BAHASA SEMBANG IN WEB FORUMS
Knowledge Management for Piles of Atopian Discourse

Mohd Zaidi Abd Rozan, Yoshiki Mikami

Nagaoka University of Technology, Nagaoka, Niigata, Japan. 940-2188

Keywords: *Bahasa Sembang,* Web Forums, Knowledge Management, *Bahasa Melayu* (BM), Written Speech.

Abstract: The popularity of web forums has led to the creation of evolutionised *bahasa* Melayu (BM) known as *sembang* language. Greatly accepted as common and sensible vocabularies for its communities, *sembang* language has emerged as an avenue that supports knowledge-sharing-procreation practice. In this paper, we are trying to explore the fact that *bahasa sembang* could serve as a melting pot for knowledge breeding since it carries special characteristics that emulate spoken expression via explicit spelling. Furthermore, the inclusion and blending of tacit and explicit knowledge are contributing to the *Socialization* and *Externalization* process in knowledge management as proposed by Nonaka's knowledge creation model.

1 INTRODUCTION

The inception of synchronous and asynchronous collaborative system has given birth to a new lingo called chat language or locally known as *bahasa sembang*. Welcomed by netizens, *bahasa sembang* is the language mostly used nowadays in web forums, on-line chat, etc. While *sembang* language is comfortably exercised for the transfer of ideas and working out problems, this gives effect to the growth of discussion forums as its channel. Since the control of such neologisms is impossible, looking at the bright side of its growth is somehow essential.

To the best of our knowledge, up to now limited research on *sembang* language was published. Existing articles tend to superficially deal on the existence and its negative aspect but not to any special functionalities it carries. Those negative aspect concerns with the numerous changes to its morphology and syntax, (Awang, 2004) stressed that Bahasa Melayu (BM) sovereignty will be facing an awfully shaky journey as it is being mutated and deteriorated in terms of its status and quality.

The following sections will relate to problem background, objective and definition of *sembang* language. This follows by our exploration based on the Nonaka's Knowledge Creation model elaborated particularly on Socialization and Externalization process for *sembang* language in web forums.

2 PROBLEM BACKGROUND

Even though *bahasa sembang* is so far infantile, yet is seen as posing threat to its primary counterpart, the standard and official BM. Linguists and concerned individuals have shown their worries ever since the beginning of the utilisation of this language. But as language tends to develop, it is not an easy task to get hold to (Rusli, 2005) as this kind of scenario is being part and parcel of the Internet technology consequences.

On the other hand, such dynamic characteristics of this transformation are viewed as a fresh start inculcating new sustainable force for BM in this cyber age. While the dominancy of other foreign languages (especially English) is regularly visible on-line, still this modified version of BM could easily be found and even vigorously used. This has shown that this new breed of BM is being nurtured rapidly without any sign of diminishing.

In this paper, our main objective will be exploring that the usage of *bahasa sembang* could possibly serve as a melting pot of knowledge breeding in web forums. We are certain with its high usage in discourse media, some special characteristic do exist and aiding to the Socialization and Externalization portion of Nonaka's (1995) knowledge creation model.

3 *SEMBANG* LANGUAGE BASICS

Originally the word "chat" is identified as "*sembang*" from Malay language and the language is well known as *sembang* language. *Sembang* language is recognized in spoken and not in written, this is in fact that *sembang* words are colloquially inclined because the nature of on-line communication at present, it is essential and only possible that it has to be written down. This has been the pushing factor for the modifications in spellings particularly representing spoken words in a written manner.

By merely looking at 10 randomly selected pages from one specific web forum (http://asamboi.org/) on the Malaysian cyberspace, there are not less than 300 unique *sembang* words that co-occurred 515 times. Based on this numbers, a considerable amount of knowledge are being created, recorded and updated utilising *sembang* language as the discourse takes place.

Table 1: BM Evolution to the Sembang Language Development.

BM words (English translation *)	*Sembang* words	Evolution Process	Transformation Factor
Bagi (give)	Bg		
Berhenti (stop)	Benti	No 1	Words economisation
Duduk (sit)	Duk		
Begini (like this)	Gini		
Ambil (take)	Amik		
Hantar (send)	Antar	No 2	Similar to spoken style
Baca (read)	Bace		
Serius (serious)	Seryuz		
Lepas (ago)	Lepaz		Similar to spoken style and melodically sound
Lupa (forget)	Luper	No 3	
Cerita (tell story)	Citer		

Only one literal meaning without listing all available synonyms

(Gani, 2000) has identified few developments of BM evolutions on the Web based on the lexical characteristics of words. In Table 1, as an example the transformations of Bahasa Melayu (BM) to sembang words are shown. Evolution process no 1 occurred due to the need of words shortening or economisation. For evolution process no 2, the transformation were caused by the need to be similar to its spoken component. Another transformation is quite equivalent to process no 2 but with the addition of melodically driven utterances. The reason behind this is the need to drive close contact or intimacy between speakers by having friendlier positive (or even negative) expressions.

In Table 2, IPA is used to represent the exact and unambiguous description for the pronunciation. Here it is shown that three types of articulation style involved; for written, spoken and *sembang*. The original dictionary words are maintained or transformed when being utilised in each of the style. For the written style, words are spelled exactly as it is. However for the spoken and *sembang*, normally words are altered in spelling, for example, the word "kenapa" are transformed to "kenape" and "kenaper" in spoken and sembang style respectively. These transformations occurred to suit the normal colloquial presentation daily practiced in conversation.

Table 2: Word Articulation Process based on International Phonetic Alphabet (IPA).

Dictionary Words (English translation)	Articulation Style		
	Written	Spoken	*Sembang*
1. Kenapa (Why)			
Normal Spelling	Kenapa	Kenape	Kenaper
IPA Spelling	kɜːnʌpʌ	kɜːnʌpɜː	kɜːnʌpɜːʳ
Similarity with sound perceived	Less similar	More similar	More similar
2. Serius (Serious)			
Normal Spelling	Serius	Serius	Seryuz
IPA Spelling	sirius	sirɜːs	sirɜːz
Similarity with sound perceived	Less similar	More similar	More similar
3. Macam ini (Like this)			
Normal Spelling	Macam ini	Camni	Camnie
IPA Spelling	mʌtʃʌm ini	tʃʌmni	tʃʌmniː
Similarity with sound perceived	Less similar	More similar	More similar

Similar to words no 2 and no 3, these newly formed words represents more expressive manner of "written speech". The words in *sembang* and spoken shared similar resemblance for the sound perceived if looked at the IPA spelling. Contradictory to its typical written style, there is less similarity compared to both *sembang* and spoken. This is because standard written style does not emulate the exact "sound" of the written words if pronounced.

4 KNOWLEDGE CREATION MODEL

In this paper, a knowledge management model (Table 3) as introduced by (Nonaka and Takeuchi, 1995) will be employed. This provides the structure for discussions by referencing to the notions of tacit and explicit knowledge introduced by (Polanyi, 1966).

Table 3: Nonaka and Takeuchi's KM Model or SECI Model.

To From	Tacit	Explicit
Tacit	Socialization (S)	Externalization (E)
Explicit	Internalization (I)	Combination (C)

Managing knowledge includes its subjective and documented type, known as tacit and explicit. Tacit knowledge is what the knower knows, which is derived from experience and embodies beliefs and values. This is an actionable knowledge and therefore the most valuable (Marwick, 2001) but frequently unstructured, intangible and difficult to codify because cumulatively stored in ones mind (Fareed H. et. al., 2004).

Explicit knowledge is represented by tangible entity, such as documents, writings or even multimedia, which is intended to be use to transfer its content to another person (Marwick, 2001). And normally codified and ready to be distributed to others without involving interpersonal interactions (Fareed H. et. al., 2004).

4.1 Socialization (Tacit to Tacit)

Typical socialization process takes place when sharing of experiences and information occurs by means of face to face meetings (or telephone) and basically due to the need of verbal communication. Anyhow, this process could also be realized without verbal interaction. According to (Bradner et al., 1999), 'chat' evaluated as much like a conversation. Since the style of words in *sembang* language is 'chat' alike, the structure of its spelling is very much similar to the phonetics of spoken words. Here, the ability to depict spoken component as conversation is the biggest advantage of *sembang* language. In this case, the spoken words that were typed by the composer are usually rooted from colloquial style writing. Furthermore, the inability to present the message in speech has motivated the utilisation of

limited input device at disposal to create and achieve spoken style as best as possible. The invention of *sembang* language has very strong link to this factor.

(Binder et al, 1997) stated that language is defined broadly to include both phonological and lexical-semantic functions and to exclude sensory, motor and general executive functions. This means that the usage of "sounds" embedded in wordings are inline with the definition. The attentiveness to a variety of signals given by a speaker such as fluency, changes in prosody have strong effect towards the control of comprehension for a message (Isaacs, 1994; Titsworth, 2001 in Piolat, 2005). *Sembang* language carries specific phonetic marks that acts as attribute in transmitting special cues and this stylistic method often signalled how important a particular information is conveyed. Again the need for "speech communication" through more informal and simpler *sembang* language carries the biggest effect for socialization process and this might ease the sharing of knowledge and expertise in this channel.

Talking about readability, a reader reads words in text and tends to emulate by following to his own rhythm or beat. By having expressions that are compounded in *sembang* words, it will be much easier to follow as such rhythms are noticeable and straightforwardly digested by the reader. This imitate as if the reader is listening to the sound of the words being pronounced and promotes to better comprehend the expressions shown.

Another point to look at is the function of language as bonding tool. This means that utilizing a substantial amount of commonly accepted language (such as *sembang*) would build good ground for linking up with web forum communities.

4.2 Externalization (Tacit to Explicit)

When tacit knowledge is transformed explicitly to written articles or even diagrams, this is called externalization. In this process, having dialogue is an effective method to articulate knowledge with others (Nonaka & Toyama, 2003). The main strength of it is the possibility of making tacit knowledge explicit by archiving. As the recording of "spoken language" takes place, this will imitate the way to disseminate knowledge via sound and speech.

And because tacit knowledge involves ideas that emerge promptly from the brain, this is a natural tacit knowledge diffusing process. A simple medium of communication such as telephone will involve listening and decoding hence understood by the

listener, but not necessarily recorded. However, by using *sembang* language the spoken style articulation could be nicely recorded.

As the slower pace of web forums compared to on-line chat, this provides opportunity mostly for critical discussions. This provides the reason that *sembang* activity in web forums is not wholly talking nonsense or about useless agenda (Mohd Zaidi & Mikami, 2006). On the contrary, synchronous discussion is often suitable for very light discussions because of the lower mental hurdles related to this type of interaction.

5 CONCLUSION

The aim of this paper is to provide an overview on how *sembang* language acts as the avenue for promoting knowledge creation. To this stage, we have shown that the usage of *sembang* language in web forums is contributing to motivate knowledge sharing and procreation. Furthermore, we have demonstrated that the tacit knowledge of individuals that are expressed in *sembang* style can be captured and specially treated for Socialization or Externalization as in Table 4.

Table 4: Utilization of Sembang Language for Socialization and Externalization Process.

Socialization (Tacit to tacit knowledge)	Act as an emulation of spoken phrase by its explicit spelling
Externalization (Tacit to explicit knowledge)	Spoken style articulation can be recorded as text for archiving purpose

Looking at the Latin expression *"verba volant, scripta manent"* means "spoken words fly away, written words remains"; here lays the fact that spoken words are likely to disappear compared to its written counterpart. However in this study, it shows that the possibility to record the *sembang* form as the verbal style would be a contributing factor for it to be kept alive and stored.

For a long time, oral mode is possible for very near distance people. But now, it is possible to bring many people who have similar interest into a room and have them talking their oral mode but in written style. The importance of analysing what is happening plus the changes in the communication media that facilitates the stimulation of evolution in a language must be taken into consideration.

The next steps remains for our future research are to implement words and phrases analysis in order to understand the value of knowledge creation

from it. The potency of utilizing *sembang* language in web forums is a strong research topic that we challenged for. This will provide insights to swerve away from being paranoid since we have less understanding of the so called "threat" we are facing. This issue is technically and also socially challenging because there are huge gaps between *sembang* language devotees and disbelievers. This paradox could be viewed as inducing towards BM deterioration or perhaps as a new prop for development of BM in the cyberspace.

REFERENCES

Awang, S., 2004. Strategi Memartabatkan Bahasa Melayu. *Kertas edaran dalam Mesyuarat Strategi Memartabatkan Bahasa Melayu,* Kementerian Pendidikan Malaysia, Putrajaya, 21 Disember 2004.

Bradner, E., Kellogg, W. A., & Erickson, T., 1999. The Adoption and Use of 'Babble': A Field Study of Chat in the Workplace, *Sixth European Conference on Computer Supported Cooperative Work*, Copenhagen (1999).

Fareed, H., et al., 2004. Managing Knowledge Effectively. *In Journal of Knowledge Management Practice*, Volume 5, No.8. ISSN 1705-9232.

Gani, M. P., 2000. *Evolusi Bahasa Melayu Internet.* Retrieved January 24, 2005, from Majlis Bahasa, Bicara Bahasa Web site: http://www.bahasawan.org/isian/bm-net.htm

Marwick, A. D., 2001. Knowledge Management Technology. *IBM Systems Journal*. Volume 40, No. 4, 2001.

Mohd Zaidi, A. R & Mikami, Y., 2006. *Knowledge Details in Web Forums: How High or Low Above the Ground*. In 19th International Federation of Information Processing World Computer Congress 2006, IFIP 2006 (Under revision).

Nonaka, I & Toyama, R., 2003. Knowledge-creating Theory Revisited. *Knowledge Management Research & Practice*, Volume 1, pp 2-10.

Nonaka, I & Takeuchi, H., 1995. *The Knowledge-Creating Company. How Japanese Companies Create the Dynamics of Innovations*, Oxford University Press. Oxford and New York.

Piolat, A., 2005. *Ecrire pour le Web : Stratégies de transformation du contenu et de mise en forme matérielle d'un texte.* Retrieved October 10, 2005, from Université de Provence-Aix-en-Provence, Professor Annie Piolat Web Site: http://www.up.univ-mrs.fr/psycle/membres/enseignants/apiolat.html

Polanyi, M., 1966. *The Tacit Dimension*, Routledge and Kegan Paul. London UK.

Rusli, A. G., (personal interview, October 5, 2005) explained the challenges that are faced by Bahasa Melayu (BM) at present and in future especially in this Internet era.