

# Lexical priming, dictionaries and Asian users of English

Vincent B Y OOI

National University of Singapore

[vinceooi@nus.edu.sg](mailto:vinceooi@nus.edu.sg)

## 1. Introduction

In keeping with the theme of the conference, "Advancing Language Teaching with Lexicography and Corpus Building", both the language teacher and the student can accelerate their priming of native speaker patterns by paying close attention to their own lexical primings vis-à-vis online lexical resources (particularly the dictionary) and corpus linguistic resources. Such close attention includes an awareness of 'frequencies of usage' and 'psychological salience', typified by the following parameters: local vs international usage, formal vs informal contexts, spoken vs written registers, Internet slang vs standard usage etc. In this paper, I draw on English primarily (as it is the language that I know best), but the observations can of course be extended to other languages.

We live in an exciting era in which there are multiple and competing free online lexical resources for the user's attention; at the same time, the plurality of lexical resources can often lead to a confusion as to which treatment is the 'correct' one. One can only surmise that this situation has arisen because lexicographers come from different traditions (e.g. British or U.S English) and emphasize particular linguistic claims based on both their intuitions and empirical evidence (corpora) used. As Lew (2011: 248) puts it, 'a great variety of dictionaries exist, and...without proper guidance, users run the risk of getting lost in the riches'.

This 'proper guidance' should foster in users a central awareness of lexical priming, defined by Hoey (2009) as 'the process of subconsciously noticing' such facts that native speakers make a mental note of when they encounter a word. For Michael Hoey, using the 'right phrase in the right context at the right time' means subconsciously noticing such factors as the following: (i) the words it occurs with; (ii) the grammatical patterns it occurs in; (iii) the meanings with which it is associated; (iv) its usage in terms of politeness vs rudeness, humour vs seriousness; (v) the style it tends to occur in; (vi) the registers that it occurs more often in: spoken vs written, academic vs novels, advertisements or newspaper writing, (vii) its usage by someone older or younger; and even (viii) its association 'with the beginnings or ends of sentences or with paragraph boundaries.'

Arguably, as the leading global language with the most functional load, English has become the most pluralized/pluricentric language. The existence of various 'Englishes' with unequal statuses ('Inner', 'Outer', and 'Expanding Circles' of

English – Kachru 1985) means that there are many potential conflicting lexical primings among countries in which English is regarded as native language (i.e. the UK, U.S, Australia, New Zealand, Canada etc). In addition, ‘Outer Circle’ countries (cf. Kachru 1985) such as the Philippines, Malaysia and Singapore have formed their own institutionalised/nativised/endonormative standards that may conflict with native speaker ones. Given the expectation that diglossia is an everyday feature in such situations, I suggest that the Concentric Circles Model (Ooi 2001) – constructed specifically for Singaporean and Malaysian contexts, but which might be extendable to other nativised Englishes – may be useful to alert users on the appropriateness of local usage in H-igh and L-ow situations, as it would also relate local norms to global ones. Lexical primings do not just exist in the minds of speakers, and a resource such as the GloWbe corpus (Davies and Fuchs 2013) reflects the collective lexical primings of 6 native speaker countries and 14 Outer Circle ones.

In this paper, a few examples will illustrate the uneven treatment, incomplete varietal usage and even cultural bias in some well-known English lexical resources. The overall ‘takeaway’ is that the Asian user of English will continue to draw on the Web for its free lexical resources and well-motivated contemporary corpora of English; he/she has to be more self-directed in reconciling lexical primings – often conflicting -- in both local usage (endonorms) and international ones (exonorms). Lexicographers have to be even more nimble in catering to much more Web-savvy audiences nowadays.

## **2. Lexical priming theory**

Hoey (2014) reiterates his theory, first expounded in 2005, that lexical priming theory seeks to ‘integrate... psycholinguistic research ... with the findings of corpus linguists.’ This theory postulates that each person’s knowledge of language is the result of all that we have heard and read repeatedly. In other words, the lexicon is probabilistic in nature. One of the more well-known proponents of a probabilistic lexicogrammar is Michael Halliday, who states that his own view of grammar (and by extension, the lexicon) is ‘inherently probabilistic’ (Halliday, 1982: 65). Also, in my own Lexical Frame Analysis, one of the tenets postulated regarding the lexicon is that it is probabilistic in nature. (Ooi, 1998).

Given that each person has varying ‘lexical frequencies’ from someone else, how is it possible that we are able to communicate with one another and have a notion of standard English? For Hoey (2005), conflicting primings are reconciled through such common factors as education, mass media, literary and religious traditions, grammars and dictionaries.

## **3. Ooi’s Concentric Circles Model**

A variety of English may be construed as one that has collective lexical primings understood and agreed upon by the speech community it typifies. Besides common frequencies of usage, the words distinctive to and give a particular cultural identity

to a speech community may be 'graded' in terms of their diglossic nature. Illustrating the complexity of English use from the small island of Singapore (see Ooi 2013), the range there may be characterised as follows: standard English (which usually means standard British/US English), SgE-H (educated, 'standard' Singaporean English) and SgE-L (Colloquial Singapore English, also popularly known as Singlish). Standard English is the benchmark set in grammar books, dictionaries, classroom texts, official media and print materials. SgE-H may be considered the local H(igh)-variety used and comprises 'invisible norms' that have to be distinguished from those of native English speakers. SgE-L, as the L(ow) variety, is used in more domestic situations and is the popular colloquial variety found in (among others) casual conversations, television humour programmes, online personal blogs, chat rooms and discussion forums. I would prefer the term SgE-L to 'Singlish' which has become a loaded term that is either hated or loved among Singaporeans. Unlike 'broken English' (which has no rules), the word order in SgE-L tends to be rule-governed. For example, the well-known SgE-L expression *Why you so liddat?* ('liddat' = 'like that', meaning 'Why do you behave in this manner?') would be regarded as ungrammatical by native SgE-L speakers if it was rephrased in one of the following ways:

\**Why you liddat so?*

\**Why so you liddat?*

While SgE-L expressions may be said to derive mainly from Chinese or Malay, native SgE-L speakers would probably also tend to agree that the expression *everyone is good* (to mean 'Hello, everyone') is not so much Singlish (if at all) as it is Chinglish. This expression is literally translated from the Chinese expression 大家好 into English.

The three aspects of English just outlined may be diagrammatically represented and 'graded' in terms of a 'Concentric Circles Model' (Ooi, 2001) as follows:

- (i) Circle 1 represents the inner circle that comprises 'core or standard English' linguistic expressions that may or may not be traditionally Germanic/ French/Latin in origin. Non-Anglo expressions that are codified and standardised in dictionaries nowadays include *kungfu*, *sari* and *lychee*. Circle 1 items are unmarked and deemed acceptable internationally.
- (ii) Circle 2 (SgE-H) is the next outer circle containing linguistic expressions from English acceptable in more formal local situations, but go beyond their conventional meaning in Western discourse. For instance, *killer litter* is a hybrid of two intriguing and yet unlikely juxtapositions in Western discourse, i.e the seriousness of 'killers' and the social irresponsibility of throwing harmless garbage on the ground. In Singaporean discourse though, this productive neologism is needed in a densely populated society of high-storey buildings. For many expressions, the influence from

- a local language or dialect is obvious. An example is *sleep late*. In US English, a song title such as *I like to sleep late in the morning* makes it clear that one gets up late in the morning; in SgE-H, it would be much more common to refer to staying up late at night. Thus, a sentence such as *I like to sleep late at 3am* would have its corresponding meaning in either Malay ('tidur lewat') or Chinese ('wan shui'). Pakir (2009: 85) reminds us that 'Singapore offers an example of a (tropical) country where spontaneous daily interaction among speakers of several languages over a long period of time has led to (various linguistic) innovation processes' and semantic shifts which differ from the conventional Western discourses that we tend to associate standard English with. Thus, "while Australians might find the description of 'windy' for homes as a negative feature in the (real estate) 'for sale' advertisements, Singaporeans consider 'windy' as 'breezy' and therefore a positive feature" (Pakir, 2009: 96). Another noteworthy point is that the sense of 'breeziness' would not tend to invite the charge of Singlish among Singaporeans.
- (iii) Circle 3 (SgE-H) is the next outer circle containing linguistic expressions that are also acceptable in formal local situations and go beyond their conventional understanding in Western discourse. However, unlike those in Circle 2, the items in this circle contain loanwords and expressions from other local languages (principally Chinese and Malay). There are no English equivalents without missing local associations. Examples include *silat* ('Malay kungfu'), *songkok* ('Malay hat'), *laksa* ('a popular curry dish') and *ice kacang* ('a dessert of shaved ice with various flavours and toppings' that can include *kacang*, a Malay word for 'peanuts'). Conceivably, Malay words such as *durian* (whose first use is attested by the Oxford English Dictionary as early as 1588) and *rambutan* (coined in 1707) would be in this circle instead of Circle 1. However, these words have stood the test of time and are now accepted by the world's English-speaking community. Hence, the words *durian* and *rambutan* rightfully belong to Circle 1.
- (iv) Circle 4 (SgE-L or Singlish expressions taken from English) is the next outer circle of English-derived expressions that are deemed suitable for local colloquial or informal situations only. In this circle, structures from colloquial Chinese or Malay have their literal English equivalents. Thus, the expression *I follow Mother to the market* does not mean that the interlocutor walks behind but instead accompanies the mother. In casual conversations, *blur* is also used as an adjective, to mean 'confused or dazed'. The late Minister Mentor Lee Kuan Yew used this example of 'blur' as a prototypical Singlish feature.
- (v) Circle 5 (SgE-L or Singlish expressions taken from other languages or dialects) represents the outermost circle of least transparency (in relation to 'core English') in having terms of non-English origin that are

primed for local colloquial situations only. *Makan* is the Malay word for 'to eat', *paktor* is the Cantonese Chinese word for 'dating', and *kiasu* is a word borrowed from Hokkien that means 'afraid to lose out'. Proponents of the 'Speak Good English Movement' in Singapore would probably be appalled to learn that the publishers of Oxford Dictionaries ("the world's most trusted dictionaries") have included both *kiasu* and the Malay word *lepak* ('loafing') – labelling them as 'South East Asian' usage. Both Circles 2 and 3 would be acceptable in more formal situations, for example classroom reports, newspaper editorials and broadcast news (see Ooi, 2007 and Low, 2010, for a range of other expressions in Circles 2 and 3). However, both Circles 4 and 5 would be acceptable in colloquial or highly informal situations only (usually speech) and contain terms that are popularly known as 'Singapore Colloquial English' or Singlish. This way of thinking about English in Singapore, ranging from Circles 1 to 5, would remove a lot of the linguistic anxiety surrounding the use of a linguistic expression as either Singlish or 'standard English'. We can use the model to relate domestic English to international English usage; at the same time, it would help foreigners to avoid cross-cultural miscommunication (as in the case of 'windy') when they first arrive in the country.

If a model is a miniature representation that shows the entire situation at a glance, then the following diagram (Figure 1) may be worth presenting to the teacher and the student of English in Singapore:

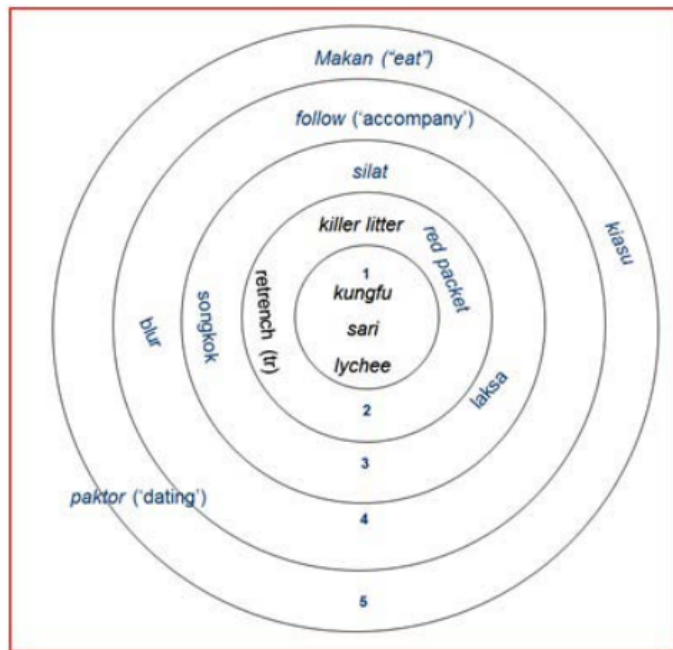


Figure 1 Concentric Circles of English for Singapore English (Ooi 2001, 2013)

In this regard, the practice of dictionaries to use mere labels ('formal', 'informal', 'colloquial' etc) for each lexical entry neither allows the user to see the gradations in formality of usage nor have a picture of the entire range of English use at a glance.

### 1 Reconciling different primings in dictionaries with corpus evidence

In this section, I would like to present four examples of how a lexicogrammatical item gets varying treatments in some free online dictionaries of English, and how complementary evidence from well-known contemporary corpora might assist in reconciling their different primings.

#### i) *brownie points*

Merriam Webster Learner's Dictionary (MWLD) defines it as "praise, credit, or approval that a person gets from someone (such as a boss or teacher) for doing something good or helpful"; example sentences list associated verbs as "earn", "win", or "get" (see Figure 2):

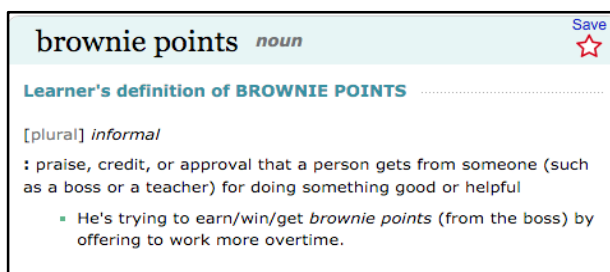


Figure 2 *brownie points* (Merriam-Webster Learner's Dictionary)

After reading this lexical entry, the learner is no wiser concerning the preferred verb associated with the term and its more appropriate context(s) of usage, i.e. spoken, written, academic etc. More puzzlingly, Cambridge Dictionaries Online (CDO) lists the term as American usage – although the British setting is used -- and "humorous" in nature (see Figure 3):



Figure 3 *brownie points* (Cambridge Dictionaries Online)

Meanwhile, the COBUILD Dictionary flags it for "disapproval", i.e. a negative semantic prosody because the approvers for brownie points can be politicians of dubious public standing (see Figure 4):

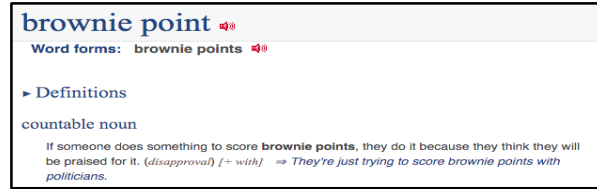


Figure 4 *brownie points* (Collins COBUILD Dictionary)

Which of these entries should the learner of English believe in? And, to further problematize the issue, the overwhelming predominance of Google as the leading global search engine (except for a country such as mainland China in which it is banned) means that a typical user is nowadays likely to use Google's dictionary (which has become a serious contender to other dictionaries) and its corresponding Wikipedia entry – and stop there (without even using CDO, COBUILD or MWLD). If Google's dictionary is to be believed, the issues raised earlier are rather non-existent: *brownie points*, it says, “are a hypothetical social currency, which can be acquired by doing good deeds or earning favor in the eyes of another, often one's superior.”

Notwithstanding Google's Dictionary as the one-stop place for lexical information, we may turn to the GloWbe corpus which shows the following frequencies for *brownie points* across 20 different countries (see Figure 5).

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK
BROWNIE POINTS	763	97	29	170	34	111	30	92	36	20	9	26	24	11	8

Figure 5 Frequencies of occurrence for *brownie points* (per million words) in the GloWbe corpus

Thus, contrary to the Cambridge Dictionary, the term is not only U.S. usage but very much so occurs in British, Australian and Indian contexts.

As for the preferred verb collocates, Figure 6 shows an edited listing for the verbs 'score', 'get', 'earn', 'win' and 'gain':

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK
SCORE BROWNIE POINTS	47	2		7		2		11	13	1		2	3		
GET BROWNIE POINTS	25	9	1	6	1	4		2			1				1
EARN BROWNIE POINTS	21	3	1	5	1	4		3				1		2	
WIN BROWNIE POINTS	17	4	3	2		1		1		2					
GAIN BROWNIE POINTS	11	1		2	2		2	1	1	1			1		

Figure 6 Edited listing for search [v\*]+ *brownie points* (GloWbe corpus)

If Outer Circle countries also get to vote, then “score” seems to emerge as the top verb collocate (as it does in the UK). In the U.S., “get” seems to be the preferred verb collocate – of course, the accuracy of these frequencies is dependent on the part-of-speech program used for the corpus.

Finally, an analysis of the concordance listings for the term suggests that its semantic prosody tends to be negative when associated with political and serious matters but

more light-hearted and positive when associated with personal matters (see Figure 7):

US B	townhall.com	A	B	C	never be proven -- only alleged. And if it is easy to give yourself <b>brownie points</b> for happy thoughts, it is even easier to demonize you
US B	...tart.bizjournals.com	A	B	C	firm, issued private warnings as far back as 2002 -- they have received few <b>brownie points</b> since the housing bust began. " Everybod
US B	wow.joystiq.com	A	B	C	purpose Balnazzar is trying to use the scarlet crusade. destroy the scourge and score <b>brownie points</b> with Kil'jaeden. so Sylvanas is d
US B	adaringadventure.com	A	B	C	write a post linking to this one, then that will definitely earn you extra <b>Brownie Points</b> . # The reason being, is I get an inbound link (\
US B	adaringadventure.com	A	B	C	the best choice to be your next mentee! # P.S. Do I get bonus <b>brownie points</b> for also having been a guest poster on your site? # By
US B	barrowjournal.com	A	B	C	play book..... your rants and personal attacks should earn you some nice Obama " <b>brownie points</b> ". Well done! D # To leave a com

Figure 7 Short concordance listing of *brownie points* (U.S. component, GloWbe corpus)

## ii) *durian*

Similar to Ooi (2010), my choice of this term is to examine the cultural bias associated with its definition and characterisation. Because the smell – and taste – of the fruit puts off many Westerners (lexicographers included), it gets a negative semantic prosody which then gets copied from dictionary to dictionary. Let us take a look at what comes up in both the Google and Yahoo search engines when the term 'durian definition' is keyed in (Figures 8 and 9):

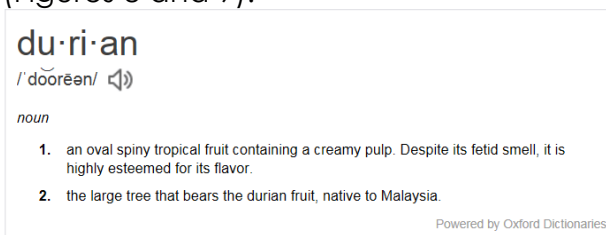


Figure 8 Keying in 'durian definition' (Yahoo search engine)



Figure 9 Keying in 'durian definition' (Google search engine)

There is a striking similarity to both definitions. In the Yahoo one, there is also the inclusion of 'powered by Oxford Dictionaries'; in the Google one, there is the inclusion of its etymology (from Malay) and its increasing frequency of use. In both cases, the choice of the term 'fetid smell' indicates that its odour is unbearable to Oxford's lexicographers but the taste is much more bearable. The Cambridge Dictionary chooses the term 'very strong smell', whereas the MacMillan Dictionary is less equivocal in characterising its 'strong, unpleasant smell but with a sweet flavour' (but, of course, there are durians that have a bitter flavour too). Despite the unpleasantness of the smell (which would be fragrant to many Asian lovers of durian), having a lexical entry for it is arguably better than its treatment in the MWLD and the Longman Dictionary – both of which exclude the term and therefore think that it would be unnecessary to their learners to acquire any



knowledge of the term. Perhaps, this is justified when is frequencies of occurrence in the GloWbe corpus are examined (see Figure 10):

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK
DURIAN	2271	33	14	29	1	115	26	20	80		3	938	826	110	48

Figure 10 Frequencies for *durian* (GloWbe corpus)

Despite the millions of people in Inner Circle countries, the fruit is clearly not celebrated in those contexts. The greatest connoisseurs of the fruit are clearly those in Malaysia and Singapore. In these countries, the term collocates productively with words such as ‘profiteroles’, “puffs”, “breath”, and even “mousse cake”.

iii) *take a look (at) vs have a look (at)*

In their study of one-million-word corpora of British and U.S. English, Leech et al (2009: 179) make the observation that ‘expanded predicates with *have*, *take* and *give* are used more frequently in fictional rather than non-fictional texts’; they also confirm the hypothesis that ‘BrE prefers *have* as a light verb in expanded predicates whereas AmE prefers *take*. These regional differences are more pronounced in the spoken than in the written data.’

Let us re-examine their claims by looking at both the BYU-COCA and BYU-BNC corpora for the phrase *take/have a look (at)*. In the Corpus of Contemporary American English (COCA), *take a look* occurs 6673 times whereas *have a look at* occurs only 446 times. In the British National Corpus (BNC), *take a look at* occurs only 263 times whereas *have a look at* occurs 905 times. So, it would certainly appear that the hypothesis by Leech et al do hold water.

Figure 11 shows that British English prefers *have a look \** in speech and to a lesser extent in fiction:

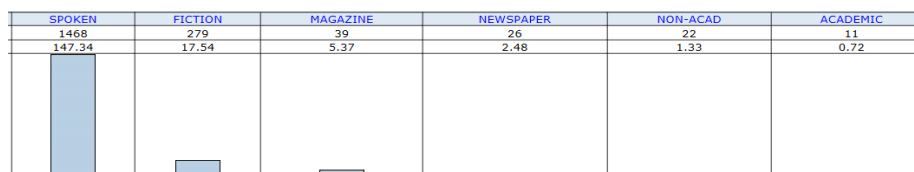


Figure 11 *have a look \** (BNC)

The picture for *take a look \** is more complex though (see Figure 12), with a preference for more contexts of use:

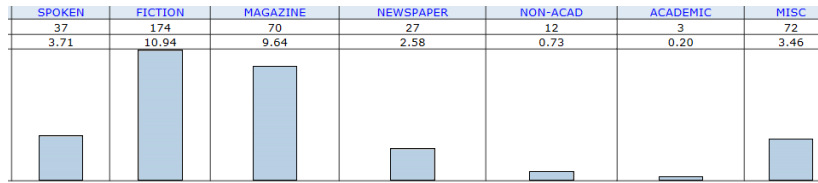


Figure 12 *take a look\** (BNC)

Unlike *have a look* in the BNC, the preferred context for the term in COCA is in the fiction and (to a lesser extent) spoken registers (see Figure 13):

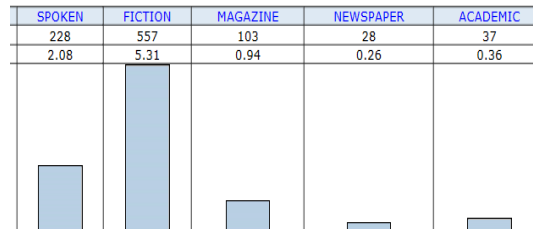


Figure 13 *have a look\** (COCA)

For *take a look\**, U.S. English also sees the term being preferred in spoken contexts mainly (see Figure 14):

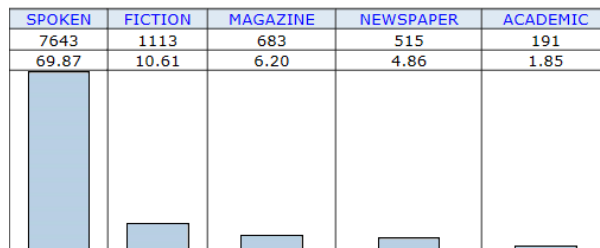


Figure 14 *take a look\** (COCA)

We would thus expect that the richness of such information in these two corpora would get reflected in the dictionary. Instead, various dictionaries merely exemplify the interchangeable usage of both verbs in example sentences (Figure 15):

[COUNTABLE] [USUALLY SINGULAR] an act of looking at someone or something

**have/take a look at:** *Can I have a look at your new skateboard?*

**have a good/close look (=look carefully):** *He got out of the car so he could have a closer look.*

Figure 15 *have/take a look* (Macmillan Dictionary)

iv) *economy rice / economical rice*

The final example here is chosen to show that there are many Asian-oriented terms which do not yet make it to Circle 1 of my Concentric Circles Model (CCM – see Section 3). For global recognition, such terms will have to be used in one or

more Inner Circle countries. A case in point is *economy rice/economical rice* which is lexically primed for Singaporean and Malaysian users. Referring to the working person's 'quick and hearty meal' of rice with a few inexpensive dishes, the term makes it to Circle 2 of the CCM and is neither regarded as 'Manglish' nor 'Singlish'. Figure 16 gives an indication of its usage in the GloWbe corpus, and many more examples will also turn up with a google search:

Dr Lai said green apples are better for this treatment. # Even my **Economy Rice** Stall lady boss, who is a nice lady in her 60? s and puff or roti prata at the staff cafeteria, for lunch he usually has **economy rice**. I don't know what else he eats the rest of the day, sandwiches, noodle soup, porridge, fruit, etc. Again choices for **economy rice** at lunch should include less fried or coconut-base items ; banana leaves. S'pore Nasi lemak has so many side dishes like Chinese' **economy rice**', nasi lemak here should called nasi padang. # I comments: # Ooo... Love this dish! Always find it at the **Economy Rice** Stall. Think this is what lunch would be:) Do you know what

Figure 16: Concordance of *economy rice* (Singapore component, GloWbe corpus)

#### 4. Some concluding remarks

In this day and age, the multiplicity of free online lexical resources and available corpora means that the Asian user of English has to navigate through and utilise such resources well. While such learners of English would like to accelerate their lexical primings in order to achieve 'naturalness' and 'nativeness', they also need to be aware of and reconcile any conflicting lexical frequencies from their first language and their own local variety of English. It would be good for the student and the teacher to triangulate the use of various lexical resources and online corpora. At the same time, lexicographers will have to improve on their lexical evidence base and codification efforts in order to maintain and enhance the richness of their dictionaries.

## References

- Davies Mark, and Robert Fuchs. (2015) Expanding horizons in the study of World Englishes with the 1.9 billion word Global Web-based English corpus (GloWbe), In *English World-Wide* 36:1, pp1-29.
- Hoey, Michael. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.
- Hoey, Michael. (2009). Lexical priming, in *MED (Macmillan English Dictionaries) Magazine*, Issue 52. URL: <http://www.macmillandictionaries.com/MED-Magazine/January2009/52-LA-LexicalPriming.htm> (Accessed: 1 May 2016)
- Hoey, Michael. (2014) Words and their neighbours, in J R Taylor (ed) *The Oxford Handbook of the Word*, Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199641604.013.39
- Kachru, Braj. (1985) Standards, codification and sociolinguistic realism: The English language in the Outer Circle, in R. Quirk & H. Widdowson (eds.) *English in the World: Teaching and Learning the Language and Literatures*, Cambridge: Cambridge University Press, 11-30.
- Leech, Geoffrey, Marianne Hundt, Christian Mair, and Nicholas Smith. (2009) *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- Lew, Robert. (2011) Online dictionaries of English, in P A. Fuertes-Olivera and H. Bergenholtz (eds) *e-Lexicography: The Internet, Digital Initiatives and Lexicography*, London: Continuum, 230-250.
- Low, Ee-Ling. (2010) English in Singapore and Malaysia: differences and similarities, in A Kirkpatrick (ed) *The Routledge Handbook of World Englishes*, Routledge: London and New York, 229-246.
- Ooi, Vincent B Y. (1998) *Computer Corpus Lexicography*. Edinburgh. EUP.
- Ooi, Vincent B Y. (2001) Globalising Singaporean-Malaysian English in an inclusive learner's dictionary, in Moore B (ed) *Who's Centric Now?: The Present State of Post-Colonial Englishes*, Oxford: Oxford University Press, 95-121.
- Ooi, Vincent B Y. (2010) English Internet lexicography and online dictionaries, in *Lexicographica* 26, Berlin: de Gruyter, 143-154.

Ooi, Vincent B Y. (2013) Branding Singapore through Singapore(an) English?,  
Commentary (Journal of the National University of Singapore Society),  
Vol.22, 41-48.

Pakir, Anne. (2009) Lexical variations in 'Singapore English': linguistic description  
and language education, in Y Kawaguchi, M Minegishi, and J Durand  
(eds), *Corpus Analysis and Variation in Linguistics*, Amsterdam: John  
Benjamins, 83-100.