

SITUATING SMS (SHORT MESSAGE SERVICE) DISCOURSE

Introduction

The advent of digital media, beginning with computers, has opened up new channels of communication, and with these, new ways of using language for communicative purposes. The Internet in particular, has shaped several new communicative genres, such as email, chat and virtual worlds (Crystal 2001). The language of the Internet, variously termed 'Netspeak' (Crystal 2001), "Internet language" (Zitzen & Stein 2004) etc. has many distinctive features that sets it apart from the conventional written and spoken varieties we are familiar with; and some of these features are now appearing in other forms of discourse as this study will demonstrate. As Crystal observes,

The electronic medium ... presents us with a channel which facilitates and constrains our ability to communicate in ways that are fundamentally different from those found in other semiotic situations. Many of the expectations and practices which we associate with spoken and written language ... no longer obtain." (2001:5)

In a similar vein, Zitzen and Stein (2004:983) go so far as to propose that in addition to the two older media of written and spoken language, we now have a third medium,¹ language in the Internet, whose status relative to spoken and written language is not entirely clear.

The focus of the present study is another type of digital discourse that has come into being with developments in mobile telephone technology—the SMS (for Short Message Service), or 'text' message as it is popularly known. Although SMS has been in existence since the 1990's (Crystal 2001), it has not been subjected to rigorous analysis in terms of its functions, features or conventions, although digital Internet genres have been researched and written about fairly extensively (See, for instance, Porter, 1996; Davis & Brewer 1997; Baron 2000; Crystal 2002; Zitzen & Stein 2004 etc.). There is possibly a perception that SMS discourse, sharing as it does many characteristics with Internet language or Netspeak, does not have sufficient distinctive features to be considered a discourse genre in its own right. However, in Sri Lanka at least, the use of these short text messages to satisfy a variety of communicative purposes points to a fairly wide popularity and

¹ For Crystal, the Internet is a fourth language medium, after writing, speech and signing.

acceptance; in spite of this, the fluidity of SMS language and conventions, and its impact on language use remains unexplored.

This paper will attempt to answer the following questions. Where should SMS discourse be situated as a means of communication? What genres of written, spoken or Internet language does it most closely approximate? What conventions or constraints is it bound by? What impact might it have on language variation and change?

Corpus and data

The data for this study comes from a corpus of SMSes collected over a period of 10 months. These messages were collected by using the researcher's professional and social networks.² Permission was obtained from the senders to include their messages in a corpus for research purposes. No names of senders or receivers were recorded in order to preserve the condition of anonymity. However, the sender's age, gender and highest level of education achieved were recorded, as sociolinguistic variables that would be taken into consideration in the linguistic analysis. In terms of length, the messages in the corpus vary from minimal response texts with an average of about three characters, such as "Ok", "Yes" and "Hmph"³ to much longer texts of up to 200 characters. The term 'character' can represent a letter (*a*), a symbol (+) or a mark of punctuation (;) in an SMS text. Length in SMS discourse cannot be quantified in terms of word units, as definitions of what constitutes a 'word'—applicable in most other forms of written discourse—cannot be applied here, as illustrated in the example below.

[1] look fwd 2 C ing u 2moro⁴

The senders of the SMSes in the corpus are males and females in the age range of 23 to 55 years. Highest educational qualifications represented ranged from a pass at the General Certificate of Education Advanced Level (GCE A/L) Examination to a PhD. In terms of language proficiency, most are native speakers of Sri Lankan English, while a few can be classified as speakers of Sri Lankan English as a second language—i.e., they are more proficient in the use of Sinhala than in English. Only

² This included a version of the 'friend of a friend' technique of data collection as outlined by Milroy (1980), where friends were asked to obtain SMSes from their friends for inclusion in the corpus, under the conditions of anonymity established for this study.

³ This is an example of what Crystal (2001) terms comic-book style interjections found in Netspeak. Other examples are 'heh heh', 'ha ha' etc.

⁴ 18 characters

SMSes where the base language is English were included in the corpus of this study. No messages where reference was made to topics of a personal nature—Eg: death, illness, injury, relationship issues, etc. or where the sender could be identified from the content of the message were included in the corpus. If references were made in an SMS where third party identification was possible, names were deleted when the message was recorded in the corpus. Since this study proposed a purely linguistic analysis, picture messages, with or without accompanying text, were not included in the corpus.

Situating SMS Discourse

Fairclough (2003) observes that discourse can be differentiated with respect to communicative features and the technologies employed in terms of two distinctions: one-way vs two-way communication and mediated vs non-mediated communication. Print, photography, broadcasting and the Internet are examples of mediated communication, which Fairclough defines as that which "... makes use of copying technologies which disseminate communication and preclude real interaction between 'sender' and 'receiver'" (2003:219). Face to face interaction is an example of two-way non-mediated communication that is synchronous, as it takes place between two or more participants in real time. SMS discourse on the other hand is an example of one-way mediated communication which is asynchronous as it takes place in what Crystal (2001:9) refers to as "postponed time"—i.e., there is a time lag between the composing/sending and receiving/reading of the message.

SMS and speech

How similar is SMS discourse to speech? One of the characteristics of any spoken discourse is prosodic features such as pitch, intonation and tone. Clearly, this cannot be achieved in a medium that uses graphic characters and a display. Senders therefore resort to a convention established in chat discourse—the use of smileys to indicate emotion, and capitalization to denote loudness of tone or emphasis. Both methods have obvious limitations; the use of capital letters, in particular, can be ambiguous, as the lack of accompanying or supportive acoustic cues could cause confusion as to whether the sender wishes to convey an aggressive tone, or simply surprise, as in the example below.

[2] REALLY?

However, the corpus also reveals that SMS discourse has many features of colloquial Sri Lankan English which would not be found in formal writing, such as code-switching, tag questions, reduplication, interjections, and idiomatic or formulaic expressions. Instances of these are illustrated in the examples below.

- [3] Good luck for tomorrow. *Kökata kiyala*⁵ it's good to take a good pair of ear plugs, no?:-) (code-switching/tag question)
- [4] Happy couple must b out shopping *no?* (tag question)
- [5] do you think _____ (name deleted) will let us record his talk? But *pl pl* take notes if you do go. Am really interested. (reduplication)
- [6] *aiyo* I can't get ovr it:-(. (interjection)
- [7] Ok. *Kokata väraya nam thiththayatath värayak né.*⁶ You wait and see how I act out the heroine of the MILLS & BOON Story.:-) (idiom/code-switching)

Code-switching between English, Sinhala and Tamil can be found in the speech of most Sri Lankan bilinguals and multilinguals, but except in the medium of print media and advertising is less common in written genres. The presence of code-switching in SMS discourse therefore poses a question: Do users feel that in spite of the use of a screen and graphic characters, communicating via SMS is more akin to speech than to writing? Certainly, the use of the tag question "no?"⁷ and the common Sri Lankan interjection/exclamation "aiyo" seem to point to such a perception.

SMS and writing

However, SMS discourse has many affinities with writing as well, although here again it is equally difficult to definitively situate it in relation to conventional genres of writing, as it has shared features as well as distinctive features of its own. For instance, it relies on alphabetic characters to convey meaning, as does any writing in

⁵ Translation: "in order to be prepared for any eventuality" or "in any case."

⁶"It is not only the strong [person] who has the advantage all the time. the weak [person] may also be in an advantageous position sometimes."

⁷ "No?" is commonly believed to be a transfer from the Sinhala negative marker "*né?*" or "*néda?*"

English, but in addition, can and does draw on a combination of alphanumeric characters, ideographs, (Eg: smileys) and a selection of symbols, the use of which would not be acceptable in most written genres.

[8] Tx. Good 2 hear from u. Having 20 guests 4 dinner. I'm cooking. So all outings out!:-(

[9] Wats dis u have become a gypsy in é metropolis. Do try a b there.

These examples also illustrate another feature of SMS discourse—the use of abbreviations or truncations of certain words (*u* and *b*) which is now sufficiently established to be considered conventional by users. This, along with a manipulation of spelling and even deliberate misspelling of words (*wats* and *dis*) is often the result of space constraints; on average, a single SMS can only have about 150-160 characters. Any text lengthier than this will be sent as two messages by the service provider—making for a very perceivable instance of mediation. Messages with more than 300 characters will be sent as three messages, up to a maximum of about 450 characters.

As is the case with writing, SMS discourse employs punctuation, but this is selective. Question marks, periods, commas and even double quotation marks and dashes do appear in SMSes, but in general, there is a laxness in punctuating the end of the last sentence, and an often complete disregard of the conventions of capitalization, as illustrated by the examples below.

[10] Yep. I rmember, shall do so pronto

[11] 2nd wed 10am to 1pm Am center

However, some messages, especially those from service providers, are extremely formal in tone, and adhere to established conventions of grammar, syntax, capitalization, punctuation etc. as illustrated below.

[12] Dear Customer; We thank you for your payment of Rs 1400.00, and giving us the opportunity to serve you.

Example 12 conforms in most ways to a formal business communication in writing between a service provider and a client, in terms of tone and the inclusion of a salutation. However, it does not have an appropriate closing, which would be a mandatory requirement if this communication was sent in the form of a letter.

Further complicating matters, we have SMS's such as example 13 below.

- [13] Nissan March for sale. WP JC.Black. Rs 11,75.pls contact mobile or _____
(*number deleted*) thx. _____ (*sender's name deleted*)

This type of business communication conforms to the genre of a classified advertisement in a newspaper in its brevity and telegraphic tone, resulting from a lack of embellishments and prepositions, articles, etc. It also contains abbreviations, the decoding or interpretation of which requires a fairly specialized knowledge of the item advertised.

SMS discourse therefore allows for what Fairclough (2003) terms genre mixing. All of the features discussed so far, both linguistic and discursal, contribute to the difficulty of situating SMS discourse using conventional parameters of speech and writing. Indeed, the closest approximation to SMS discourse appears to be Internet language, with its mixing of writing and speech, which Baron aptly describes as "... an emerging language centaur—part speech, part writing" (2000:248).

Any comparison that involves Internet language must, however, be cognizant of the existence of distinctions between the discourse characteristics of different Internet genres such as email, chat, virtual worlds and the world wide web in general, as illustrated by Crystal (2001). In terms of these distinctions, the language of SMS as found in the examples listed so far in this study most closely approximates the language of chat as described by Zitzen and Stein (2004) in their extensive study of the genre. However, some differences still remain. Chat is synchronous, while SMS is asynchronous. Asynchronicity renders turn-taking—a vital part of conversational structure even in a chatroom—impossible. Adjacency pairs, while existent in SMS discourse (as illustrated below), are adjacent only in meaning, not in time; if one SMS sender is slower than the other, even an intended adjacency pair may not be realized as intended.

- [14] A: Yes that's fine-see you then. Goodnight.
B: Ok, g'night

Linguistic variation

A further complexity encountered in trying to situate SMS discourse within conventional parameters is the unusually high degree of linguistic variation it allows. This may be because SMS discourse is still a relatively new genre, like the language of the Internet which Crystal claims is still in a state of transition and therefore "... has no rules in the sense of universally agreed modes of behaviour

established by generations of usage” (2001:15). This lack of ‘rules’ has resulted in variations in spelling, abbreviations and truncations. Some of these abbreviations and truncations are transfers from email and chat discourse, such as *tk*s (thanks), *u* (you), *nite* (night) etc. and are used by SMS senders for both ease and quickness of typing as well as in awareness of the limits on the number of characters a single message can contain. As Crystal observes, however, the small screen size and limited character space has “motivated the evolution of an even more abbreviated language than that of chat groups and virtual worlds” (2001:228) in the language of SMS.

The table below gives a list of lexical items and their variants as found in the corpus of this study.⁸

Lexical item (word)	Variant
about	abt, bout
and	&. n, +
because	bec. cos
Colombo	CMB, col. cbo
could	cd, cld, cud
from	fr, frm
have	hav, hv
morning	morn, morng, mrng, am
okay	k, ok
please	pl, pls
should	shd, shld, shud
thank (you)	tk, thks, tx, thanx
the	é, da, d
tomorrow	2mrw, 2moro, tom, tomo, 2morrow, tmrw
will	wil, wl, l
with	w/, w
would	wd, wld, wud
your/you're	yr, ur, u're

Some of these variants are clearly based on phonetics or the application of phonological principles, as in *shud*, *wud*, *ur*, *2moro*, etc. Other abbreviations are the result of vowel deletion (*wld*, *hv*, *thks*, etc). A third type of variant is a letter of the

⁸ It should be noted that this is by no means a comprehensive list of all variants found in SMS discourse, but the ones found in the corpus.

alphabet or a number which is a homophone of a particular lexical item, such as *v* (*we*), *r* (*are*), *2* (*to* or *too*), *b4* (*before*) etc. However, a large part of the variation found in SMSes is individualistic—i.e., the result of creativity and manipulation of spelling conventions on the part of individual message composers. Indeed, some messages show a combination of several types of variation—vowel deletion, combinations of alphanumeric characters, phonetic spelling and conventionalised abbreviations, as evidenced in the example below.

[15] Thx 4email. Gd u recovrg+bk 2 routine. Sis come 2SL aug.may join briefly.
how r u?

Even in a chatroom, such a high degree of variation and departure from the conventions of spelling would be unusual in a single turn. It is perhaps the higher degree of intimacy involved in SMS discourse—i.e., the sender and receiver are acquainted with each other (except of course in the case of a service provider and customer) that allows for a greater amount of creative variation. Between friends or acquaintances, idiolects are more easily decoded and understood; in a chatroom, however, because chatters do not necessarily have prior acquaintance with each other, a greater degree of adherence to established or conventionalised variation is demanded. In this sense, instant messenger services on the Internet which function on the basis of the sender and receiver being acquainted with each other are perhaps the closest relation to communication via SMS.

What effect, if at all, might such variation have on language use in general? To put it differently, might features of a newly emergent discourse (as SMS is in Sri Lanka) be transferred to other communicative genres in terms of grammar, spelling or vocabulary? Will, for instance, the following be permitted in any other medium of communication but an SMS?

[16] don't worry ur presntatn wil go ok an no1 wd u'stand anyway. i 4 l wont!⁹ ;)

The answer to this question lies in how prescriptivist a stance we wish to take on language use. Crystal (2001) claims that 'errors' of spelling, grammar and punctuation are tolerated in email; this appears to be the case with SMS as well. Because of SMS users' awareness and exploitation of such acceptance, it is almost impossible to ascertain a sender's English language proficiency, social or

⁹ Don't worry your presentation will go okay and no one would understand anyway. I for one won't!

professional status, age, etc., all of which are factors that usually correlate with linguistic variation. Take for instance, the following messages:

- [17] Cnt contact sir on mob.ringin no answer.Cald hme n a lady told he has gone out frm CMB.just try on mob again.if nt going v l in4m sir ok.reply if u gt 2 knw ntin
- [18] hope u got my sms b4 i left SL.how are u now? am v sorry couldn't come 2c u.last days were v. stressful.take care. _____ (*name deleted*) xx.

Both these messages flout conventions of spelling, punctuation and grammar. The sender of Example 17 is in her early 20's, uses Sri Lankan English as her second language, and has successfully completed the GCE A/L Examination. The sender of Example 18 is in her early 40's, is a native speaker of Sri Lankan English, and has a PhD. In terms of social factors, it is very unlikely that these two subjects would be grouped together under any criteria of categorisation as they show no commonalities – except of course, in language use. As is the case with email (Crystal 2001) therefore, there appears to be a common understanding among all users that when composing an SMS, one does not need to be as prescriptivist as demanded by most other forms of written communication.

From a sociolinguistic point of view, this is evidence of two remarkable effects. First, factors such as age, educational level, language proficiency, profession, etc. which are known to have an effect on language use and variation, appear to have no visible influence on the language of SMS in Sri Lanka. This challenges a fundamental notion underlying linguistic variation—that variants are associated with social factors, and that variation is therefore predictable (Chambers 1995). Second, SMS discourse in Sri Lanka appears to indicate a “levelling” process, a term used to describe varieties of a dialect or language moving towards or approximating a standard. In this particular case, however, the leveling is happening in reverse, where highly proficient users of Sri Lankan English with educational qualifications as high as a PhD express themselves in language that is indistinguishable from that of a second language learner of English. (See Example 18). The reason for this phenomenon demands more investigation. Why have non-standard usages become more common in an increasingly prescriptivist environment of English language use, with even those highly proficient in English flouting the conventions of ‘standard’ grammar, syntax and spelling? What is it about SMS discourse that allows such non-standard usage without the danger of stigmatization?

Is it because this is still a relatively new type of discourse, and therefore does not have an established tradition or convention to draw upon?

SMS discourse therefore appears to defy explanation or categorization not only in relation to other types of discourse whose features have been already analysed and described, but also in relation to patterns of sociolinguistic behaviour we are accustomed to. The wide range of communicative purposes SMS is used for in Sri Lanka apart from that of simple conversational interaction—i.e., exchanging news and information, sales and marketing, making and confirming appointments, even voting—indicates that it has become a preferred method of communication. It remains to be seen whether SMS texts will continue to show a fluidity of expression and a strong sense of individuality in language use as they do now or if any current non-standard or variant usages will become conventionalized over time.

Acknowledgements

I wish to thank the anonymous reviewer whose insightful comments helped to make this paper more comprehensive. Thanks are also due to all providers of SMS messages for inclusion in the corpus.

Bibliography

- Baron, Naomi S. (2000). "Alphabet to Email." London: Routledge.
- Chambers, J.K. (1995). Sociolinguistic Theory. Blackwell.
- Crystal, David. (2001). Language and the Internet. Cambridge: Cambridge University Press.
- Davis, Boyd H. and Jeutonne P. Brewer. (1997). Electronic Discourse: Linguistic Individuals in Virtual Space. Albany, New York: State University of New York Press.
- Fairclough, Norman. (2003). Analysing Discourse. London: Routledge.
- Milroy, Lesley. (1980). Language and Social Networks. Oxford: Basil Blackwell.
- Porter, David. (1996). (Ed.). Internet Culture. New York and London: Routledge.
- Zitzen, Michaela and Dieter Stein. (2004). "Chat and Conversation: A Case of Transmedial Stability? *Linguistics* 42-5, 983-1021.

DUSHYANTHI MENDIS