Title	Technological Bleed-through: The Transformative Effects of Technology on Text
Author(s)	CHRISTOPHER MUELLER
Citation	大阪学院大学　国際学論集（INTERNATIONAL STUDIES），第 34 巻第 1・2 号：53-67
Issue Date	2023.12.31
Resource Type	Article/ 論説
Resource Version	
URL	
Right	
Additional Information	

Technological Bleed-through:
The Transformative Effects of Technology on Text

CHRISTOPHER MUELLER

ABSTRACT

As technology progresses, communication has become an increasingly online affair, and language is adjusting itself to match.   Social and workplace interactions are often constructed, written, and shared via technological means, be it keyboard, website, keypad, or text chat.   This paper presents research from a variety of languages, contexts, and writing systems to show that this technological progress has altered not only the more obvious linguistic spheres, such as vocabulary and morphology, but also infiltrated the actual text itself.   The technology we use to communicate online is blending and merging with the writing systems which use it, forming creative new ways of writing.   Examples are drawn from research in English, Arabic, Japanese, and Taiwanese in an effort to represent the universality of this phenomenon, as well as bring attention to scripts which are less represented in CMC literature.

Introduction

If culture can be defined, as Goodenough（1957）and Wardaugh（2010） suggest, as "whatever a person must know to function within a particular society"（p. 229）, then technology has certainly established itself as an integral aspect of modern culture.　Social media, computers, and smartphones are becoming increasingly widespread,[1] and as a result, vocabulary and expressions with roots in these technologies,（e.g., *selfie*, *hashtag*, *download*, etc.）have rapidly filtered into our daily lexicons.　Technology-infused language has begun to color much of our communication, be it through metaphor（e.g., *computer virus*）, instances of so-called netspeak or chatspeak （e.g., *u2*, *LOL*, *tysm*）, or even the sudden proliferation of the *e-* prefix（e.g., *e-mail*, *e-Arabic*）（Crystal, 2001; Dauodi, 2011）.　Such linguistic phenomena have not gone unrecognized, historically.　Sapir（1929）and Worf's（Carroll, 1956）seminal theory of linguistic relativity, for example, posits that there is some link between languages, their content and structures, and how speakers of these languages process and understand the world around them.　Using this theory as a locus of analysis, and considering the apparent interrelations between technology and culture, intriguing questions begin to develop. Specifically, if language can serve as a reflection of culture, then to what degree is language itself directly affected by technology?

A growing number of authors have made efforts to tackle these and similar questions, much of the research coalescing under the broader umbrella of computer-mediated communication（CMC）.　Until more recently, however, the bulk of CMC literature was "focused almost exclusively on emergent practices in English"（Danet & Herring, 2003, para. 1）, neglecting instances of online communication in other languages.　While there has been some intriguing research done on differing scripts（Daoudi, 2011; De Oliveira, 2003; Hård af Segerstad, 2002）, there remains a noticeable lack of attention placed on languages which do not make use of the Roman alphabet（e.g., Chinese, Thai, Arabic, etc.）.　These languages have quite literally billions of users in combination, so how is it they have avoided a wider representation in the literature?

1）It is now estimated that around 60% of the world's population are regular internet users, for example（Internet World Stats, n.d.-b）.

The reasons for this are likely twofold.　Firstly, the text-transmission protocol used on the internet is based on the ASCII character set.　ASCII is an acronym for American Standard Code for Information Interchange and was established in the 1960s in the United States（Danet & Herring, 2007）.　As the name suggests, the character set is based on the Roman alphabet and the phonemic inventory of the English language.　The most portable and basic textual input of email and chat, called "plain text", contains only unformatted ASCII characters.　Though this is slowly changing, and non-standard forms of computer input are becoming more available, the language of online communication（and computation in general）remains, on the whole, an English one.　Secondly, most CMC takes place online, and until recently, the vast majority of internet-connected users were from the United States.　In 2003, for example, the United States made up 36.5% of internet users worldwide, compared to the second-ranked China（10.9%）and third-ranked Japan（9.7%）（Nishimura, 2003）.　Though this has changed in more recent years（China and India have since overtaken the US）, for the majority internet's nascence, its users were primarily English speakers（Internet World Stats, n.d.-a）[2].

At this point, it may also be useful to further define "online communication".　"Online" is perhaps something of a misnomer, as this form of communication need not be purely internet-based（Danet & Herring, 2003）.　Short Message Service（SMS）（i.e., text messaging）, mobile chat applications（e.g., Whastapp, Skype）, and the like all fall within this category.　Much of it does include different varieties of communication on the internet, however, including bulletin board systems（BBS）, social media sites such as Reddit or Facebook, or text chat applications such as Instant Messenger or Internet Relay Chat（IRC）.

While there have been many important articles and studies within CMC regarding the intersection of technology and language, much of this research has focused on documenting and categorizing the emergent jargon（Crystal, 2001; Squires, 2010）or the unique hybridization（or 'in-between-ness'）of the spoken and written word in online communication（Ferrara, Brunner, &

---

2) Additionally, while the number of internet users in non-English speaking countries has
　risen significantly, as of October 2022, 60.9% of the top 10 million websites on the
　world wide web are in English（W3Techs, n.d.）.

Whitmore, 1991; Herring, 1996, Section 1）． There has been far less attention paid, however, to scripts themselves（the alphabets, syllabaries, and logograms）, the devices used to input these scripts（keyboard, keypad）, and how they interact． This article, then, will attempt to collect and synthesize some of the currently available research available in this area, drawing attention to the various research gaps present concerning the transformative effect of technology on the constructed elements of language itself, with a secondary focus on writing systems which do not use the Roman alphabet（i.e., the alphabet of ASCII）． Of specific interest are instances of technological 'bleed-through', where word-on-page（or more aptly 'word-on-screen'）text is directly influenced or altered by the underlying technology． This should also serve as visible, concrete evidence for a Whorfian view of language, quite literally shaped by its cultural（and by extension, technological）context.

LEET Speak

One particularly cogent area of exploration is that of computer gaming, specifically massively multiplayer online role-playing games（MMORPGs）. MMORPGs require an online connection and maintain a persistent online world meant to be explored with other players． A popular example is World of Warcraft, which at its height boasted 12 million subscribers（Blizzard Entertainment Inc., 2010）． Interaction between players is encouraged（and often required）to successfully complete in-game tasks, and many tools for socialization, such community hubs and user-organized guilds, are present within the game． Communication is accomplished primarily through keyboard-based text chat, and as such, online gaming culture is filled with examples of this splicing of technology and language． In a 2005 article, for example, Blashki and Nichol investigated "LEET speak" or "1337 5p34k", a form of language "specifically developed for, and by, a community of young people who play computer games"（p. 77）． Again, since communication in the game is usually entered via keyboard and computation has a numerical aspect（hexadecimal/binary/etc.）, numbers are often used as visual analogs for letters． The number "1337", then, would correspond to "LEET", a truncation of "elite".

Following the research of Beavis（2005）and Gee（2003）, the authors set out to discover how language served to define a specific cultural group, i.e.,

"gamer geeks", and vice versa.　The project focused on a group of forty first-year students in a tertiary education program.　Participants self-selected from a large course on gaming fundamentals, and data were collected from an online discussion board prepared for the course.　Both the researchers and participants contributed to the online forum discussions, which were centered around leet speak.　Many examples of different features and intricacies within leet speak were collected, including, as in the following sentence, a lack of standard spelling or grammar rules:

"1|=\| 0u |{4N r34|）t|-|15t|-|3 N\| 0u i5 t3|-| |_337"

（Translation: 'If you kan read this then you is teh leet'）（Blashki & Nichol, 2005, p. 83）

Also of note, the word "teh" here is not a true misspelling.　In multiplayer gaming, communication between individuals often takes place during competitive and high-action gameplay.　As such, typographical errors are not uncommon.　In many cases, however, these 'typos' have entered the lexicon and become the default.　The article "the" is one example of this, often intentionally written as "teh"（Tavosanis, 2007）.　Another common case would be the word "pwned", which can represent the word "owned".　The 'P' and 'O' keys are located side-by side on most English-language（QWERTY）keyboards, allowing for such an error.　In these instances, the method of input（i.e., keyboard）in concert with the technological medium（i.e., computer gaming）has altered certain spelling conventions.　The devices used to compose the text, the placement of the keys and the position of the fingers, have created a technology-specific spelling for this community of users; it would not exist were it not for the specific technologies used.

Blashki and Nichol（2005）were also interested in how leet speak had come to be, and how its users perceived the language.　While interpretations varied, one prevailing theory was that leet speak originated in an attempt, as one of the participants stated, "to stay under the radar when it came to pirating"（p. 81）, as search engines were unable to process text written in such a manner.　In this way, the edited and purposely malformed text, combining numbers, letters, and punctuation, came to be as a response to（and circumvention of）the technological medium in which it was used.　It also 'looks' geeky, visually incorporating numbers and symbols like those used in programming languages and text chat.　Indeed, there was a general consensus among both participants and researchers that leet speak also performed an

affiliative function, reinforcing social bonds within a specific outsider group
(i.e., "geeks"). Leet speak, then, "evolved from the game geek's need to
express and communicate within a supportive community of like-minded
participants" (p. 85). The technological foundations of both the computer
game genre and piracy/hacking communities amalgamated to produce a form
of English which could address both the technical and social concerns within
geek/gamer culture.

Digital Sarcasm

Another interesting example of such an amalgamation involves the
production and perception of sarcasm in online environments (Greenwood &
Maynard, 2014). Two researchers from the United Kingdom attempted to
create a software program which could accurately detect sarcasm on the social
media service Twitter. While sarcasm in speech is detectable by changes in
prosody, timing, posture, or facial expression (Attardo, Eisterhold, Hay, &
Poggit, 2003), when sarcastic utterances are processed digitally they become
much more difficult to distinguish (Kunneman, Liebrecht, & van den Bosch,
2013). As one way of compensating for this, the hashtag symbol (or '#')
developed as an indicator of sarcastic tone.

The hashtag symbol's current role in online communication is rooted in
an earlier technology. While it was originally used as a symbol to designate
a number (e.g., #1), with the advent of online chatrooms, specifically IRC,
the hashtag became a method of indicating interest groups and topics (e.g.,
#comicbooks) (Channel Scope, n.d.). Its usage has since extended beyond
online chatrooms to platforms such as Instagram and Twitter. It now
functions as a metadata tag, transforming a word or phrase into a clickable
link which then leads to additional groups of posts including the same
hashtagged word(s) (Twitter Inc., n.d.). In their study, Greenwood and
Maynard (2014) collected a corpus of 134 Twitter posts (or 'tweets')
containing the phrase "#sarcasm". They discovered that this expression was
frequently used to reverse the polarity of the preceding phrase. In other
words, the hashtag served to 'tag' an utterance as sarcastic in tone. One
sample from their data is the following sentence:

"You are really mature. #lying #sarcasm" (p. 4239)

Without the added tag(s), the intended meaning is lost. Thus,

hashtagged expressions function as a replacement for the ancillary or paralinguistic elements which would be present in 'offline' speech. The research suggests that the hashtag has become an integral part of the utterance, as it cannot be correctly parsed without it. In this way, technological features have become somehow fused with the semantics of the language, while simultaneously incorporating technological elements（i.e., the metadata tags）into the text itself. The technology is thus directly reflected in the text, and it remains usable（clickable）as a technology. In other words, the utterance is itself a form of technology.

Additionally, tagged expressions are not limited to the word "sarcasm" or even any single word, as hundreds of instances of "#notreally" or "#irony" were also discovered（2014）. Neither is this a purely English phenomenon; Dutch researchers（Kunneman et al., 2013）performed a similar study and came to similar conclusions regarding the role of hashtags. It is also interesting to note that unlike in the previous studies, this specific linguistic feature was not borne out of technology's facilitation（i.e., the advent of computer gaming）, but rather from its failure（i.e., the insufficiency of computer interaction）. In short, written expressions are filling semantic gaps with what was originally a feature of online chatrooms.

In my own research（Mueller, 2016）, I have encountered a similar phenomenon. On Reddit, a social media and content-aggregating website, users can leave comments on posted submissions in BBS-style 'threads'. In many of these comment threads, users would end their sentences with "/s" to indicate a sarcastic or ironic tone[3]. Over 50 examples of /s-tagged comments were collected over a period of several months. A representative sentence from the data set can be seen here:

"That's a really valid point you made there! What a great comment! /s"（p. 89）

As in Greenwood and Maynard's（2014）findings, a 'tag' has been placed on the sentence to indicate the speaker's sarcastic intent. Instead of the '#' symbol, however, we have a different token: a scrap of pseudo-HTML code. The "/s" in this instance serves as an abbreviated version of "</s>" which is

---

3) These s/-tagged posts were not limited to Reddit, however, and can be readily seen in a variety of other online spaces such as discussion forums, blogging websites（e.g., Tumblr）, and Youtube comment sections（Mueller, 2016）.

itself a truncation of the HTML code "<s>TEXT</s>"（Mueller, 2016）. In the HTML coding language, "to initiate a certain function, such as italicizing text, a tag such as <i> is used. In turn, to end this tag, a corresponding </i> tag is used. In that sense <i> equates to 'begin italics' and </i> equates to 'end italics', with the text to be italicized nested in-between. In this manner, the </s> tag is functionally stating 'end sarcasm' or 'the sarcasm ends here'"（p. 96）. The two languages, one for communication（English）, and one for coding（HTML）, have joined here to perform a specific semantic function. The HTML code, the very technology used to construct the website, is "seeping out" into the text itself, becoming a part of the utterance.

Syllabaries and Logograms

Though far less research exists on languages which do not make use of the Roman alphabet, some interesting studies have still surfaced. One such study（Nishimura, 2003）focuses on the innovative use by Japanese speakers of *kanji* and other scripts on internet BBS-style message boards. Nishimura grounds her study in previous research on online communication in Japan （e.g., Katsuno & Yano, 2002）, but is unique in her focus on how the online medium has changed not only the message, but the actual script itself. Her data were collected from comments on Japanese fan-websites featuring popular actors, films, and music. Nishimura noticed that writing styles on such websites were markedly different from Japanese as it was traditionally written. One such feature was the interspersion of English letters（known as *romanji* script）and words with traditional Japanese *kanji*, *hiragana*, and/or *katakana* script. One possible reason for this is that Japanese keyboards did not function in a similar way to English keyboards. To write in Japanese on certain computers, "users must use so-called *romaji-kana-kanji* conversion software like Microsoft IME or Justsytem's ATOK"（para. 25）. In these cases, text must first be entered in *romanji*, the Roman alphabet, based on their Japanese pronunciation. The software then searches for a corresponding word in *hiragana* and converts the *romanji* into the *hiragana* syllabary. If the desired word is in *katakana*, a "*katakana* key" is pressed; however, if the desired word is in *kanji*, then a "conversion key" must be pressed, and a menu of recently used *kanji* symbols which correspond with the *hiragana* word are displayed. The user may then select the desired symbol. The system is

complex, and potentially time consuming; it is often simpler to default to the English character, as it is the first point of input.　Moreover, in cases where a *katakana* word would be preferred or more correct, *hiragana* may be retained "for the purpose of speeding up typing"（para. 82）.　The keyboard input, in combination with the available software, has thus had a profound effect on the choice of script used.

　　Nishimura（2003）also found that in online communication, kanji verb stems would often appear isolated between parentheses.　When used as verbs, *kanji* "are not normally used in isolation; they are used with *hiragana* for inflectional endings, tense, and so on"（para. 73）.　In these cases, *kanji* were being used between parentheses to indicate an action, for example（涙）, *namida*, or "tears", would be used to indicate that the speaker was sad or crying（para. 74）.　This usage was specific to online communication.　Users were detaching kanji from their traditional grammatical structures（while adding symbols available on their keyboards）for quickness and ease of use.

　　In addition to this, Japanese speakers communicating online were making use of a system of letters-for-numbers reminiscent of the previously discussed "leet speak".　The sequence of numbers "4649" in the Japanese, for example, can represent the word "*yoroshiku*".　In Japanese, the number 4 is pronounced *yon* or *shi*, 6, *roku*, and 9, *ku*.　The four sounds combine to produce *yoroshiku*（para. 78）.　As in leet speak, numbers here are serving as replacements for letters（or more precisely, syllables）.　Unlike leet speak, however, this system is based on pronunciation of the numbers rather than their visual characteristics.　Here, once again, there is evidence of serious linguistic changes which are endemic to the electronic medium in which they are consumed and produced; the language has been altered to facilitate the technology used（keyboard/computer）and the technological context（online message boards）.

　　A second study of a non-Roman writing system involves the concept of "playfulness" online in speakers of Taiwanese（Su, 2009）.　The author focuses on two specific languages utilized in Taiwanese online communication, namely Stylized Taiwanese（ST）and Stylized Taiwanese-accented Mandarin（STM）.　Her study examines the creative use of these stylized forms on three recreational BBSs and a Web-based BBS belonging to two Taiwanese universities.　Both languages are used in similar ways in online communication.　Taiwanese does not have a standardized writing

system, so Chinese characters are used in both ST and STM（p. 319）.　When writing on the message boards, however, Su noticed that university students were using a novel format, replacing character-based understanding of the words with a more phonological approach.　An example of this from her data is the sentence "Who can compete with me?"（p. 325）.　The meaning- or character-based sentence would be constructed as follows:

　　誰人甲我比　　〔ʂei zən tɕia wo pi〕　　who, person, to-better, I/me,
　　　　　　　　　　　　　　　　　　　　　compete

However, what appeared in the BBS was quite different:

　　瞎郎尬挖畢　　〔ɕia laŋ ka wa pi〕　　blind, man, to-embarrass,
　　　　　　　　　　　　　　　　　　　　　to-dig, to complete（p. 325）

　　While the meaning of the first sentence would be clear to a reader, the second sentence, when viewed semantically, character-by-character, is gibberish.　Phonologically, however, the second sentence is in fact very similar to the first; said aloud, it would be readily intelligible.　This sound-based writing style was unique to the online community, and "could be hardly decipherable to someone unfamiliar with the BBS"（p. 324）.　The author equates this manipulation of the writing system to "a token of group solidarity between BBS users"（p. 324）, but it is clear that it is also a product of the BBS environment, online and text-based; such a system of wordplay, for example, could not exist orally.

Arabizi

　　Another example of text-on-page transforming to suit technological advancement has recently appeared in written Arabic.　Many younger speakers of Arabic have foregone traditional Arabic script completely where technology is concerned（al Khalil & Palfreyman, 2010; Yaghan, 2008）.　In media such as texting and instant messaging, the Arabic alphabet is often unavailable（or is available but unwieldy）as until more recently, "the Arabic language was not supported by the widespread technology"（2008, p. 44）. Similar to the situation in Japan, the Roman alphabet is much more readily found on keyboards and keypads.　As such, a novel form of written Arabic has developed to compensate, using a combination of Roman（ASCII） characters and（Arabic）numerals.　If a specific Arabic phoneme has an equivalent in the Roman alphabet, the equivalent is used（e.g., *b* for ب）.　If

no equivalent is available, a specific number is used（e.g., 7 for ح）．For
example, خُمُص becomes *7ommos,* or خالد becomes *5aled.*

In a 2008 study, Yaghan performed extensive group interviews with
university students in Jordan in an effort to document this form of Arabic and
describe its rules and features. The slang term for the writing style is *Arabizi,*
a portmanteau combining the words *arabi*（Arabic）and *engliszi*（English）．
Yaghan discovered that the use of *Arabizi* was widespread among young
people throughout Arabic speaking countries, and that each country had their
own specific form（or forms）depending on the local dialect. Some
regularities, however, did surface from within the data. Vowels, for example,
are optional in Arabizi, and can be removed at the user's discretion. This was
a direct result of its origin within text messaging; parsimony was needed to
allow for a maximum number of characters. True to its name, *Arabizi* also
makes use of both English and Arabic, though the English is limited to
"common World Wide Web and cellular phone messages（"plz" for "please,"
"thnx" for "thanks," etc.）"（p. 42）．Stressed consonants are written twice to
indicate their emphasized status（e.g., *ennik*），and certain symbols are used as
affixes to indicate plurality（@）or the past tense（8）．

Yaghan was also interested in why young Arabic speakers were
gravitating so strongly towards *Arabizi.* Technological factors remained
valid, as some cellular phones did not have Arabic script capability, and many
that did allowed fewer characters. However, it was also expressed by the
students that "using one set of 'English' keys that can be applied to the two
languages is more convenient, and less confusing"（p. 45）．*Arabizi* was also
viewed as a more open, "freer" writing system than formal, standardized
Arabic（الفصحى or *fussha*）．Because it was not taught at any official level,
and is acquired solely through practice, *Arabizi* has more fluidity in its rules,
especially when compared to the rigidity of *fussha.* There is, for example,
nothing akin to 'typos', as words can be spelled in a variety of different ways
depending on the user's country, region, or individual preferences. The
author also observed that *Arabizi* was becoming increasingly common, and
had begun to appear outside of its original online context in the form of
handwritten notes, CD covers, movie posters, and even music videos（e.g.,
Nancy Ajram, 2013）．

Yaghan also notes that in the last two centuries, multiple explicit attempts
have been made to standardize Arabic using the Roman alphabet, the earliest

recorded example having occurred in 1880, and the most recent in 1943.　All such attempts have failed.　It is interesting that a standardized writing system seems to be in the process of generating organically.　The technological advances which served as the impetus for *Arabizi*'s development may have helped to create a fresh stage in the standardization of modern Arabic.

　　In addition to Yaghan's study, more and more articles on *Arabizi* have begun to appear.　Al Khalil & Palfreyman（2010）investigated patterns of use in young female students in the United Arab Emirates.　The authors were interested in how consistent the ASCII representations were, what influences shaped spelling choice, and what purposes *Arabizi* served for its users.　Using a combination of corpus data, surveys, and informal interviews, they arrived at similar conclusions to Yaghan（2008）, namely that the use of Arabizi was extremely flexible, informal, and "easier to type"（Al Khalil & Palfreyman, 2010, p. 17）.　As in Nishimura's（2003）study, a technology's ease of use was the impetus for a change in the writing used.

Conclusion

　　The studies in this paper represent research on a variety of writing systems, language backgrounds, and technological contexts.　They are common, however, in their representations of written language as subject to （and often inspired by）the flow of technological advancement.　In each of the languages discussed, written text has been affected visually, structurally, or even semantically by means of technological input（e.g., keyboards, cellphones）and/or technological environment（e.g., bulletin boards, social media）.　As technology has become ever more a part of the global day-to-day, some change to language is, of course, expected; new words become necessary and old words become obsolete.　I hope I have shown here, however, that the effects technology has on language run much deeper, and in many of these cases, even manage to infiltrate the appearance of the language itself.　Changes to sentence structure and grammar（e.g., Nishimura, 2003）, spelling（e.g., Yaghan, 2008）, word choice（e.g., Su, 2009）, and the incorporation of technological symbols into our writings（e.g., Greenwood & Maynard, 2014; Mueller, 2016）all stand as examples of how technology has managed to permeate the very text it is used to construct, to 'bleed through' into the written word.

Though continued research is undoubtedly needed, especially with regards to smaller languages and languages not compatible with the ASCII format, I believe that the（albeit limited）examples discussed above have served to show that language is indeed profoundly affected by the culture（and by extent, technology）in which it is suspended. But culture, language, and, perhaps most of all, technology are highly dynamic creatures, and research in these areas must be constantly refreshed lest it risk stagnation. To this end, studies in CMC, such as those included here, remain an important addition to language literature at large. In an increasingly 'online' world, these studies, and others like them, reveal that the technology we use to communicate is changing not only what we write, but how we write it.

References

Al Khalil, M., & Palfreyman, D.（2003）. 'A funky language for teenz to use': Representing Gulf Arabic in instant messaging. *Journal of Computer Mediated Communication, 9*, 1.

Attardo, S., Eisterhold, J., Hay, J., & Poggi, I.（2003）. Multimodal markers of irony and sarcasm. *HUMOR: International Journal of Humor Research*, *16*（2）, 243-260.

Beavis, C.（2005）. Pretty good for a girl: Gender, identity and computer games. In S. de Castell & J. Jenson（Eds.）, *Changing views: World in play*. Vancouver, British Columbia, Canada: University of Vancouver.

Blashki, K., & Nichol, S.（2005）. Game geek's goss: Linguistic creativity in young males within an online university forum（94/\/\3 933k'5 9055oneone）. *Australian Journal of Emerging Technologies and Society*, 3, 77-86.

Blizzard Entertainment Inc.（2010, October 7）. *World of Warcraft® subscriber base reaches 12 million worldwide*［Press release］. https://investor.activision.com/news-releases/news-release-details/world-warcraftr-subscriber-base-reaches-12-million-worldwide

Carroll, J. B.（ed.）（1956）. *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.

Channel Scope.（n.d.）. *Internet relay chat: Channel management*. Retrieved October 8, 2022, from http://tools.ietf.org/html/rfc2811

Crystal, D.（2001）. *Language and the internet*. Cambridge: Cambridge University Press.

Danet, B., & Herring, S.（2003）. Introduction: The multilingual internet. *Journal of Computer Mediated Communication, 9*, 1.

Danet, B., & Herring, S.（2007）. Multilingualism on the Internet. In M. Hellinger & A. Pauwels（Eds.）, *Handbook of language and communication: Diversity and change*（pp. 554-585）. New York: Mouton de Gruyter.

Daoudi, A.（2011）. Globalization, computer-mediated communications and the rise of e-Arabic. *Middle East Journal of Culture and Communication, 4 (2)*, 146-163.

De Oliveira, S.（2003）. Breaking conversational norms on a Portuguese users network: Men as adjudicators of politeness? *Journal of Computer Mediated Communication, 9*, 1.

Ferrara, K, Brunner, H., & Whittemore, G.（1991）. Interactive written discourse as an emergent register. *Written Communication, 8*, 8-34.

Gee, J.（2003）. *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Goodenough, W. H.（1957）. Cultural anthropology and linguistics. In P. L. Garvin（Ed.）, *Report of the seventh round table meeting on linguistics and language Study*. Washington, DC: Georgetown University Press.

Greenwood, M., Maynard, D.（2014）. Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. *LREC 2014*, 4238-4243.

Hård af Segerstad, Y.（2002）. Effects of mobile text messaging on the Swedish written language: Human adaptability made visible? In F. Sudweeks & C. Ess（Eds.）, *Proceedings of cultural attitudes towards communication and technology*（pp. 355-360）. Montréal, Canada: Université de Montréal.

Herring, S. C.（ed.）（1996）. *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives*. Amsterdam: John Benjamins.

Internet World Stats.（n.d.-a）. *Internet 2022 usage in Asia*. Retrieved October 3, 2022, from https://www.internetworldstats.com/stats3.htm

Internet World Stats.（n.d.-b）. *World Internet Users and 2022 Population Stats*. Retrieved October 3, 2022, from https://www.internetworldstats.com/stats.htm

Katsuno, H. & Yano, C.（2002）. Face to face: On-line subjectivity in contemporary Japan. *Asian Studies Review, 26*, 205-232.

Kunneman, F., Liebrecht, C., & van den Bosch, A.（2013）. The perfect solution for detecting sarcasm in tweets# not. *WASSA 2013*, 29.

Mueller, C.（2016）. Positive feedback loops: Sarcasm and the pseudo-argument in Reddit communities. *Working Papers in TESOL & Applied Linguistics*, *16（2）*, 84-97.

Nancy Ajram.（December 6, 2013）*Aamel aekla（official music video）*［Video］Youtube. https://www.youtube.com/watch?v=6mWD96Rx5Ds

Nishimura, Y.（2003）. Linguistic innovations and interactional features of casual online communication in Japanese. *Journal of Computer Mediated Communication, 9*, 1.

Sapir, E.（1929）. The Status of Linguistics as a Science. *Language, 5*, 207-214.

Squires, L.（2010）. Enregistering internet language. *Language in Society 39*, 457-492.

Su, H.（2009）. Reconstructing Taiwanese and Taiwan Guoyu on the Taiwan-based internet. *Journal of Asian Pacific Communication 19*（2）, 313-335.

Tavosanis, M.（2007, January 8）. A causal classification of orthography errors in web texts. *IJCAI-07 Workshop on Analytics for Noisy Unstructured Text Data（AND-07）*, 102.

Twitter Inc.（n.d.）. *How to use hashtags*. Retrieved October 10, 2022, from https://help.twitter.com/en/using-twitter/how-to-use-hashtags

Wardaugh, R.（2010）. *An introduction to sociolinguistics*. Sussex, UK: Wiley-Blackwell.

W3Techs.（n.d.）. *Usage statistics of content languages for websites*. Retrieved October 8, 2022, from https://w3techs.com/technologies/overview/content_language

Yaghan, M. A.（2008）. Arabizi: A contemporary style of Arabic slang. *Design Issues, vol. 24*（2）, 39-52.