

Digital Tongues: Internet Language, Collective Identity, and Implications for Human-Computer Interaction

Zi-Xiang Lin

Kang Chiao International School
{zixiang.ryan.lin@gmail.com}

Abstract

Nowadays, internet languages, including emojis, memes, hashtags, and slang, have become vital in constructing online communities' collective identities. However, all these forms of internet language can sometimes disempower people from other generations or cultures. This position paper presents an argument explaining how online forms of communication create social belonging for specific groups at the expense of users, and especially elderly people, due to interpretation hurdles. The present study aims to evaluate the relationship between the internet language and online collective identity, highlighting how patterns in internet language can inform human-computer interaction (HCI) by revealing how users' express identity, inclusion, and exclusion online.

1 Introduction

Language is a powerful tool to describe events, record dialogue, and form a collective identity. Indeed, language can be in any form that stands as a method of human communication. Theatre and Dance Assistant Professor Deborah Paredez's book "Selenidad: Selena, Latinos, and the Performance of Memory" captures how Selena utilized song, a musical performance, as a "language" to express Latino identity. Born on April 16, 1971, Selena is a Texas Tejano singer whose songs deeply reflect her Mexican-American experiences and cultural pride. As recorded by Paredez (2009), Selena not only resonates with other Latinos by featuring elements of her culture but also forms a collective identity through her actions and voice. Shifting into the 21st century, where language is utilized digitally, internet languages play a key role in building collective identity and redefining how people

interact, create, and share cultural narratives, just like how Salena uses her voice. Unlike traditional communities bound by geographical borders, online communities exchange or share information on social media platforms every day at any time.

Specifically, four main types of internet languages have been created to communicate online: slang offers casual vocabulary for quick communication, emojis convey emotions and thoughts, memes capture experiences in visual formats, and hashtags create themes for posts (Barron and Bollen, 2022; Graham, 2019; Oliseyenum and Oghenetega, 2023; Petrova, 2021). Becoming more prevalent in online social platforms, internet languages' emergence raises the question: To what extent does internet language contribute to collective identity formation in online communities?

As the internet evolves as the primary global interaction, the ability to engage diverse users through the internet language is crucial for building a collective identity among online communities. When analyzing the linguistic contribution to online communities, the generational factor appears to disintegrate the online community as a whole; understanding future trends is also essential since digital transformation is rapid, cross-generational, and on the cutting edge for new lifestyles. This paper explores how internet language influences the formation of collective identity in online communities, particularly using slang, emojis, memes, and hashtags. While not proposing design solutions, the present position paper highlights how patterns in internet language can inform human-computer interaction (HCI) by revealing how users' express identity, inclusion, and exclusion online. These insights may support the future HCI and NLP research designs toward more inclusive, culturally aware interfaces that bridge generational and social gaps.

2 Internet languages and Collective Identity in Online Community

Internet languages are writing language conventions and linguistic features in online communication, conveying a message (Squires, 2010). Although internet languages usually appear in “word text” such as slang and hashtags, they can be presented in graphs like emojis or memes. Specifically, slang and emojis convey ideas or emotions, while memes and hashtags connect users with shared experiences. According to English professor Graham (2019) and foreign languages associate professor Petrova (2021), the internet language bridges cultural barriers in real life as people resonate with others based primarily on feelings. Specifically, individuals can interpret the emotions of emojis and memes from their selection without knowing the context. They promote emotional resonance, allowing users to share and experience similar emotions. This ability to communicate nuanced emotions through simple graphics and texts fosters cross-cultural understanding as they promote openness, flexibility, and creativity, especially when people are lonely and can find online companionship. Thus, the widespread use of emojis and memes blends creativity with cognitive development, forming a collective identity in the communication medium.

Also, to build emotional connections, the hash character (#) implies the intention to establish a connection with others, as indicated by computational social science postdoctoral associates Barron and Bollen (2022). For instance, the hashtag in social media posts underlines the media user’s identity, like “#metoo,” subtly inviting others to resonate with the sender. As the number of viewers increases, more and more people will feel the same and join the discussion with similar beliefs. Ultimately, this fosters individual consciousness and expands it to the internet users’ collective identity.

Yet, several studies hold the opposite argument. Social psychology professor Hogg (2016) linked online users’ behavior to Social Identity Theory (SIT), enhancing in-group members’ traits and reducing characteristics similar to those outside the group. This suggests that affiliation to specific groups increases, while the association between diverse groups diminishes, leading to party isolation and trait dissociation. The dynamic

highlights how online interactions can undermine the online community wholly and impact social behavior, ultimately influencing users’ engagement and perspectives in digital spaces.

Extending this idea to internet language, English literary study doctors Oliseyenum and Oghenetega (2023) claim that slang is “a specialized form of language variation unique to a particular social group.” That is, only people in the “community,” aside from additional learning, can understand the meaning of the slang. While the essences of emojis, memes, and hashtags are often easy to recognize, they also require users to grasp subtle meanings for full understanding. As a result, these “cognitive limits” restrict internet language from being used outside a specific group, supporting SIT and building barriers to online communication.

Similar observations are concluded by associate psychology professor Bäck et al. (2018). The researchers found that singular pronoun (I) usage among online platforms decreases, while plural pronoun (we/they) significantly increases, “distancing to more outgroups.” This shift in pronoun use illustrates in-group bonding and highlights how digital language norms exclude newcomers, creating barriers to unity. Thus, the study reflects the decline of individual consciousness and the rise of group awareness within similar parties, supporting the SIT. Concluding the linguistic perspective, although internet languages strengthen the bonds in each group with similar traits, they aggravate the gap between different ethnic parties, instead of integrating the online community as a whole.

3 Generational Gaps and Digital Literacy

Binney (2004) indicates the interpretation of text and narratives will be “reinterpreted by another generation” as they “will continue changing and be changed.” This underlines the characteristics of language, which will be interpreted differently in different eras. The study also implies that people of different generations have different views on the same language, justifying that older and younger people may have different interpretations or understandings of internet language.

Indeed, economics and informatics assistant professor Hysa et al. (2021) discovered the daily social media usage of Baby Boomers (born in 1945–1964), Generation X (born in 1965–1980), Generation Y (born in 1981–1994), and Generation

Z (born in 1995–2010) are 28.6%, 40%, 50%, and 90% respectively. Specifically, social media usage is higher among Generation Z and Y than Generation X and Baby Boomers. The findings indicate that social network utilization decreases as age increases, illustrating the generational communication differences. As younger generations seek online media as the primary transmission source, this potentially causes separation of generational groups in online communities, since youth surf on online platforms more often.

In addition, the selection of internet language by different generations also contributes to the discrepancy. According to language and literature assistant professor [Azad et al. \(2023\)](#), Generation Z embraces highly informal language, abbreviations, and internet language. In contrast, Baby Boomers, the older generation, prefer more formal language and traditional communication norms on social media ([Puspita and Ardianto, 2024](#)). This reflects Generation Z's adaptation to digital communication's fast-paced, visually-oriented nature, challenging the elders to comprehend the meaning. Combining [Hysa et al. \(2021\)](#) and [Azad et al. \(2023\)](#), the two studies underline the destruction of collective identity between generations due to diverse internet language preferences.

Moreover, digital generational differences also affect real-life relationships among different age groups, causing digital gaps to widen. Unlike older generations, Generation Z relies on digital communication and integrates remote work ([Pichler et al., 2021](#)). Since work behaviors differ among generations, with the younger generation employing more social platforms, the fundamental digital differences affect both internet usage and physical human interactions, creating online media barriers. As time goes on, the younger generation will embrace more digital technologies, while the older generation continues to employ the traditional working style. Thus, the elders will not collaborate with youth and therefore do not use internet language, demonstrating that internet language cannot unite people of different ages. Concluding the generational perspective, online and offline interactions will exacerbate the generational gaps, underlining that different internet language utilization dissociates generational collective identity.

4 Future Digital Platform

For future predictions, researchers believe that online communication will resolve the barriers as they integrate languages on the internet to develop a collective identity. After dozens of years, the current generational barriers will be mitigated as the current youth emerge and dominate the society, resulting in almost everyone employing internet language to communicate. In fact, world studies professor [Godwin-Jones \(2018\)](#) hypothesizes that internet language, the informal language choices, will gradually replace traditional formal communication as digital communication becomes more integral to everyday life. That is, the employment of slang, hashtags, emojis, and memes will all significantly increase as younger generations adopt them progressively.

Nevertheless, while the original generational barriers are mitigated, new generational barriers will emerge. For instance, [Nurhayati \(2025\)](#) points out Generation Alpha (born in 2011–2024) continues to grow their language repertoire, preferring conventional idioms and less complex syntax. This contrasts with Generation Z, who demonstrates a better contextualized and multi-layered comprehension of internet-based phrases. Furthermore, [Melissa et al. \(2024\)](#) attributes this phenomenon since Generation Alpha interacts more with short-form video platforms, including YouTube and TikTok, that makes them act differently with Generation Z, who are influenced by earlier social media culture.

Regarding prediction on blending languages, Spanish and Portuguese assistant professor [Dickinson \(2023\)](#) observed codeswitching, the process of changing language, in online blogs. By combining Spanish and English, [Dickinson's \(2023\)](#) case study reached more internet users since understanding was not limited to only English or Spanish speakers. The tendency to attract more social media users will encourage others to follow, foreshadowing higher employment of codeswitching. However, [Dickinson \(2023\)](#) fails to acknowledge the challenges of regional barriers or integrating languages from distinct language families. Dickinson observes residents' behaviors who live in the United States to evaluate codeswitching. Since both Spanish and English are spoken in the United States, Dickinson's study can only support language integration in places that already use those languages, failing to resolve barriers in online communication across countries

and continents that speak distinct languages. As a result, the future norm of combining languages in online platforms is unlikely to happen and form a worldwide collective identity.

Proposing alternative future predictions on online communication developments, several researchers claim that the growth of emojis and memes will be faster because they do not necessarily require the viewer to know the language and understand the meaning. Specifically, computer science researchers [Raj and Balachandran \(2020\)](#) and Artificial Intelligence safety researcher [Weng et al. \(2014\)](#) predict that emojis and memes in online communication will rise, becoming more embedded in digital culture as their overall usage is expected to increase. To expand viewership, emojis and memes will experience more significant changes because they can acquire a wider range of online users, as they often transcend language barriers. Their language-independent nature allows users to easily grasp meaning, breaking communication barriers and fostering a global collective identity. Comparing [Dickinson's \(2023\)](#) and other researchers' predictions, future online platforms are more likely to shift toward image-based Internet languages because they demand fewer language restraints.

Elaborating on fewer text-based languages in the future, computer information researcher [Penni \(2016\)](#) further forecasts that video, with fewer verbal limitations, will dominate online interactions, as it is more dynamic and easier to attract attention. [Penni's \(2016\)](#) prediction of video evolution eliminates the possibility that viewers do not understand the emojis, memes, or text in Internet languages. As users increasingly prefer video content, communication on online social networks will shift toward video-based interactions rather than text or image-based internet language. Nevertheless, video content cannot completely avoid text usage because videos can only better engage the viewer but cannot replace verbal communication to exchange information.

To solve this, psychology professor [Gernsbacher \(2015\)](#) discovered that video captions benefit viewers, especially non-native speakers, who watch videos to improve comprehension by more than 60%. To spread information, both internet languages and video require the viewers to at least understand the meaning, emotion, or experience. However, a video without any language that comprehends these elements will decrease its

employment as online users cannot comprehend the video, failing to engage users. Thus, although internet communication will shift toward videos, internet languages will still be required.

Combining [Penni's \(2016\)](#) and others' predictions, the future social media platform will employ more emojis and memes along with videos, as they require fewer language recognition abilities while engaging more participants. These new forms of online communication will allow more people to join, understand the post, and share the same ideas, promoting the collective identity of the online community.

5 Implication for Human-Computer Interaction

The online collective identity dynamic and cognitive diversity among different linguistic or social groups present problems for Human-Computer Interaction and Natural Language Processing. Specifically, internet language evolves too quickly and symbols that are not in traditional corpora. Hence, the NLP models tend to be inaccurate when performing sentiment analysis or in the task of translation ([Ishita and Mamidi, 2025](#); [Khurana et al., 2022](#); [Raiaan et al., 2024](#)).

[Omar et al. \(2022\)](#) mentions that with these dynamic, creative, and sometimes adversarial internet languages, the models will not generalize, will misunderstand meaning, or will become vulnerable to adversarial attacks (e.g. minor changes of phrasing that evade spam filters or chatbots). In fact, [Ishita and Mamidi \(2025\)](#) explains Generation Alpha usage of slang has not yet been sufficiently translated by AI, wherein 89% of inaccuracies came from contextual misinterpretations. Since internet language depends on context, cultural references, and pragmatic cues, such as sarcasm or irony, this fast-evolving feature urges NLP systems and programmers to bridge the gap in real-world user-generated content scenarios to improve sentiment analysis and text translation.

While bridging cognitive gaps is evident, it is also important to take note of individuality to maintain online identity. [Van Der Meer \(2024\)](#) highlights that NLP systems and large language models tend to capture the dominant issues or opinions raised by the majority while often neglecting the minority or dissenting voices. This issue becomes a distortion of a complete collective identity, impeding the formation of diversity upon

which any robust online community is based (Burton et al., 2024). Thus, as the algorithms are trained over static or majority-centered corpora, the algorithms risk misinterpreting context-dependent expressions, neglecting minority voices, and accommodating emergent linguistic expressions. Moving forward, systems should be assisted by adaptability, inclusivity, and cultural sensitivity such that AI tools not only interpret languages but also respect and maintain diversity in online collective identities.

6 Conclusion

While internet languages will bridge groups with similar traits and develop primarily on image-based and video in the future to foster collective identity, they will disintegrate the connection between each online community, especially generations with different internet employment. As online languages evolve, the need for digital literacy will increase. Unlike traditional language learning, which often occurs in formal education settings, understanding internet language requires users to adapt to a fast-changing, informal digital environment. Effective communication in online spaces depends not only on users' ability to decode these forms but also on their skill in using them to engage and inform others meaningfully. Insights from this study may inform future HCI approaches that aim to bridge generational gaps, promote digital literacy, and support diverse forms of online expression without reinforcing exclusionary norms.

Limitations

Due to the fast-paced nature of online platforms, future internet languages may emerge rapidly, replacing existing ones. This creates linguistic gaps between regions, subcultures, and communities, making it challenging to form a cohesive collective identity online. To understand the future trends of internet languages, future research should monitor and investigate new forms of online communication continuously, accounting for the evolution of these languages as they develop.

Ethical Considerations

This position paper raises ethical concerns for future research about how internet language is interpreted in HCI design. Standardizing emojis, memes, or slang for usability risks erasing cultural and generational diversity. Future systems should

avoid reinforcing dominant norms or excluding marginalized groups. Researchers should also be cautious when applying automated analysis, ensuring that meaning is not oversimplified or misused.

References

- Alexander T. J. Barron and Johan Bollen. 2022. [Quantifying collective identity online from self-defining hashtags](#). *Scientific Reports*, 12(1).
- Beata Hysa, Aneta Karasek, and Iwona Zdonek. 2021. [Social media usage by different generations as a tool for Sustainable tourism Marketing in Society 5.0 idea](#). *Sustainability*, 13(3):1018.
- Deborah Paredez. 2009. *Selenidad: Selenia, Latinos, and the Performance of Memory*. Duke University Press.
- Diksha Khurana, Aditya Koli, Kiran Khatter, and Sukhdev Singh. 2022. [Natural language processing: state of the art, current trends and challenges](#). *Multimedia Tools and Applications*, 82(3):3713–3744.
- Emma A. Bäck, Hanna Bäck, Marie Gustafsson Sendén, and Sverker Sikström. 2018. [From I to We: Group formation and linguistic adaption in an online xenophobic forum](#). *Journal of Social and Political Psychology*, 6(1):76–91.
- Helen Wilfred Raj and Santhi Balachandran. 2020. [Future emoji entry prediction using neural networks](#). *Journal of Computer Science*, 16(2):150–157.
- Iis Kurnia Nurhayati. 2025. [Vernacular Evolution: A Comparative Study of Language Use in Generation Z and Generation Alpha on Social Media](#). *International Journal of Language and Cultural Linguistics*, 1(1):18–27.
- Inamul Azad, Sugandha Chhibber, and Azra Tajhizi. 2023. [How do different generations communicate on social media? A comparative analysis of language styles, emoji usage, and visual elements](#). *Language, Technology, and Social Media.*, 1(2):22–33.
- Ishita and Radhika Mamidi. 2025. [The Evolution of Gen Alpha Slang: Linguistic Patterns and AI Translation Challenges](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop)*, pages 678–686, Vienna, Austria. Association for Computational Linguistics.
- Janice Penni. 2016. [The future of online social networks \(OSN\): A measurement analysis using social media tools and application](#). *Telematics and Informatics*, 34(5):498–517.

- Jason W Burton, Ezequiel Lopez-Lopez, Shahar Hechtlinger, Zoe Rahwan, Samuel Aeschbach, Michiel A Bakker, Joshua A Becker, Aleks Berdichevskaya, Julian Berger, Levin Brinkmann, Lucie Flek, Stefan M Herzog, Saffron Huang, Sayash Kapoor, Arvind Narayanan, Anne-Marie Nussberger, Taha Yasseri, Pietro Nickl, Abdullah Almaatouq, et al. 2024. [How large language models can reshape collective intelligence](#). *Nature Human Behaviour*, 8(9):1643–1655.
- Judith Binney. 2004. Maori oral narratives, Pakeha written texts: two forms of telling history. *New Zealand Journal of History*, 38(2):203–214.
- Kendra V. Dickinson. 2023. [What does it Meme? English–Spanish codeswitching and enregisterment in virtual social space](#). *Languages*, 8(4):231.
- Lauren Squires. 2010. [Enregistering internet language](#). *Language in Society*, 39(4):457–492.
- Lilian Weng, Filippo Menczer, and Yong-Yeol Ahn. 2014. [Predicting successful memes using network and community structure](#). *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1):535–544.
- Maledo Richard Oliseyenum and Edobor Helen Oghenetega. 2023. [A morpho-semantic analysis of some Nigerian internet-based slangs](#). *Kampala International University Interdisciplinary Journal of Humanities and Social Sciences*, 4(1):110–126.
- Marwan Omar, Soohyeon Choi, Daehun Nyang, and David Mohaisen. 2022. [Robust natural language processing: recent advances, challenges, and future directions](#). *IEEE Access*, 10:86038–86056.
- Michael A. Hogg. 2016. [Social Identity Theory](#). In *Understanding peace and conflict through social identity theory: Contemporary global perspectives*, pages 3–17. Springer International Publishing.
- Michiel Van Der Meer. 2024. [Facilitating Opinion Diversity through Hybrid NLP Approaches](#). In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 4: Student Research Workshop)*, pages 272–284, Mexico City, Mexico. Association for Computational Linguistics.
- Mohaimenul Azam Khan Raiaan, Md. Saddam Hossain Mukta, Kaniz Fatema, Nur Mohammad Fahad, Sadman Sakib, Most Marufatul Jannat Mim, Jubaer Ahmad, Mohammed Eunus Ali, and Sami Azam. 2024. [A review on large language models: architectures, applications, taxonomies, open issues and challenges](#). *IEEE Access*, 12:26839–26874.
- Morton Ann Gernsbacher. 2015. [Video captions benefit everyone](#). *Policy Insights From the Behavioral and Brain Sciences*, 2(1):195–202.
- Paula Melissa, Maurenta Bunga Novia Siregar, Fathiha Malika Shakira, Lailan Haz, and Rahmadsyah Rangkuti. 2024. [Sociolinguistic Study: Variation of Slang Words between Gen Z and Gen Alpha](#). *Philosophica Jurnal Bahasa Sastra Dan Budaya*, 7(2):114–125.
- Robert Godwin-Jones. 2018. [Chasing the Butterfly effect: Informal language learning online as a complex system](#). *Language Learning & Technology*, 22(2):8–27.
- Sage L. Graham. 2019. [A wink and a nod: The role of emojis in forming digital communities](#). *Multilingua*, 38(4):377–400.
- Shaun Pichler, Chiranjeev Kohli, and Neil Granitz. 2021. [DITTO for Gen Z: A framework for leveraging the uniqueness of the new generation](#). *Business Horizons*, 64(5):599–610.
- Vinka Ganita Puspita and Ardik Ardianto. 2024. [Code-Switching and Slang: An analysis of language dynamics in the everyday lives of Generation Z](#). *Linguistics Initiative*, 4(1):76–87.
- Yulia Petrova. 2021. [Meme language, its impact on digital culture and collective thinking](#). *E3S Web of Conferences*, 273:11026.