage/.

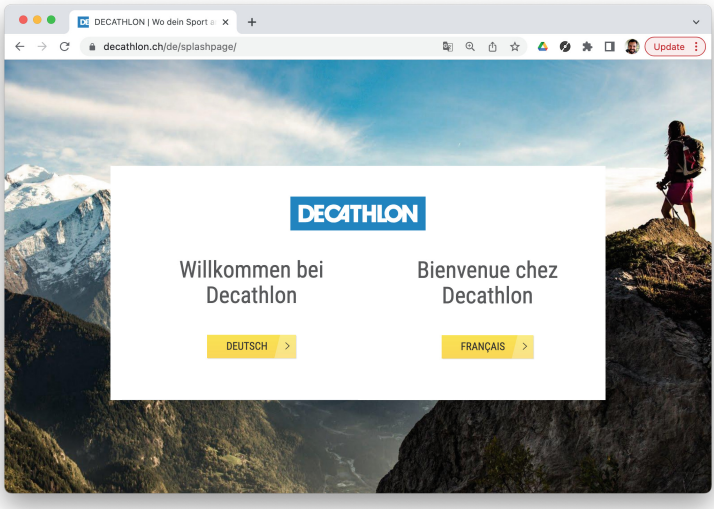

Fig. 3. Language selection page of Decathlon Switzerland website.

5.2 User-to-user Computer-Mediated Communication

In addition to the designer-user communication described in the previous section, HCI design also impacts all forms of user-to-user computer-mediated communication. Figure 4 shows a model for many-to-many communication typical of social media environments. This model is based on the model of intercultural communication proposed by Haworth and Savage [8] and the model for hypermedia marketing proposed by Hoffman and Novak [10]. One-to-one and one-to-many interpersonal communication can be seen as a special cases of this model.

The diagram in Fig. 4 presents eight users from four different countries interacting through a social platform. The platform can be a social network, an online forum, an interactive e-commerce website, or any other form of Web 2.0 application. The multiple countries emphasize the multicultural characteristic of many of these systems. A certain number of languages is attributed to each user representing both the multilingualism of the individuals (who may have skills in one, two, or more languages), and the multilingualism of countries (which are home to individuals with a variety of language skills). The observed increase in language diversity intensifies the challenges of the interaction design process.

Users communicate by posting user-generated content to the platform, which acts as a repository and/or distributor of these content units. The content generated by users can be a post on a social network, a message in a forum, a product review in e-commerce, etc. In this model, user-generated content is equivalent to the message in the traditional model of communication. The social platform, in turn, is the communication channel through which messages are propagated.

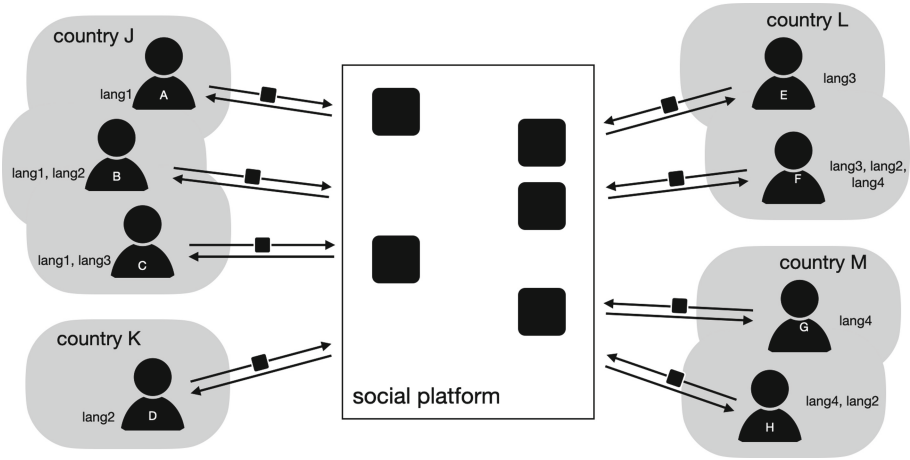

Fig. 4. Intercultural communication through social media (based on Haworth's and Hoffman's works).

Multiculturalism is the first aspect to be observed in the heterogeneous environment of a social media platform. In order to maintain a cordial relationship between members from many different cultures and nationalities, most of these systems publish a set of rules or guidelines that should be followed by every member (e.g., Facebook Community Standards[2], Reddit Content Policy[3], Booking.com Reviews Guidelines[4], etc.). These guidelines are then enforced by the companies' staff and by other members of the online community through abuse reporting.

Multilingualism is also remarkable in social platforms. The level of multilingual support will depend on many aspects of the system design. Some systems may limit the available languages that users can choose due to software/technical constraints. In online forums, it is common to find communities determining which is the language of choice for a certain group/topic of discussion [2]. Multilingual users can also vary the language of messages and posts according to the context [13]. Some systems offer features that facilitate access to multilingual content, such as content filtering by language and automatic machine translation.

Figure 5 shows the user reviews page of an accommodation listing on Booking.com. The image highlights various language-related aspects of the interface. The language settings in the top right corner indicate that the device is configured for English (EN tag). On the left side of the interface, the Booking website itself has been switched to Italian using its configuration options (IT tag). The user reviews are displayed in their original languages. In the given example, one

[2] https://transparency.fb.com/policies/community-standards/.
[3] https://www.redditinc.com/policies/content-policy.
[4] https://www.booking.com/reviews_guidelines.html.

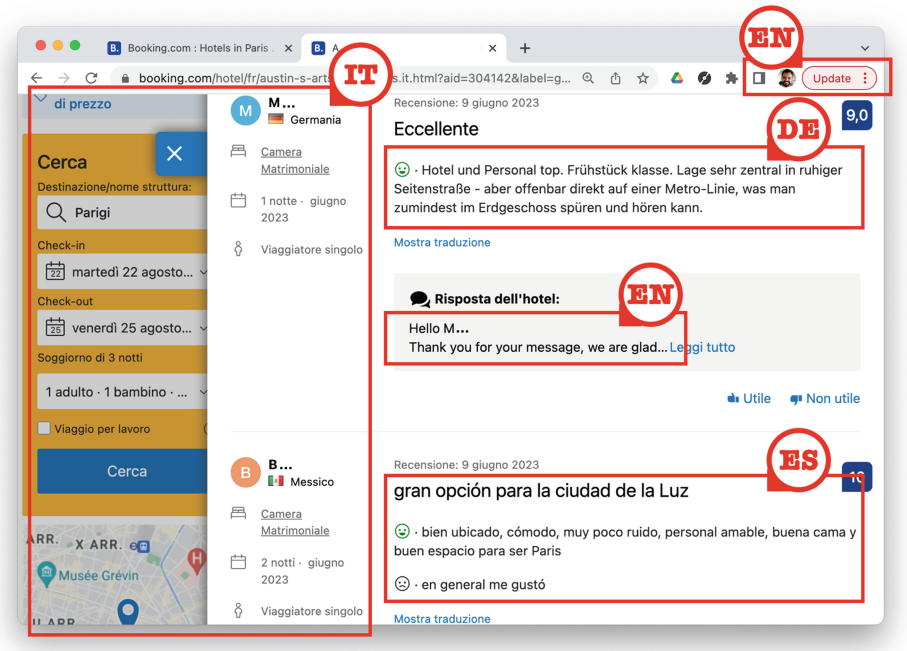

Fig. 5. International user reviews on a hotel page of Booking.com website (accessed in Jun 2023).

review is shown in German (DE tag), while another review is in Spanish (ES tag). Notably, the Booking website permits hotel managers to respond to user reviews. These responses can be composed in different languages based on the specific situation. An English response is provided for a German review (EN tag in the center). Additionally, the system offers several features to assist users in dealing with multilingual content. An automatic translation feature allows users to translate other users' reviews. Furthermore, a language filter allows users to sort and visualize reviews based on their preferred language. Nonetheless, it is important to note that the automatic translation feature is not available for hotel management responses.

Booking.com is an exemplary platform in terms of multilingual support, as it offers localized interfaces in 45 languages. This allows a significant portion of the global population to access the platform in their native language. Furthermore, the platform actively encourages the display of international user-generated content. Notable features include facilitating users' access to multilingual content and providing tools to effectively manage and filter such content.

6 Conclusion

This theoretical review provides valuable insights for Human-Computer Interaction researchers interested in exploring language use in interactive system design.

By establishing a common terminology through a communication model, it lays the foundation for discussions on language and HCI. The model's components, such as sender, receiver, message, encoding, decoding, channel, feedback, code, context, and noise, are instrumental in understanding and defining communication in two key contexts: designer-user metacommunication during interaction design (as described by Semiotic Engineering), and user-to-user communication on social media platforms. In the latter context, the review emphasizes the significance of multiculturalism and multilingualism in contemporary systems.

The globalization of information and communication technologies increased the levels of multiculturalism and multilingualism in software usage. Offering localized interfaces is still an important task that guarantees access to digital systems for many users. Nonetheless, design teams now have bigger challenges that go beyond software localization, such as proposing solutions for the presentation and navigation of multilingual user-generated content.

References

1. Adler, R.B., Rodman, G.R., Sévigny, A.: Understanding Human Communication, vol. 10. Oxford University Press, Oxford (2006)
2. Birnie-Smith, J.R.: Ethnic identity and language choice across online forums. Int. J. Multiling. **13**(2), 165–183 (2016)
3. Blankenberger, S., Hahn, K.: Effects of icon design on human-computer interaction. Int. J. Man Mach. Stud. **35**(3), 363–377 (1991)
4. De Souza, C.S.: The semiotic engineering of user interface languages. Int. J. Man Mach. Stud. **39**(5), 753–773 (1993)
5. De Souza, C.S.: The semiotic engineering of human-computer interaction. MIT press (2005)
6. Esselink, B.: A practical guide to localization, vol. 4. John Benjamins Publishing (2000)
7. Fast, J.: Body language, vol. 82348. Simon and Schuster (1970)
8. Haworth, D.A., Savage, G.T.: A channel-ratio model of intercultural communication: The trains won't sell, fix them please. J. Business Commun. (1973) **26**(3), 231–254 (1989)
9. Hébert, L., Tabler, J.: The functions of language. In: An Introduction to Applied Semiotics, pp. 232–240. Routledge (2019)
10. Hoffman, D.L., Novak, T.P.: Marketing in hypermedia computer-mediated environments: Conceptual foundations. J. Mark. **60**(3), 50–68 (1996)
11. Huffaker, D.A., Calvert, S.L.: Gender, identity, and language use in teenage blogs. J. Comput.-Mediated Commun. **10**(2), JCMC10211 (2005)
12. Jakobson, R.: Linguistics and poetics. In: Style in language, pp. 350–377. MA: MIT Press (1960)
13. Karusala, N., Vishwanath, A., Vashistha, A., Kumar, S., Kumar, N.: "only if you use English you will get to more things:" Using smartphones to navigate multilingualism. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–14 (2018)
14. Lyons, J.: Language and linguistics. Cambridge University Press (1981)
15. Narula, U.: Communication models. Atlantic Publishers and Dist (2006)
16. O'Reilly, T.: What is web 2.0. "O'Reilly Media, Inc." (2009)

17. Sapir, E.: An introduction to the study of speech. Language 1 (1921)
18. Schram, W.E.: The process and effects of mass communication. University of Illinois Press (1954)
19. Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**(3), 379–423 (1948)
20. Singh, N., Boughton, P.D.: Measuring website globalization: a cross-sectional country and industry level analysis. J. Website Promot. **1**(3), 3–20 (2005)
21. Waugh, L.R.: The poetic function and the nature of language. Verbal art, verbal sign, verbal time, pp. 143–68 (1985)
22. Yunker, J.: Have we reached "peak" language? (2023). https://globalbydesign.com/2023/03/02/have-we-reached-peak-language/. Accessed Apr 2023