

Soyjaks as forms of In-Group Communication and Identity Signaling in Online Communities

AP Research

April 30, 2026

Word Count: 3547

Introduction

Internet memes have become a major form of communication in modern culture, growing rapidly in popularity across social media and online communities. On the internet, people share ideas not only through written language but also through images and remixable formats. This shift has led to new ways of expressing and spreading ideas online that rely on visual and participatory content. One of the most evident examples of this are memes, which function as fast changing forms of cultural expression that spreads and changes through user interaction. Researchers have conceptualized memes as more than just humorous images, instead describing them as complex forms of communication. Limor Shafman defined memes as “(a) a group of digital items sharing common characteristics of content, form, and/or stance; (b) that were created with awareness of each other; and (c) were circulated, imitated, and transformed via the internet by multiple users” (Jenkins, 2014). This emphasizes that memes are not static but dynamic cultural systems shaped by participation. Memes play an important role in identity formation and group communication. Online communities often use shared meme formats to show belonging, reinforce social norms, and establish insider knowledge. Online spaces often develop their own cultural logics in which participation requires familiarity with certain patterns and symbols (Boyd, 2008). Certain memes evolve beyond simple image sharing into more structured systems of communication. One example is the soyjak meme community, primarily associated with forums and mainstream social media, which is built around recurring images, continuously modified character variants, and the constructed linguistic system known as Soyspeak. Within this community, users not only consume memes but actively create new trends and linguistic forms. The purpose of this study is to examine how soyjaks function as forms of in-group communication and identity signaling within online communities. It is hypothesized

that soyjak culture operates as a system in which meaning is constructed through shared knowledge of meme templates and linguistic rules, reinforcing group cohesion while limiting accessibility to outsiders.

Literature Review

Research on internet memes has increasingly focused on their role as systems of communication rather than just forms of entertainment. Memes function as a type of language that influences digital culture and collective thinking by shaping how individuals interpret and communicate meaning online (Petrova, 2021). Online communities develop shared cultural norms, which become systems for understanding and interacting with content. Within these communities, language becomes indicators of belonging and participation. Petrova further suggests that memes not only reflect communication but actively shape it by influencing how users think, express ideas, and engage with others. This supports the idea that memes operate as systems of communication rather than just being jokes. Expanding on this, the theory of participatory culture helps explain why memes spread so effectively in online environments (Janine, 2025). Digital media has shifted users from passive consumers to active contributors who reshape cultural content. In this context, memes function as tools for participation, allowing people to engage in shared creativity while contributing to evolving online communities. This explanation is essential for understanding meme ecosystems where constant variation and change are important. In addition to their cultural function, memes also have a significant role in identity formation within online spaces. Digital environments create shared spaces where identity is formed through interaction rather than fixed characteristics (Boyd, 2008). In these spaces, users rely on shared symbols, language, and references to establish group membership. Similarly,

language variation is tied to social identity, as individuals use specific linguistic styles to align themselves with particular communities (Eckert, 2016). Further research on internet language shows that online communities often develop their own structured forms of communication. Digital interaction encourages the creation of new linguistic patterns including abbreviations, slang words, and altered spellings that become meaningful within specific groups (Crystal, 2017). These patterns can develop into sociolects or community specific language systems which follow their own internal rules. Finally, researchers have highlighted the multifaceted nature of memes, meaning that they combine both visual and textual elements to communicate meaning. Rather than relying only on written language, memes use recognizable images, formats, and repeated structures to convey complex ideas quickly. This allows users who are familiar with the format to understand the meaning immediately, while those outside the community might struggle to understand it. Despite the increase of research on memes and online language, there is limited research on structured meme communities in which images, language, and user generated variations operate together as a unified system. In particular, communities centered around soyjaks demonstrate a combination of evolving character variants and linguistic patterns such as soyspeak. This study aims to address this gap by analyzing how these elements function together as mechanisms of in-group communication and identity signaling within online communities.

Methodology

This study examines how soyjaks create and sustain systems of in-group communication and identity within online communities. It focuses on three interconnected parts of soyjak culture: language patterns (Soyspeak), visual meme formats (soyjak variants), and participatory

behaviors. The hypothesis is that meaning within this community is constructed through shared customs that reinforce group identity while limiting accessibility to outsiders. To investigate this, the study uses a qualitative exploratory method. This approach is appropriate because the goal is not to measure frequency or test cause and effect relationships, but to analyze how meaning is created and understood within a specific subculture. Exploratory research is useful when studying communities that do not have existing research, as it allows for a deeper examination of patterns, structures, and interactions. Rather than relying on numerical data, this study focuses on identifying recurring features across different forms of content. This allows for a deeper understanding of how communication happens in soyjak culture. The data for this study consists of publicly available content from discussion forums, archives, and mainstream social media platforms where users create and share soyjak related content. These spaces were selected because they contain consistent examples of the language, imagery, and interaction patterns relevant to the study. In addition to identifying relevant spaces where soyjak content is produced and shared, this study applies a purposeful sampling strategy to select examples that best represent the core components being analyzed. Purposeful sampling is commonly used in qualitative research to focus on cases that illustrate the phenomenon in depth rather than attempting to generalize across a large population (Palinkas et. al, 2016). In this case, selected data includes threads, posts, and images that demonstrate consistent or repeated patterns of communication rather than isolated or random examples. The research setting is entirely digital, consisting of online platforms where soyjak culture is actively produced and circulated. This includes imageboards, forums, archival sites, and mainstream social media platforms such as Twitter, Tiktok, and Instagram. While soyjak culture is often associated with more niche platforms, it has increasingly spread into mainstream media. Including both types of spaces

allows for a broader understanding of how in-group communication systems are maintained internally while also being adapted or diluted when exposed to wider audiences. This dual setting approach is important because it highlights the tension between exclusivity and diffusion, which is central to how these communities construct identity. No direct human participants were used for this study. Instead, all data was collected from publicly accessible content, meaning that ethical concerns related to consent and privacy are minimized. Furthermore, to ensure ethical integrity, no usernames or identifying information are included, and all content is synthesized rather than focusing on individual users. This aligns with standard ethical practices for digital ethnography and online discourse analysis, where publicly available data can be used as long as individuals are not personally identified or harmed (Markham & Buchanan, 2012). The primary variables examined in this study are patterns in Soyspeak, such as grammar, syntax, and vocabulary, as well as recurring visual features found in soyjak variants. In addition, the study analyzes forms of user interaction including quoting, editing, and reacting to content, along with indicators that distinguish in-group from out-group communication. Together, these variables contribute to understanding how meaning is constructed within the community. To collect data, examples were gathered over an observation period by manually browsing selected platforms and identifying posts that had relevant features. Screenshots and textual transcriptions were used to preserve the original format of the content, which is especially important given the visual nature of soyjak memes. The inclusion of both text and image data allows for a more comprehensive analysis, as meaning is often created through the interaction between the two. The study also incorporates document based analysis using existing descriptions of soyspeak, soyjak variants, and community practices. These descriptions, while not always formally academic, provide insight into how members of the community understand and define their own

systems. For example, Soyspeak is described not simply as slang but as a structured form of communication with its own grammatical rules, mutation patterns, and conventions such as apostrophization and acronymization. As a result, it functions more like a language than typical internet slang, showing features commonly associated with dialects or sociolects. The primary tools used in this study are coding frameworks and thematic analysis techniques. After collecting data, examples were categorized based on recurring features. For instance, examples of soyspeak were grouped according to specific linguistic patterns, such as the use of apostrophes to remove syllables or the replacement of consonants. Similarly, soyjak variants were categorized based on visual characteristics and also through its system of variants and subvariants. This categorizes them based on whether they function as recognized variants which are widely recognized or subvariants which are derivative adaptations of those forms. This includes examining how variants are classified within the community through distinctions such as “IAS” (Is a Soyjak) or “NAS” (Not a Soyjak), as well as how they are evaluated using shared standards. Participatory behaviors were analyzed by identifying repeated interaction patterns, such as back and forth exchanges between users, the modification and reuse of existing images, and the application of specific rating systems such as “gem” and “coal”. These behaviors were treated as part of a broader communicative system, where meaning is not only conveyed through individual posts but through ongoing interaction. This coding process follows a thematic analysis approach, which involves identifying, analyzing, and interpreting patterns within qualitative data (Braun & Clarke, 2006). The goal is not to quantify how often something appears but to understand its significance within the system. For example, the mineral rating system is not just a way to evaluate content quality, it also functions as a shared language that reinforces group norms and values. To ensure reliability, the study focuses on consistency across multiple examples rather

than relying on a single instance. If a pattern appears repeatedly across different platforms or contexts, it is considered more significant. Additionally, examples were cross referenced between platforms to see how certain features change or remain stable when moving from niche communities to mainstream spaces. This helps strengthen the validity of the findings by showing that observed patterns are not isolated. The analysis also considers how accessibility affects communication. While the study does not focus primarily on exclusion, it acknowledges that many of these systems require prior knowledge to fully understand. This creates a distinction between insiders and outsiders, even without explicitly stating it. By examining how these systems operate, the study can infer how they contribute to identity formation within the group. Another important part of the analysis is the relationship between structure and participation. Unlike many meme cultures that rely primarily on shared references, soyjak culture involves active modification and creation. Users are expected not only to understand the culture but also contribute to it by producing new content. This participatory element is analyzed by examining how users build on existing formats, respond to each other, and follow or challenge established conventions.

Results/Findings

The analysis of the collected data reveals that soyjak culture operates through a structured system of communication in which meaning is constructed through repeated visual, linguistic, and participatory patterns. One of the most consistent findings is that Soyspeak functions less like informal internet slang and is more like a semi-structured linguistic system with its own rules. Across multiple examples, users consistently applied recognizable modifications such as syllable removal with apostrophization, consonant substitution, and abbreviation based

communication. These choices were not random but followed predictable patterns that remained stable across different contexts and platforms. In practice, Soyspeak appears to serve two functions. First, it reduces linguistic complexity for in-group communication, allowing for faster and more compressed expression. Second, it creates a barrier to interpretation for users unfamiliar with the format. Posts using Soyspeak were often only fully understandable when the reader had prior exposure to its rules, suggesting that knowledge of language functions as a marker of group membership. Furthermore, the analysis of soyjak images shows that variants function as visual categorization rather than random meme edits. Variants such as Chud, Bernd, Gapejak, and Feraljak consistently displayed recognizable visual traits that allowed them to be identified and categorized by community members. These traits included facial structure, expression, and additional features such as accessories or line width. A key finding is that classification systems like IAS and NAS determine whether a variant is accepted into the culture. Beyond this classification, the mineral rating system which ranges from brimstone, coal, iron, gem, to gemerald, further reinforces evaluative structure, where variants are not only recognized but ranked in terms of cultural value. A third major finding is that participation itself is a central communicative mechanism within the community. Users do not simply consume content but actively modify and respond to existing material. One of the most common interaction patterns was "soydueling", where users continuously quote each other with various soyjaks. In addition, terms such as "gem" and "coal" were frequently used as a form of social signaling, allowing users to collectively evaluate content while reinforcing shared standards of quality and humor. Across all platforms studied, participation was not optional but structurally encouraged. Content that was not modified or responded to was less likely to be circulated, suggesting that engagement is a key mechanism through which the system sustains itself. A final finding is that

soyjak culture varies across different platforms. On imageboards and forums, communication is more complex and highly structured, while on mainstream platforms such as Twitter, Tiktok, and Instagram, elements are often simplified or partially stripped of their original context. Despite this adaptation, core features such as visual variant recognition and rating systems remain intact suggesting that the underlying structure is resilient even when partially detached from its original environment.

Discussion/Analysis

The findings suggest that soyjak culture operates as a self contained semiotic system in which meaning is not fixed within individual memes but constructed through shared rules of interpretation. Rather than functioning as isolated memes or slang terms, the system integrates visual classification, linguistic modification, and participatory interaction into a unified communicative framework. This supports the hypothesis that meaning within the community is constructed through shared customs that reinforce group identity while limiting accessibility to outsiders. An additional pattern that emerged in relation to in-group and out-group distinction is the emergence of labels used to distinguish between perceived insiders and outsiders. One such term observed in the data is "usois," which is used to refer to newer or external users who discovered soyjak culture through mainstream platforms rather than through imageboards where the memes originated. This distinction highlights how accessibility plays a role in shaping community boundaries. Users within more established spaces often frame knowledge of internal systems such as soyspeak, variants, and slangs as markers of legitimacy, while outsiders are positioned as lacking the contextual knowledge required to interpret or participate meaningfully. As a result, identity formation within the community is not only based on participation but also

on the perceived authenticity of how that participation begins. However, the data also suggests that exclusion is not always intentional, instead it emerges naturally from the increasing complexity of the system over time. The presence of structured language rules in Soyspeak aligns with theories of digital sociolect formation, where online communities develop internal linguistic systems that evolve independently from mainstream language norms. Similarly, the soyjak variant classification system demonstrates that memes can function as categories rather than purely expressive images. One of the most significant implications of these findings is that meme communities can evolve into highly organized cultural systems with their own internal logic. This challenges the assumption that internet memes are inherently chaotic or unstructured. However, the system also depends heavily on participation. Without continuous user modification, classification, and interaction, the structure would lose its coherence. This indicates that identity within the community is not only communicated but actively produced through engagement.

Limitations

This study has several limitations that should be considered when interpreting its findings. First, the research relies entirely on publicly available content rather than direct interaction with participants. While this approach is appropriate for digital ethnographic research, it limits the ability to confirm how users themselves interpret or understand the systems being analyzed. As a result, interpretations of meaning in Soyspeak, soyjak variants, and participatory behaviors are based on observable patterns rather than verified user intent, which may introduce some degree of interpretive bias. A second limitation is the lack of demographic verification for users within the sampled data. Although prior descriptions of soyjak

communities suggest a predominantly teenage userbase, this study cannot independently confirm the age, background, or identity of individuals producing the content. This makes it difficult to generalize findings beyond the observed material or draw conclusions about how identity characteristics may influence participation in the community. Additionally, the sampling method may still introduce selection bias. The process of manually selecting posts, threads, and images means that the dataset may overrepresent more popular examples of soyjak culture while underrepresenting less documented interactions. This could lead to an incomplete picture of how communication functions across the full range of the community. Another limitation is the rapidly evolving nature of internet meme cultures. Soyjak related content changes quickly across platforms, meaning that the data collected represents only a snapshot in time. As a result, some patterns identified in this study may disappear or shift into new forms after the observation period. Finally, because the analysis focuses on identifying patterns rather than measuring frequency, the findings are interpretive and qualitative in nature. While thematic analysis allows for a deeper exploration of meaning, it does not establish causation or statistical significance. Therefore, conclusions should be understood as descriptive insights into communication structures rather than definitive or universally generalizable claims.

Conclusion

Overall, this study demonstrates that soyjak culture functions as a structured system of in-group communication that integrates visual classification, linguistics, and interaction between users. These elements operate together to form a coherent system through which meaning and identity are constructed. The findings suggest that online meme communities can develop

complex internal systems that resemble linguistic and cultural structures, reinforcing group identity while simultaneously shaping how information is interpreted and shaped.

Future Directions

While this study provides insight into the internal structure of soyjak communication systems, there are several directions for future research that could expand or deepen these findings. One potential direction is a comparative analysis between soyjak culture and other meme based communities that also rely on structured formats or recurring templates. Examining whether similar linguistic or visual classification systems emerge in different online subcultures could help determine whether this level of structure is unique or part of a broader pattern in internet culture. Another area for future research involves the role of platform migration and mainstream exposure. Since soyjak content appears across both niche imageboards and mainstream platforms such as Tiktok, Twitter, and Instagram, future research could investigate how meaning changes as content moves between these environments. Specifically, it would be valuable to analyze whether systems like Soyspeak and variant classification become simplified, distorted, or reinterpreted when they are removed from their original context. Future studies could also explore the process of learning and entry into such communities. This would involve examining how new users acquire knowledge of internal systems and how long it takes to develop fluency in both linguistic and visual components. Such research could help clarify whether in-group status is achieved gradually through participation or whether it depends on exposure to specific entry points within the community. Additionally, a more participatory method could provide deeper insight into user interpretation and intent. While this study focused on observable patterns in publicly available content, direct engagement with community

members could help clarify how users themselves understand systems like Soyspeak, variant classification, and rating structures. This would help address one of the main limitations of the present study, which is the reliance on interpretation rather than firsthand explanation. Finally, future research could examine the sustainability of these systems over time. Given the rapidly changing nature of internet meme culture, it would be valuable to study whether structured symptoms like those found in soyjak communities persists, evolve, or disappear as platforms and user demographics shift. This would contribute to a broader understanding of how digital cultures remain coherent in environments that face constant change. In conclusion, soyjak culture provides a compelling example of how online communities can develop complex, self-reinforcing systems of communication that combine language, imagery, and participation into a unified structure. Further research into similar online communities and their evolution would deepen our understanding of how meaning is constructed and sustained in digital environments.

Bibliography

- Crystal, D. (2011). Internet linguistics: A student guide. *Language Learning & Technology*, 15(3), 1–3. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4012002/>
- Boyd, d. (2014). *It's complicated: The social lives of networked teens*. Yale University Press.
<https://www.danah.org/books/ItsComplicated.pdf>
- EBSCO Information Services. (n.d.). Participatory culture.
<https://www.ebsco.com/research-starters/communication-and-mass-media/participatory-culture>
- Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation.
<https://web.stanford.edu/~eckert/Courses/ParisPapers/EckertInPress.pdf>
- Jenkins, H. (2014, February 26). A meme is a terrible thing to waste: An interview with Limor Shifman (Part one).
<https://henryjenkins.org/blog/2014/02/a-meme-is-a-terrible-thing-to-waste-an-interview-with-limor-shifman-part-one.html>
- Markham, A., & Buchanan, E. (2012). Ethical decision-making and internet research: Recommendations from the AoIR ethics working committee (Version 2.0). Association of Internet Researchers. <https://www.aoir.org/reports/ethics2.pdf>
- Petrova, K. (2021). Meme language: Its impact on digital culture and collective thinking.
https://www.researchgate.net/publication/352646607_Meme_language_its_impact_on_digital_culture_and_collective_thinking

