ReCALL **16** (2): 308–329. © 2004 Cambridge University Press
DOI: 10.1017/S0958344004000527 Printed in the United Kingdom

How to chat in English and Chinese: Emerging digital language conventions

HEATHER LOTHERINGTON
Faculty of Ediucation, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada
(*e-mail: hlotherington@edu.yorku.ca*)

YEJUN XU
Department of Epidemiology and Biostatistics, McGill University, Purvis Hall, 1020 Pine Avenue West, Montreal, Quebec H3A 1A2, Canada
(*e-mail: yejunxu@yahoo.com*)

Abstract

Rapid changes in language form and function occurring in digital environments present teachers and students of second languages alike with conundrums as to language and discourse standards. Factors affecting the changes that are emerging in digital English include the spatial and temporal possibilities and constraints of the medium, digital facilitation of case-creativity and iconic incorporation, and new social network configurations. This paper analyzes evolving changes in orthographic, syntactic, discourse and sociocultural conventions occurring in English and Chinese in digital environments, based on a small scale study conducted at York University in 2002–2003, noting trends across these languages as well as more limited, culturally and linguistically specific evolutions. The converging conventional changes occurring in these two major world languages suggest that similar transitions are happening generally in languages used for online communication, which has serious implications for second language instruction.

1 Introduction

The smiley faces began to appear in student writing a few years ago, penciled apologetically on notes appended to late assignments. Then at a recent TESL conference, a respected colleague confessed that she had noticed her teenaged son peppering his academic work with the sort of unconventional spellings growing in online usage. Yet, at a major international conference dedicated to language learning and teaching through digital media, paper after paper treats the language of their web-based and digitized language courses according to established written standards.

The conservative standards of written language prescribed in language course books, whether digitized or not, have been radically reshaped in the conversations of digital

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

chat users. Crystal (2001) illustrates how Netspeak is a separate form of communication distinct from written or spoken communication. Given the inevitable coalescence of language conventions in youth competencies, which have already begun to cross over into academic domains, it is critical for teachers of languages to think about how language use and standards are changing online.

In a little over a decade, the English used in virtual space has radically altered the orthographic conventions that had been used in print for centuries with only minimal, superficial change; it has coined a growing lexicon that even computer spell-checkers cannot keep up with; and authored dictionaries of emoticons, e.g. NetLingo[1]. A number of factors have prompted changes in the form of English into what Crystal calls Netspeak, as well as communicative function, including:

- the ephemeral nature of online writing that mimics aspects of speech in both synchronous and asynchronous environments;
- the effects of medium possibilities and constraints on discourse in synchronous and asynchronous environments (Lotherington, 2001b; Werry, 1996) ;
- case-creativity in digital expression rooted historically in early software case insensitivities[2] (e.g. WordPerfect, iMac);
- ASCII keyboard possibilities, which facilitate the use of icons, abbreviations and emoticons;
- emerging social networks and the identities that form within them (Merchant, 2001; Hawisher & Selfe, 2000).

Prior research[3] indicates that Chinese, which uses logographic rather than alphabetic writing, is also changing in digital environments to include, for instance, obviously hybridized orthography (Gao, 2001; Lee, 2002).

This paper reports on the results of a study conducted at York University in 2002–2003[4] in which English and Chinese digital communication was analyzed for conventional language use patterns in selected digital environments. The paper examines emerging language conventions occurring in online conversations in two major world languages: English and Chinese. Though a detailed analysis of the pedagogical implications of emerging and changing language conventions in English and Chinese (and indeed any language being taught and learned as a second language today) is outside of the scope of this paper, the results of the analysis are instructive to all who are teaching, learning and using language in digitally-mediated environments, and particularly to teachers of languages, for whom they pose serious questions about prescriptive language standards, conventional norms, and frozen notions of communication as speech and writing.

1. www.netlingo.com
2. Russell Smith, column in the Globe and Mail (25/11/2000: R11)
3. The study, *Assessing public conceptions of literacy*, was conducted by myself assisted by Jean Shi Jing Xu, Franca Vani and Maria Carmen del Carrero de Salazar, and funded by a Faculty of Education Minor Research Grant.
4. We are indebted to the Faculty of Education for funding the project: *Digitization and language change* through a Minor Research Grant; and to the Graduate Program in Theoretical and Applied Linguistics in the Faculty of Arts for supporting this project through the generous allocation of an extended graduate assistantship.

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

2 Rewriting the frontiers of speech and writing

Speech encodes language through sounds that are systematized and organized to represent recognized structures of language, such as words and phrases. These segmentals are infused with prosodic features (e.g. pitch, contour, stress, volume, tempo) to express meaning. Identifying social variables such as age, sex, regional and social identity are signified through suprasegmentals in speech, such as pitch and stress. If speech is face-to-face, paralinguistic features, such as gestures, proxemics and facial expressions also contribute to the transmission of meaning.

In contemporary life speech is frequently disembodied through the agency of technology, such as the telephone or voice recording equipment, making paralinguistic information unavailable. Contemporary speech may also be synthesized, as on computer voice alerts, elevators, and digitized telephone responses, in which case, suprasegmental information may also be reduced, particularly above the level of the word, such as phrasal intonation and sentence stress.

Speech is no longer by nature ephemeral, given voice and image recording technology, but in face-to-face encounters we expect speech to be unrecorded, and thus, transitory. Writing is typically contrasted with speech in terms of permanence and, by implication, formality. However, there has always been a spectrum of permanence in writing from disposable memory aids, such as shopping lists to professionally published books.

Writing uses symbols to orthographically record either sound, as in alphabets and syllabaries, or meaning as in logographies. Writing has tended to be more formal in conventional usage as it can be edited. However it is also more limited in conveying meaning and, especially, sentiment, as it lacks the expressive complexity of speech. Punctuation is conventionally used in writing to clarify meaning by indicating syntactic information, such as phrasing (, ; :), interrogative intent (?), and direct speech (“”); or emotional reaction, such as surprise (!), analogous as far as writing is capable to suprasegmental speech information. However, punctuation can be easily shanghaied as a scan of real estate marketing language indicates:

> South exposure loft!
>
> Trendy King West Lofts! Magnificent view! Beautiful, two bedroom unit in the upcoming area with 2 full baths! Shows like a model! Wood floors, quality broadloom, mirrored closet doors, open concept kitchen, top of the line appliances, balcony! Appliances and parking included in the price! (*Real Estate Exchange,* 2003, 3 [4]: 22)

In an early linguistic analysis of the language used in digital bulletin board systems, conducted in the early 1990s, Collot & Belmore (1996: 21) noted that:

> Electronic Language displays some of the linguistic features which have been associated with certain forms of written language, and others which are more closely associated with spoken language.

Warschauer’s (1999) pedagogical study of electronic literacies views digital language as

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

melting the borders of speech and writing, noting that "[t]he historical divide between speech and writing has been overcome with the interactional and reflective aspects of language merged in a single medium" (op. cit.: 6). Crystal (2001), on the other hand, in his study of language and the Internet, which comparatively analyzes digital language against speech and writing, concludes that Netspeak is a functionally separate mode from either speech or writing. Randall (2002: 12) pulls together these views:

> The Internet has given us a new medium in which to communicate, and it has demanded of us that we develop a new process for doing so – a speaking-writing hybrid.

Digital chat is a fusion of speech and writing in that it utilizes a literate interface to create a conversation, which may be treated as ephemeral or may be "saved" in print form. Online language use is certainly less formal than that used in edited printed texts, though in asynchronous mode, it can be edited both manually and electronically by programmable spellcheckers and grammarcheckers. Digitization also enables iconic representation, adding a paralinguistic dimension to conventional orthographies.

3 Characteristics of chat shorthand

A factor that complicates the analysis of language use online is the variety of digital environments in which communication can take place. In synchronous environments, such as chat rooms and online instant messaging systems (e.g. MSN, ICQ), both/all communicators are present in real time, and conversational flow is affected by limits to turn-taking in both space and time. On the other hand, in asynchronous environments, such as email, bulletin boards, group conference folders, and listservs, participants are not necessarily present, so discourse is not spatially and temporally constrained in the same way.

Chat shorthand refers to the abbreviated language used in synchronous digital chat environments. As Merchant explains:

> Chatroom interactions are essentially rapid written conversation. They combine features of face-to-face talk with explorations in interactive writing. Turn-taking is determined by the speed of electronic transmission as well as the speed of participants' responses. Interruption and overlapping are not possible, but some typographic devices are used to create paralinguistic effects...(2001: 300)

Randall (2002: 31) found that use of abbreviations can "speed up the typing of messages and function as a symbol of group belonging". According to Lee (2002: 10) "shortenings may be used to indicate familiarity and intimacy between users" as well as saving typing time.

Since onliners do not communicate with each other face to face, they cannot "take advantage of non-linguistic cues such as facial expressions, intonations, or gestures" (Gao, 2001: 18). Accordingly, devices have been developed to simulate the paralinguistic features of oral communication. Capitalization and punctuation are used widely for stress (Gao, 2001). Emoticons, e.g. ;-) (winking face), are employed to add emotional information.

Thus, chat shorthand is typified not only by abbreviated forms but also by

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

extralinguistic symbols used as emoticons. Some servers provide a preset emoticon configuration (E#24[5]). Recent word processing software[6] translates common smileys such as :) into the ☺ symbol through their autocorrection function. Reference guides to chat shorthand are now appearing, such as www.netlingo.com, which has documented smileys and emoticons in a dictionary format. A Canadian online translation service transL8it.com offers to help make sense of the confusing proliferation of shorthand.

4 Methodology

In the fall of 2002 and the early winter of 2003, the authors embarked on a small scale study to investigate the changing nature of language conventions in digital environments. Online chats were contributed by study participants, who included native online users of English, native online users of Chinese, and bilingual Chinese-English online users. Participants were regular users of a variety of synchronous chat fora.

Data were collected in English and Chinese. Our data set comprised a number of preserved chats and emails from three distinct communities:

1. A small group of high school students, physically located across Toronto, aged 13–14, using MSN.
2. A pair of female undergraduate students in their early twenties, studying at a large urban Canadian university, using AIM.
3. A group of young professionals located in China and in continental North America, using OICQ.

Based on preliminary analysis of the synchronous chat data collected from online users in English and Chinese in the late fall and early winter, we wrote, piloted and electronically distributed a questionnaire on language convention preferences to faculty and students in selected departments at our university through conference and listserv facilities, and to our informants on AIM, MSN and OICQ, with an invitation to forward the questionnaire to other users chatting in English and/or Chinese. The questionnaire was available in both English and Chinese.

Though Internet chat data were collected from interlocutors who we had approached for permission, we realized that collecting language data in virtual space begs many ethical questions. Just as the issue of copyright is thrown into chaos with new texts in virtual space that are published and accessed outside the regulated mechanisms of capitalist modern era text publication; so, too, Internet conversations in chat rooms are more diffuse, anonymous and less regulated than face-to-face encounters. Given that multiple people engage in digital chats using pseudonyms, and that a single individual may use a number of pseudonyms that can change or be rotated, particularly common with young teenagers, anonymity was truly protected.

5. The questionnaire respondents were identified as follows: E = English; C1= Chinese written in English; C2=Chinese written in Mandarin; #N indicates number assigned to respondent within these three categories.
6. MicroSoft Word, for instance, automatically interprets :) as ☺

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

5 Analytical framework

The data analysis is based on the online chat data and 54 questionnaire responses: 27 in English (E) and 27 in Chinese [14 written in English (C1), and 13 written in Chinese (C2)]. The age range of questionnaire respondents was 14 to 62. Nineteen males and thirty-five females responded.

Our analytical framework for investigating chat emerged from a basic set of categories, which was informed by previous research and continually revised according to the data received in both languages. The analysis focused particularly on code changes but it also considered discourse conventions.

In his study of linguistic innovation in the Internet chat of British teenage girls, Merchant (2001: 301–2) locates in chat shorthand innovation in:

- punctuation;
- formality;
- spelling;
- abbreviations :
  - "the use of non-alphabetic characters to construct icons – usually relating to emotions (hence the coinage: emoticon);
  - simple abbreviations based on initials, e.g. *swalk* = sealed with a loving kiss;
  - combinations of numbers and letters "to create an approximate phonetic rendering of the message", e.g. *gr8* = great; and
  - phonetic spelling.

Crystal's examination of digital language yielded a number of observations about language use in synchronous chats (2001: 157– 65):

- participant overlap in chats;
- use of nicknames ("nicks");
- visually distinctive features:
  - smileys, e.g. :-) ;
  - rebus-like abbreviations/ colloquial elisions, e.g. *r* = are;
  - transcription of emotional noises, e.g. *hehehe*;
  - comic-book style interjections, e.g. *ugh*;
  - unusual combination of symbols, e.g. *TwoHands*;
  - innovative punctuation;
  - frequent typos;
- non-standard grammar.

In her study of online communication in Hong Kong, Lee (2002: 9–10) builds on Crystal (2001) to identify various linguistic shortenings, which "may be used to indicate familiarity and intimacy between users, so as to facilitate fast typing within a short period of time":

- acronyms;
- initialisms;

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

- acronym of sentence, e.g. *gtg* = got to go;
- letter homophone, e.g. *u* (you);
- number homophone, e.g. *2* = too; *99* = Nite nite (Cantonese);
- combination of letter and number homophone, e.g. *b4* = before;
- reduction of individual word, e.g. *cuz* = because;
- combination of letter initial and letter homophone, e.g. *oic* = oh I see.

Our framework for investigating and reporting on emerging orthographic and syntactic conventional use in digital environments (see Table 1) was based on these analyses of linguistic innovations in chat shorthand, and subject to continual revision according to the data gathered.

In addition to orthographic and syntactic changes, we noted many innovations in discourse and sociocultural conventions. Observations included code-switching and mixing, and orthographic mixing patterns; "writing" a spoken vernacular; and variation in concordance and parallelism in conventional usage patterns across conversational partners, as well as startling innovation in identity marking. Our description of these emerging conventions in online language use in English and Chinese follows.

6 Analysis of chat data

Innovations proliferated in the digital chat conventions of both languages, many of which were language or code specific. It was evident in both chat and questionnaire data that the young teenagers in our study (13–14 year olds) used far more innovation than

Table 1 *Analytical framework for emerging orthographic and syntactic conventions*

Orthographic conventions
spelling innovations
acronyms
emoticons
abbreviations
hybridized codes
innovations in capitalization
homophones
character homophones
numeric homophones
numeric quasi-homophones
onomatopoeia
variant stylistic spelling and synonymy
Syntactic conventions
sentence/utterance shape
sentential acronyms
holophrastic expressions
alphanumeric phrases
extended phrases
sentence mechanics
punctuation
capitalization

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

older users on balance. Indeed, the older the user, the less innovative the language they tended to use. However, there was variation in the degree of innovation within a single chat. Our data contain interestingly unparallel conversations in which one conversational partner used relatively standard language and his/her interlocutor/s, more innovative language, clearly indicating that individual discretion is a factor in innovation alongside medium-inspired latitude. Indeed perhaps what digital media inspire is the latitude to be innovative without reprisal rather than a particular conventional standard.

According to our data, digital language use occurs within a broad spectrum of conventional use ranging from standard print conventions at the conservative end to highly original innovation at the other (see Figure 1).

Discussion of linguistic innovations follows the framework outlined in Table 1.

6.1 Orthographic conventions

Orthographic innovation was a primary interest. Previous research had indicated incipient hybridization in Chinese orthography online. The degree of innovation in English spelling is one of the first observations made about online usage, with emoticons and acronyms peppering email and chat, especially of the younger generation.

The degree of inventiveness within orthographic conventions in our data was very high. Spelling worked on a spectrum as illustrated in Figure 2.

We found the following innovations in English spelling and Chinese writing:

Acronyms

Hypothesized to be speed-driven shorthand, phrasal and even sentential acronyms are written as words online. This practice predates digitization; for example, *scuba* is an acronym for 'self-contained underwater breathing apparatus'. Many online acronyms are now very common, e.g. lol (laugh out loud); btw (by the way); ttyl (talk to you later). Also, acronyms taken from pre-digital business usage are included in respondents' questionnaire responses, e.g. na (not applicable), fyi (for your information). Examples of acronyms used across digital chat users include:

Conservative print-based standards	→	→	*Innovative screen-based usage*
standard ←	spelling	→	innovative
careful ↙		↘	careless
optimal ←	use of punctuation	→	minimal
conventional ↙		↘	creative
grammatical ←	use of capitalization	→	stylistic
text-based ↙		↘	speech-emulating
considered ←	discourse content	→	topical
edited ↙		↘	ephemeral

Fig. 1. Variation in language conventions in digital communication environments.

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

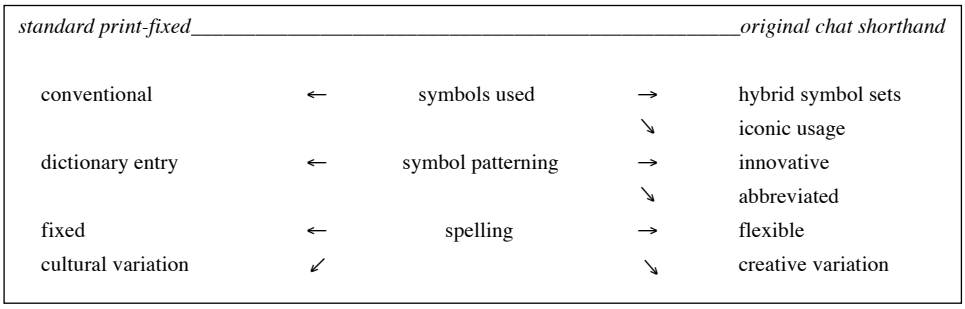


Fig. 2. Spelling conventions in digital communication environments.

lmao	laugh my ass off	(E#2; E#20; E#27)
brb	be right back	(E#4; E#5; E#26; E#27)
gtg or g2g	got to go	(E#5; E#26; E#27; Channah)
wb	welcome back	(E#4; E#26)

We asked about coinages: personally authored original acronyms and found synonymy with more frequently used acronyms (e.g. lmao):

lmpo	laugh my pants off	(E#2)
rotflmao	rolling on the floor laughing my ass off	(E#5)

Chinese acronyms exist as well. They are constructed from the initials of the Pinyin transliteration of characters[7]. Common acronyms include:

gg	ge^{1} ge	哥哥	(elder brother) (C1#3; C1#8; C2#2; C2#7)
mm	mei^{4} mei	妹妹	(younger sister) (C1#3; C1#8; C1#12; C2#2; C2#4; C2#7)
tmd	ta^{1} ma^{1} de	他妈的	(Chinese curse) (C1#12; C1#4; C2#4; C2#7; C2#12)

Emoticons

Smiley classics such as: ☺ and ☹ were endemic across both language groups, but a variety of smileys appeared in our data, including, for instance (glosses included are those of users):

;)	wink; just kidding	(E#1; E#2; E#9; E#23; E#26; E#27; C1#3)
:P	tongue hanging out	(C1#3; E#5)
:@	mad	(E#2; E#27)
:O	shock	(E#1; E#2; E#27)

Less commonly reported emoticons included:

OX	shut mouth	(E#4)
:O*	kiss	(E#4)

7. Tones are indicated with superscripted numbers.

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

:$	confused	(E#4; E#5)
%-l	stressed out	(E#23)

A fascinating finding confirming earlier observations is a variant of smileys in Chinese that does not occur in the English data. This variation is based on orientation of the emoticons. In English, emoticons are written and read perpendicular to the page orientation, e.g. ;-). In Chinese, some emoticons are created in this plane, but others are read head on, such as: (^_^). Though all emoticons are read from left to right, this orientation difference recalls the directionality differences in the writing of traditional script, which in Chinese moves from top to bottom and from right to left, and in English moves from left to right. Some examples of Chinese smileys include:

_	(C1#12; C2#6)
^Q^	(C2#1)
^..^	(C2#1; C2#2)
`	(C2#5)
v	(C2#7)

Abbreviations

The plentiful abbreviations in digital chat data are created in a number of ways:

Homophonic spellings

u	you	(E#13; Gary)
y	why	(Elizabeth)
r	are	(Clarence)

Truncated homophonic spellings

k	ok	(E#13; Leah & Channah; Elizabeth)

Borrowed shorthand

w	with	(E#18)
b/c	because	(E#18)
fyi	for your information	(E#10)
iou	I owe you	(C2#7)

Reduced spellings needing a gloss

ft	faint	(Chinese film reference)
wo	without	(E#18)

Simplified but recognizable spellings

nite	night	(E#18)
bakpak	backpack	(Leah & Channah)
plz	please	(E#26)
sed	said	(Gary)
g'nite	goodnight	(Channah)

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

Hybridized codes

Hybridization comes from a variety of keyboard sources: numeric, symbolic; and other orthographies. We found:

- alphanumeric rebus writing, with numbers functioning:
  1. as a "morpheme";
     sk8; l8r; gr8 (E#2; E#5; Channah)
     4got; b4 (Channah)
     2day; 2morrow; 2nite (Clarence; Elizabeth; E#27; E#50
     1derful (E#5)
     no1; every1 (E#2)

  2. and as a lexeme:
     4 u (C1#3)

- ASCII keyboard creativity
  ~ * ^ *HeAtH~ * ^ * (Elizabeth)

- Hybridized script usage in Chinese

Inclusion of Chinese characters, Japanese kana, and Roman letters, numerals and keyboard symbols in OICQ conversations, especially names:

のッ メRain さ
happy 李
protoss(趴趴熊)
洪兴の甩甩*兔兔Ⅲ
╰☆╮老√婆-也╲疯メ狂 ╰☆╮
╰★红狐狸卡→可爱靓靓 ╭★

Innovations in capitalization

Digital language use is increasingly case-creative. Indeed, the use of word medial capitalization brands hardware and software in the marketing world, e.g. WordPerfect, iMac. Capitalization may also reveal coinages of mixed heritage, e.g. WorldCALL, which creates a compound from a word and an acronym. Word punctuation and capitalization tend to be more or less optional in our data, and the same word may or may not include capitalization within the same chat. Proper nouns are frequently left in lower case, e.g. christina aguilera; judaism; tuesday (Channah - Leah)

Homophones

Character homophones

The Chinese writing system, used by a quarter of the world's population, has been in use since before 1,000BC (Taylor & Taylor, 1995). Chinese is written in a logographic script, i.e. the lexicon is the primary level of representation (Coulmas, 1996). There are six classes of Chinese characters, but 90% of Chinese is written using one class of characters: compounds including an element of meaning coupled with an element of sound (Coulmas, 1996). Given that Chinese is isolating and monosyllabic, there is substantial propensity for homophony.

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

In the Chinese data, homophones constituted a very important category. It is hypothesized that the use of "incorrect" homophonic characters, i.e. characters chosen on the basis of sound rather than meaning, is due to the more popular Chinese character input systems, based on Pinyin, which retrieve characters according to their phonological realization: Zhi4 Neng2 Shuang1 Pin1 (智能双拼), and Zhi4 Neng2 ABC (智能 ABC). According to the survey data, only one out of fourteen respondents used the radical-based Wu3 Bi3 Ai4 Xing2 (五笔字型) input system, which retrieves characters according to their meaning component.

Interestingly we located the use of many of these "wrong" characters: those chosen homophonically, rather than according to their denotative meaning, e.g.:

"incorrect" homophonic character	*correct character*	*meaning*
shou3 ji2 手即	shou3 ji1 手机	cellphone
fu4 zi3 miao4 父子庙	fu1 zi3 miao4 夫子庙	Confuscian Temple

Chat partners who confirmed they would use Chinese homophones, which sounded like what they wanted to say, rather than locating the correct written Chinese characters gave the following reasons:

1. The communicative goal is to make sense; these characters will do.
2. Laziness: using homophones is easier.
3. This is a time-saving convenience.

Numeric homophones

Numeric homophones are regularly used in English, e.g.:

2 = to; too
4 = for

More commonly, numeric homophonic spellings are used to represent part of a word, functioning as a unit of sound, e.g. sk8; b4; every1.

Numeric quasi-homophones

A highly original innovation found widely in our Chinese data but nowhere in the English language data is a category that works on the principle of approximate homophony – recognizable but not exact pronunciation (see also Lee, 2001). This trend is thought to have originated in the popularity of pagers some years ago, which had capability for displaying only numbers, not words. Hence:

Numeric homophone	*Pinyin*	*Character*	*Pinyin approximation*	*Character*	*Meaning*
520	wu3 er4 ling2	五二零	wo3 ai4 ni3	我爱你	I love you
88	ba1 ba1	八八	bai2 bai2	拜拜	bye bye
3166	san1 yao1 liu4 liu4	三幺六六	_	_	sayonara
9494	jiu3 si4 jiu3 si4	九四九四	jiu4 shi4 jui4 shi4	就是就是	that's it, that's it

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

Onomatopoeia

Onomatopoeia is used extensively in the Chinese data, both character onomatopoeia, and numeric onomatopoeia.

- Chinese reduplicated characters are used onomatopoeically to represent sounds:
  haha 哈哈 (wittgenstein)

- Numeric constructions are also employed as quasi-homophones to indicate sounds, e.g:
  5555 wu¹ wu¹ wu¹ wu¹ 呜呜呜呜 (sound of crying)(C2#9)

Variant stylistic spelling and synonymy

Variant spelling creates a challenge in comprehending a chat.

Emphatic spellings were indicated by elongation, e.g:

soooooo; waaaay (Channah; Leah)

A further complication in variant spellings is the interchangeability of lexical items indicating synonymous usage. For instance, laughter was indicated in numerous ways, including acronyms, onomatopoeia, and symbolic constructions. Some of these are cultural:

lol
lmao
hehe 呵呵
haha 哈哈
hehhehhehheh
heehee
~~~~~ (jingjing; ivy; happiness)

Common emoticons have synonyms, too, e.g.,

Smiley face:
☺
:)
:-)
:D
:>

Shock face:
=-O
:0
:-O

Tongue hanging out:
:p
:-p
:op

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press
~~~~~

Spelling was often very unstable even within the same chat, as if the conversation was an evolving pidgin. For instance, Leah and Channah, two university students in discussing a religion project, varied their spelling of god and G-d throughout the chat.

Understanding the meanings of acronyms and smileys thus requires understanding a range of typical constructions and seeing the interchangeability within the range. Thus chat interpretability depended on a high degree of literally understanding each conversational partner's digital language. This is highly challenging to second language learners.

Syntactic conventions

One of the questions that emerged from studying the data was about new grammatical shapes. What functioned as a word or a sentence in digital chat? We found:

Sentence/utterance shape

Sentential acronyms

In terms of new sentence shapes, we found many sentential acronyms; gtg/ g2g (got to go); ttyl (talk to you later), bringing a phrase, even of single utterance length, down to a word shape. Again there was a formality spectrum with more formal chats using more text-like sentence equivalents and less formal encounters literally transcribing speech.

Holophrastic expressions

In Chinese, we found holophrastic expressions, e.g. faint, also indicated as a reduced spelling: ft, functioning as utterance length intertextual references e.g. to a recognized Hong Kong movie.

Alphanumeric phrases

me 2	(Gary)
4 sure	(Channah)
g2g	(Channah)

Extended phrases

In the Chinese data, we found an extended use of words or phrases to indicate a larger thought:

xi xi shui ba: "wash and sleep" means 'I'm tired; I want to take a rest or go to sleep.' It does not necessarily imply taking a shower.

Yumen: "Melancholy, morose", a word, no longer current in everyday use, is frequently used to express: 'I don't feel good. This sucks.'

Sentence marking/ mechanics

Punctuation

The spectrum of conventional grammatical punctuation runs from lexical requirements, e.g. I'm, don't, 'cello; to the sentence, e.g. She asked, " Where is it?" and he answered, "How do I know!" in response.

We found little punctuation being used in general.

- irregularly with wordshapes

  youre; ive (Gary)

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

im (Elizabeth)

- seldom with "sentences"
  can u answer my q cuz I gotta tell my mom (Clarence)
  sorry for my poor knowledge (smile)

- for emphasis
  !!!...im comin too!!! (Elizabeth)
  shit!!! I just remembered that I have the stupid word assignment due on Tuesday!:-((Leah)
  welcome to home!^@^ (smile)

Capitalization

Innovative capitalization is a hallmark of digitization, e.g. WorldCALL. In the early days of computerization, DOS was not case-sensitive, so people began to use capitalization optionally for style. Upper case simply wasn't required for email. Clever advertising people have borrowed this stylistic nuance to market youth-oriented products, so now we're seeing the result in print (Lotherington, 2001a):

hot docs (Canadian international documentary film festival)
eye (Toronto weekly entertainment tabloid)
alt studio (L'Oréal hair product)

In our data, we found:

- diminished lexical capitalization:
  i; i'm; i'll (Gary; smile; Channah)

- diminished grammatical capitalization:
  are you having a happy chanuka? (Leah)
  its ppl from canada I think selling stuff. (Channah)
  i mean like beliefs (Gary)

- paralinguistic/ prosodic
  FAINT (jingjing)

Discourse conventions

Changing discourse conventions are most evident in digital chat, apart from innovations in orthographic conventions. Where the media offer new avenues for creativity, such as the coining of iconic vocabulary, alongside technological limitations, such as speed of typing constraining conversational input, interlocutors will create new ways of conversing.

Politeness conventions have evolved to label behaviours such as spoofing (injecting spurious information into a chat), trolling (sending provocatively irritating messages into a group), lurking (reading the contents of a chat but not responding), spamming (sending unsolicited mail, such as junkmail), and flaming (sending a pointedly aggressive message) (Crystal, 2001). Politeness norms include typographical conventions, e.g.

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

interpreting ALLCAPS as shouting.

Turn-taking is shaped by media possibility as well. In synchronous environments, contributions do not appear on screen until the communicator presses <send>; accordingly in multiple participant chat rooms, a number of potentially confusing conversational threads may result, depending on the timing and choice of communicators' responses, greatly affecting topic maintenance. This can be especially trying for those who are not native users of the language.

There are major differences in discourse owing to synchronicity. With asynchronous communication, such as email, interlocutors can consider and edit their responses or even choose not to respond, bringing their conversations closer to writing, whereas in synchronous environments, where they must respond instantly, the conversations move towards the speech end (Lotherington, 2001b). Figure 3 provides a spectrum.

Code mixing/switching

In the English chats we found anglicized Hebrew used. The questionnaire data revealed that bilinguals and multilinguals used other languages on line, either as medium or mixed, usually for solidarity purposes with a comprehending interlocutor. The languages used by our respondents online included: Mandarin, Cantonese, Yangchou dialect; Japanese; Hebrew; French, German, Italian; Spanish; Russian; Slovak; Tamil; Hindi/Urdu; Swahili; Jamaican Patois and Latin.

The Chinese data indicated a great deal of code-mixing and code switching, as well as transfer from other languages, predominantly English but also Japanese. The Chinese responding in the C1 questionnaires were all bilingual in English and at least one Chinese language. However, the C2 surveys also had a surprising amount of transfer and code-mixing, especially from English. Orthographically, we found inclusions of Japanese characters, Roman letters, numerals and keyboard symbols.

There was very interesting code-mixing at the level of the word, transcribing the Chinese predilection to reduplicate verbs (Binyong and Felley, 1990) in English words:

think think
say say

and most interestingly;

show show

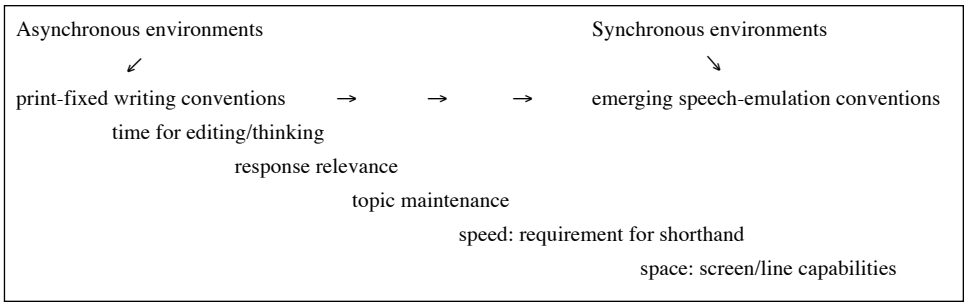


Fig. 3. Digital environments and discourse constraints.

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

where the reduplication in English is joined by the Chinese link yi (meaning one). We found:

- importation of English and Japanese words into Chinese written script:
  hoho (ivy)
  first love (smile)
  OK (wittgenstein)

- importation of English and Japanese words into compounds (names):
  Jiangsmile
  洪兴 の 甩甩*兔兔M

- incorporation of English chat shorthand
  CU (wittgenstein)
  u (smile)

- incorporation of Chinese shorthand, based on Romanized Chinese writing (Hanyu Pinyin)

supersgg			super handsome elder brother
s	shuai^{4}	帅	handsome
gg	ge^{1} ge	哥哥	elder brother

- homophonic characters representing words in other languages:

tomorrow	土毛肉	tu^{3} mao^{3} rou^{4}
me too	米突	mi^{3} tu^{1}

- morphemic mixing:
  待业 ing to emphasize present (tense) status of verb (happiness)

- code switching: importing whole English sentences into Chinese conversation
  How do you feel that? (wittgenstein)

- use of non-word symbols in names
  ☆
  ~~~~~

- Hanyu Pinyin used for onomatopoeia
  hoho

Writing a spoken vernacular

In Chinese, we found character representation of elements of spoken vernacular dialects with no standard written form in both Yangzhou Hua, and vernacular Cantonese. Two examples from Yangzhou Hua:

那又怎么样撒？	na^{4} you^{4} zen^{3} me yang^{4} sa?	So what?
今天运气不好塞	jin^{1} tian^{1} yun^{4} qi^{1} bu^{4} hao^{3} sai	Bad luck today.

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press
  ~~~~~

Parallelism in style across interlocutors
In one pair of regular chat partners there is a very interesting lack of parallelism in chat mechanics; honeygarlic is consistently more formal than sk8celine:

sk8celine	g2g work....ttysoon....
honeygarlic	ya. bye

Sociocultural conventions
Our chats included participants within a range of relationships from established real time friendships to encounters between people who only met online.

Identity markers
We were astonished with the names we encountered in the chat data, which, in some cases, constituted whole narratives more interesting than the chat itself. With this in mind, we asked about naming in our questionnaire and found a brave new world of identity formation.

It is important to realize that participants often have more than one online name, and that these names may be transitory. Indeed this revolutionary way of understanding naming, which in real time normally indicates a permanent marker of identity[8], complicated our permissions procedures as we were not always sure whose transcripts we were reading from chat participants who had agreed to participate. One thing was certain: anonymity was protected.

Naming conventions
We found much longer, narrative names in chat data where the technological parameters allow for longer names. In email, there are length and syntactic limitations. Email names were normally devoid of capitalization, spaces or punctuation, owing at least historically, to the limitations of the technology, but we also found creative capitalization:

- real or possibly real names:
  ivy
  kevinkong
  jingjing 静静

- use of another person's name:
  BinLaden
  Rothenberg (C2)
  wittgenstein (C2)

- nicknames:
  handsome 小帅
  tumptymoose

8. Names can be legally changed. Also there are cultural traditions for naming and name changing, such as the tradition of a woman taking her husband's surname upon marriage (now highly optional).

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

romaniac
boo

- words or fake words:
  feelings
  smile
  bdellovibrio (bacteria)
  blanket
  raining

- real or fake compounds:
  honeygarlic
  MasterDebater
  veganclimergrll
  Melonhead
  Cheshirecatonacid
  paddlerON
  waxpants
  Enfamous
  ufoundme
  Youjumpijump

- acronyms
  CBWY

- mixed character names that can be said:
  sk8er
  Roxstar69

- mixed character names that can only be read:
  ~ * ^ * HeAtH~ * ^ *
  &**^)(^))&%(
  洪兴の甩甩*兔兔 M
  丶☆、老√婆-也↘疯メ狂丶☆、
  ★红狐狸ㄑ→可爱靓靓ノ★

- real or fake noun phrases:
  genital herpes
  The Koz
  Who am I
  Kooth Decay

- numeric phrases
  9527

- narratives:
  Don't mess with me 2day or ill ghetto u up

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

Go Bush go!
My ugliness stirred the Chinese Communist Party
YAY!!!!!!!!!!!! I'm Learning Kanjis!!!!!!!!!!!!!
Illusion never changed into something real *Gouldie* locks
If cheerleading were any easier it would be called football.
R friends, foes in disguise?

- poetry; song lyrics or jingles:
  It's hard love, but it's love all the same. Not the stuff of fantasy but more than just a game.
  Care...you wear nothing but you wear it so well.

My teenage online users inform me that though online names change, often some element of the name will be constant, thus recognizable, such as the design frame, constructed of symbols, within which the name is put, or the signature format, or the interweaving of upper and lower case letters in the name. In these cases, the identity is not in the name but in the patterning of the graphic symbols.

Code and orthographic choice
Another very interesting identity marker was code choice: the language used to chat in. We created two sets of Chinese questionnaires: one translated into English; the other in Mandarin, in order to cater to two different demographics: Chinese using chatrooms functioning in English, and those using OICQ, which functions mainly through the medium of Chinese.

8 Conclusion and implications: Netspeak meets the second language learner

In our small scale study, we located tremendous creativity in online chat language, in terms of orthographic, syntactic, discourse and sociocultural conventions. Some innovations were exclusively Chinese, viz., the widespread use of numeric quasi-homophones and head on emoticons. The creative numeric usage owes something to the fact that Chinese logographic conventions are more cumbersome to write on a keyboard, and the limitations set inspire creative usage. The head on emoticons have no obvious roots in technological limitation, but recall obliquely directionality differences in traditional written script, and as such, may indicate "Asian" identity.

What is obvious is that chat shorthand is creative, yet logical; it is sufficiently systematic to allow for two (or more) way communication, but it is highly flexible and welcoming of innovation. Given the effects of digitization on both Chinese and English, it is fair to hypothesize similar revolutionary changes are happening in all languages used in virtual space.

The chat shorthand we studied in this project is reminiscent in some ways of pidginization and creolization processes in the creation of mixed forms within a global multilingual community having the desire to communicate across cultural and/or linguistic borders. New forms are unstable; invention and synonymy abound. Substrate and

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

superstrate linguistic contributors are discernible in code mixing.

An insistent question behind this demonstration of innovation in English and Chinese in online environments is how language teachers should react to these rapid changes in digital communication. Clearly, substantial conventional changes are happening in the two major world languages we investigated, suggesting that such transitions are happening generally in languages used for online communication. CALL materials focus on language as a spoken and written phenomenon, utilizing digital media to describe, explain, illustrate, and facilitate these communication modes. However, in online environments, such examples of language in use are inauthentic, as is the characterization of language into dual written–spoken modes (see Lotherington, in press). In authentic digital communication, a large and flexible range of conventional use is growing; Crystal (2001) labels this communication mode Netspeak – neither writing nor speech.

If indeed linguistics describes language in use and language teachers teach authentic language, then teachers of CALL have much to ask about the authenticity of their digitized course materials.

Acknowledgements

This is a revised paper based on presentation at *WorldCALL 2003*, originally entitled: "Evolving language conventions in digital environments: A look at English and Chinese".

References

Binyong, Y. and Felley, M. (1990) *Chinese romanization: Pronunciation and orthography*. Beijing: Sinolingua.

Collot, M. and Belmore, N. (1996) Electronic language: a new variety of English. In: Herring, S. C. (ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*. Amsterdam: John Benjamins, pp. 13–28.

Coulmas, F. (1996) *The Blackwell encyclopedia of writing systems*. Oxford: Blackwell.

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

Gao, L. (2001) Digital age, digital English. *English Today*, **17**(3): 17–23.

Hawisher, G. E. and Selfe, C. L. (2000) Inventing post-modern and transgressive literacy practices on the web. In: Hawisher, G. E. and Selfe, C. L. (eds.), *Global literacies and the world wide web*. London: Routledge, pp. 277–289.

Lee, C. K. M. (2002) Literacy practices in computer-mediated communication in Hong Kong. *The Reading Matrix*, **2**(2): 1–25.

Lotherington, H. (in press) Authentic language in digital environments: Research perspectives in CALL. In: Egbert, J. and Petrie, G. (eds.), *CALL research perspectives*. Mahwah, NJ: Erlbaum.

Lotherington, H. (2001a) ESL literacy and new media: reshaping literacies in the age of information. *Contact*, **27** (2) special issue: 4–11.

Lotherington, H. (2001b) Going virtual: lessons learned when a strike forced an undergraduate education course into virtual space. *The International Journal of Learning*, **8**. http://HeatherLotherington.Author-Site.com/

Merchant, G. (2001) Teenagers in cyberspace: an investigation of language use and language change in internet chatrooms. *Journal of Research in Reading*, **24** (3): 293–306.

Randall, N. (2002) Lingo online: a report on the language of the keyboard generation.

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press

http://www.arts.uwaterloo.ca/~nrandall/lingo-online.htm

Taylor, I. and Taylor, M. M. (1995) *Writing and literacy in Chinese, Korean and Japanese*. Amsterdam: John Benjamins.

Warschauer, M. (1999) *Electronic literacies: Language, culture and power in online education*. Mahwah, NJ: Erlbaum.

Werry, C. C. (1996) Linguistic and interactional features of Internet relay chat.In: Herring, S. C. (ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*. Amsterdam: John Benjamins, 47–63.

https://doi.org/10.1017/S0958344004000527 Published online by Cambridge University Press