

WORD FAMILIARITY AND LEXICAL CHANGE: THE CASE OF SARAWAK MALAY DIALECT

Radina Mohamad Deli¹

Rosnah Mustafa²

Monaliza Sarbini-Zin³

Centre for Language Studies, Universiti Malaysia Sarawak

¹mradina@cls.unimas.my

²mrosnah@cls.unimas.my

³szmonaliza@cls.unimas.my

Abstract

This preliminary study looks at the familiarity rating of words in Sarawak Malay Dialect (SMD). Although familiarity ratings of language items are usually utilised in psycholinguistic research, they can be very useful for studies in the area of language change. The aim of this study is twofold: (1) to compare the perceptive familiarity rating of Sarawak Malay words; and (2) to document Sarawak Malay words that are undergoing lexical change. Fifty SMD words were used in this study consisting of those with meanings that can be considered as medium to high in frequency for everyday speech. Questionnaires were designed using a 5-point Likert-type scale to rate word familiarity and distributed to 15 participants who were native SMD speakers between the ages of 20 and 25. Across word items, more than one third were found to be rated as less familiar and unknown and thus were not actively used in daily conversations. There were also a number of words perceived as familiar to highly familiar but were not widely used in everyday speech. Evidently, it is crucial to document and preserve these SMD words as they are fast becoming passive vocabulary for the young and may eventually be lost in their lexicon.

Keywords: Word familiarity, Sarawak Malay dialect, language change, passive vocabulary

Introduction

It is a generally accepted notion that languages are alive and dynamic in the sense that they are systems which evolve over time. In line with the global changes that affect communities' sociocultural and socioeconomic conditions, it is inevitable for the lexicon, morphology and syntax of a particular language to undergo constructive and even detrimental changes. In an ever-changing world, the lexicon of a less dominant language or dialect undergoes lexical change processes which affect the meaning of a word in its usage. There are also instances in which a shift or variation in word preference occur across generations of speakers. Oftentimes, observable changes in the lexicon are the most common indication of language change.

There are also other more serious language phenomena affecting the lexicon. One of it is when words cease to be used by a linguistic community and disappear gradually from the current lexicon. The loss of words or meanings is attributed to factors such as urbanisation and language contact.

In the case of Sarawak Malay dialect (hereafter, SMD), numerous works which are descriptive in nature have focused primarily on linguistic systems and patterns of variations observable in SMD (e.g., Collins, 1987; Zainal Abidin Merjan, 1992). The present study, however, examines SMD from a different perspective given the view that a large number of words in the SMD lexicon are fast becoming passive vocabulary among the young and old SMD speakers alike. A shift from SMD to a more standardized Malay lexicon or Standard Malay (hereafter, SM) has also been identified (e.g., from the word *manok*, meaning chicken in SMD, to the word *ayam* in SM). One of the causes of such a shift is the use of SM as the primary medium of instruction in Malaysian national schools. A Malay child growing up in Malaysia (in the case of this study, Sarawak) will most likely learn SM in primary and secondary schools, and even the university. This study, however, will not explore the child's SM acquisition and learning processes. Instead, the study examined 50 SMD words that are in danger of disappearing from the SMD lexicon by using a psycholinguistic method of rating words and their meanings based on perceived familiarity.

The only study to gauge familiarity ratings of SM words was done by researchers from National University Singapore in which a total of 530 words were examined (Pareira, Liow & Mohd Saniff, 1992). Another study from the same university looked at the familiarity ratings of standard Malay affixes (Lee, Liow, & Wee, 2007). Thus far, there are no other studies found using similar methods for rating word familiarity of a regional dialect. Compared with other known languages in the world, European languages have been documented extensively not just based on corpus and frequency counts, but also in terms of familiarity. The method of rating words is important in measuring how well language users know or understand certain language items. The results of such measurement can be used to validate studies investigating the use of these items in language comprehension and production. Other more advanced longitudinal studies use computational modeling methods involving large numbers of language items to explore intra- and inter-generational processes of language change and variation (e.g., Smith, 2007).

Using familiarity rating to document words will provide substantial information about certain language items. This include: 1) the language items that are less known or relevant for a specific language community; 2) the differences in familiarity rating between generations of speakers; and 3) the percentage of loss of words overtime. This study will focus primarily on the first type of information. The objectives are: a) to compare the perceptive familiarity rating of Sarawak Malay words; and b) to document Sarawak Malay words which are undergoing lexical change.

Methodology

A questionnaire was designed to rate the perceived familiarity of words based on the following scale and description which have been translated into English from their original SM version:

- 1 – not familiar with the word and its meaning*
- 2 – not very familiar with the word and its meaning*
- 3 – rather familiar with the word and its meaning*
- 4 – familiar with the word and its meaning*

5 – very familiar with the word and its meaning

In general, the familiarity rating scale used in psycholinguistic assessments is based on 7 levels of ratings (e.g. Lee, Liow, & Lee, 2007). However, due to the preliminary nature of the current study, the items were examined using a 5-point Likert scale rating system. If the respondents chose either 3, 4 or 5 as their answers, they were encouraged to give written meanings or synonyms to the words both in SMD and SM. A language background section and a separate final question asking respondents to inform whether or not they use these words in their daily conversation were also included. The words used were collected through observations and interviews with native speakers of SMD in and around Kuching, Sarawak. A mixture of 50 words from the SMD lexicon that are frequently and rarely used by native SMD speakers were chosen as target items. These words, however, were not controlled for frequency, type or meaning.

Participants

The participants of this study consisted of 11 female and 4 male undergraduates between the ages of 20 and 25 who are native speakers of SMD. They were grouped as a single generation of speakers closely associated with the Z generation. These are individuals born in or after 1990, which coincides with the birth of graphical Web resembling the internet of today. Twelve of them are from Kuching whilst the rest are from Kota Samarahan – a division 25km from Kuching town. Only 2 out of 15 respondents reported using SM instead of SMD as their most frequent spoken language at home.

Results and Findings

A mean count for each of the 50 SMD words resulted in about 38% of the total number of words rated as having low to very low familiarity rating. Those rated familiar and highly familiar made up 62% of the total items. The 19 words shown in Table 1 are under the familiarity mean of 3.0 and therefore can generally be considered as less known and not widely understood. These words have a high tendency of disappearing from the SMD lexicon if not documented thoroughly and systematically. For the complete table, please refer to the Appendix.

It was observed that words between the scales of 1.0 to 2.0 are mostly specialized terms used in carpentry with the exception of the adjective *ganjo* (tall) and the nouns *tadeng* (earrings), *tekoan* (teapot) and *godang* (clinic). Consequently, it is hardly surprising for the terms to be very unfamiliar to most participants. The four words mentioned earlier have meanings that are still relevant in Sarawak today and are usually substituted with words phonetically closer to those in SM. Words between the scale of 2.0 and 3.0 are also made up of adjectives, verbs, and nouns frequently used in everyday speech of native SMD speakers in and around Kuching and Kota Samarahan. These are possibly words that the participants may or may not have heard at some points in their lives but would not understand and would not know how to use these words in their daily conversations. None of the participants reported of having used any of the 19 words in their daily speech. Given the less familiar nature of these words, there were no synonyms or meanings provided by the participants.

Table 1

Ascending mean for perceived familiarity rating of SMD words below 3.00

Words	Mean	Standard Deviation
suit	1.2000	.41404
tadeng	1.2667	.79881
perepat	1.2667	.45774
kaso	1.4000	.82808
langkok	1.4667	1.06010
ganjo	1.6000	1.40408
tekoan/ tengkoan	1.6000	1.40408
godang	1.6667	1.39728
gelegar	1.7333	1.27988
menais	2.1333	1.80739
rembo	2.1333	1.68466
langar	2.2000	1.69874
ayong	2.2000	1.78085
ngansa	2.2667	1.62422
kolom	2.2667	1.75119
kelido	2.3333	1.75933
ladin	2.4667	1.88478
kenceng	2.4667	1.88478
pagu	2.6000	1.91982

Most of these unfamiliar words, however, can generally be classified as having high frequency meanings. The word *ladin* (knife) has largely been influenced by the SM word *pisau* but with an identifiable glottal sound for words ending in SMD, thus resulting in the term *pisok*. The word *ladin* may still be used by older generations of native SMD speakers (those born in 1950) when they speak with those from their own age group and their own children.

In the case of extreme word loss, it often involves words or phrases with meanings that are no longer relevant to a particular linguistic community or generation as in the case of the Inuit community in North Baffin Island (Bordin, 2009). These dying words are closely related to the Inuits' traditional culture and religion. Thus, a clear lexical discontinuity between generations of Inuit speakers is observable. For the present study, the word meanings of the 50 SMD words are still very relevant for the SMD speech community of today. Therefore, the unfamiliarity of these words to the Z generation, the younger generation of native SMD speakers, should not be viewed as a case of a total word loss but rather a generational shift in which a word or a phrase is used in preference over other words or phrases. It would be interesting for future research to examine urbanisation, language contact, and language policy as possible contributing factors to the phenomena affecting SMD.

Table 2

Ascending mean for perceived familiarity rating of some SMD words above 3.00

Word	Mean	Standard Deviation
kaleng	3.0000	1.96396
bicu	3.3333	1.79947
dasan	3.7333	1.75119
rangkak	4.2000	1.42428
jamek	4.2667	1.38701
mupok	4.5333	1.24595
lantoh	4.5714	1.08941
kumbuk	4.6000	1.05560
saruk	4.8667	.35187
uras	4.9333	.25820
mantak	5.0000	.00000
ngesor	5.0000	.00000

As for words with ratings above 3.0 (Table 2), they can be generally considered as familiar SMD words that are understood by the participants. These words range from nouns mostly consisting of household items such as *kaleng* (food container made of iron or steel) and *dasan* (place for storing things) to verbs and adjectives with high frequency meanings like *tapuk* (to put) and *mantak* (raw or uncooked). For this category of words, those rated comparatively low may have meanings that are less used in everyday speech. The word *kaleng*, for instance, has a relatively lower frequency meaning as opposed to the more frequently occurring meaning of the word *ngesor* (to show off).

All these familiar words, however, are not necessarily used for everyday conversation based on the feedback from some participants in the study. A possible reason is their preference for using synonyms and words phonetically nearer to those in SM with or without some noticeable dialect influence. Some examples are the words *uras* (garbage) and *mupok* (go home) which are often substituted with the words *sampah* and *balit* in SM (*balit* derives from the SM word *balik* but with an identifiable plosive “t” for words ending in SMD, thus resulting in word *balik* being pronounced as *balit*).

Conclusion

From the familiarity rating database, it is evident that there are a number of SMD words that are neither understood nor actively used by the younger generation of native SMD speakers. Previous claims of such words being gradually phased out by more standardised forms can now be supported by systematic numerical data. About one third of the words presented to the participants in this study were either not widely understood or non-existent in their lexicon. The use of SM as the primary medium of instruction in national schools may have affected their word choice. The familiarity rating method used in this study proves to be useful for investigations of linguistic phenomena such as lexical change. It can further strengthen information about the knowledge and usage of language items using solely observational methods. For dialects such as SMD which does not have sufficient written and historical documentation, a current frequency count database would not provide much information about the words besides those that are presently in active use in social media

sites by users of limited age groups. Although the language change phenomenon occurring in SMD is hardly a new finding, a study of this nature allows researchers to systematically document and build a repository of words in hope of preserving those in danger of being lost to a new generation of speakers. It is hoped that future research on SMD can address other aspects that are not covered in this study such as diverse and variegated population SMD speakers and lexical attributes (e.g., spelling, frequency, and regional differences). Other potential research can involve cross-generational investigations of language items, to see whether there are differences or similarities in terms of their understanding and usage.

References

- Bordin, G. (2009). *Lexical Discontinuities between Generations: Recent Inuit Cases from North Baffin Island. Anthropological Linguistics, 51(3), 191-208.*
- Collins, J. T. (1987). *Dialek Melayu Sarawak.* Kuala Lumpur, Malaysia: Dewan Bahasa dan Pustaka.
- Lee, L. C., Liow, S. J. R., & Wee, M.-L. O. (2007). Morphological structure of Malay: Using psycholinguistic analyses of rated familiarity. In M. Alves, P. Sidwell & D. Gil (Eds.), *SEALS VIII Papers from the 8th Annual Meeting of the Southeast Asian Linguistics Society 1998* (pp. 109-119). Canberra: The Australian National University.
- Pareira, V., Liow, R. S. J., & Mohd Saniff, S. (1992). *Familiarity ratings for 530 Malay words.* Working Paper No. 42, Department of Social Work and Psychology, National University of Singapore.
- Smith, A. D. M. (2007). Language change and the inference of meaning. In C. Lyon, C. L. Nehaniv & A. Cangelosi (Eds.). *Emergence of Communication and Language* (pp. 323-337). London: Springer.
- Zainal Abidin Merjan. (1992). *Daftar kata bahasa Malaysia Dialek Melayu Sarawak.* Kuala Lumpur, Malaysia: Dewan Bahasa dan Pustaka, Kementerian Pendidikan Malaysia.

Appendix

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
suit	15	1.00	2.00	1.2000	.41404
tadeng	15	1.00	4.00	1.2667	.79881
perepat	15	1.00	2.00	1.2667	.45774
kaso	15	1.00	4.00	1.4000	.82808
langkok	15	1.00	5.00	1.4667	1.06010
ganjo	15	1.00	5.00	1.6000	1.40408
tekoan_tengkoan	15	1.00	5.00	1.6000	1.40408
godang	15	1.00	5.00	1.6667	1.39728
gelegar	15	1.00	5.00	1.7333	1.27988
menais	15	1.00	5.00	2.1333	1.80739
rembo	15	1.00	5.00	2.1333	1.68466
langar	15	1.00	5.00	2.2000	1.69874
ayong	15	1.00	5.00	2.2000	1.78085
ngansa	15	1.00	5.00	2.2667	1.62422
kolom	15	1.00	5.00	2.2667	1.75119

kelido	15	1.00	5.00	2.3333	1.75933
ladin	15	1.00	5.00	2.4667	1.88478
kenceng	15	1.00	5.00	2.4667	1.88478
pagu	15	1.00	5.00	2.6000	1.91982
kaleng	15	1.00	5.00	3.0000	1.96396
telok	15	1.00	5.00	3.0667	1.83095
sandit	15	1.00	5.00	3.2000	1.69874
lang	15	1.00	5.00	3.2667	1.94447
ngeteng	15	1.00	5.00	3.3333	1.91485
bicu	15	1.00	5.00	3.3333	1.79947
tumpik	14	1.00	5.00	3.3571	1.90575
belunyor	15	1.00	5.00	3.5333	1.76743
sep	13	1.00	5.00	3.5385	1.85362
percak	15	1.00	5.00	3.6000	1.91982
ngepok	15	1.00	5.00	3.6000	1.76473
dasan	15	1.00	5.00	3.7333	1.75119
rutit	15	1.00	5.00	3.7333	1.86956
acap	15	1.00	5.00	3.8000	1.61245
ngintu	15	1.00	5.00	3.8667	1.64172
ampus	15	1.00	5.00	3.9333	1.53375
ipak	15	1.00	5.00	3.9333	1.62422
rangkak	15	1.00	5.00	4.2000	1.42428
jamek	15	1.00	5.00	4.2667	1.38701
betekang	14	1.00	5.00	4.2857	1.43734
aruk	15	1.00	5.00	4.4667	1.06010
tapuk	15	1.00	5.00	4.5333	1.06010
mupok	15	1.00	5.00	4.5333	1.24595
lantoh	14	1.00	5.00	4.5714	1.08941
kumbuk	15	1.00	5.00	4.6000	1.05560
saruk	15	4.00	5.00	4.8667	.35187
uras	15	4.00	5.00	4.9333	.25820
engkah	15	4.00	5.00	4.9333	.25820
cerik	15	4.00	5.00	4.9333	.25820
mantak	15	5.00	5.00	5.0000	.00000
ngesor	15	5.00	5.00	5.0000	.00000
Valid N (listwise)	12				