

*Words, Dictionaries and Corpora: Innovations  
in reference science*

# PROCEEDINGS OF ASIALEX 2015 HONG KONG

Edited by:

LI Lan

Jamie Mckeown

LIU Liming



rm in office but  
ad. The lads had  
nest of \$70m and  
a suicidal and  
little more than  
first set after  
on the strike.  
the latest in

a string of  
a string of  
a string of  
a string of  
a string of  
a string of  
A string of

ectoral  
email excha  
embarrassing fail  
emergency operatio  
endorsements from  
epileptics whose c  
episodes showing C  
errors before winn  
evenly

Department of English

Hong Kong Polytechnic University

Words, Dictionaries and Corpora: Innovations in reference science

# **Proceedings of ASIALEX 2015 Hong Kong**

Edited by **LI Lan, Jamie MCKEOWN and LIU Liming**  
Hong Kong Polytechnic University  
June 2015

©The Hong Kong Polytechnic University, 2015

All rights reserved. No part of this book may be reprinted or reproduced or utilize in any form or by any means without permission in writing from the publisher.

ISBN 978-962-367-790-5

Printed and bound in Hong Kong

# A Dictionary of Hong Kong English vs The GloWbe Corpus?: Triangulating sources of lexicographic evidence

Vincent B Y Ooi

National University of Singapore

vinceooi@nus.edu.sg

## Abstract

This paper addresses the general question of reconciling differing and competing sources of evidence for ensuring that both frequency of occurrence and ‘socio-cultural salience’ are achieved in codifying a particular variety of English. In so doing, it compares the efforts of Cummings and Wolf (2011) and Davies and Fuchs (2015) with regard to Hong Kong English. Cummings and Wolf (2011) look at various sources of evidence for their selection of entries in the dictionary, and lexical items are also given a frequency rating based on their ‘brief Google.hk surveys’ (the highest two tiers having more than 1000 occurrences on Google.hk). In this paper, ten selected lexical items from Cummings and Wolf (2011) that have these high frequencies are compared with their respective (non-) occurrences in the Hong Kong component of the GloWbe corpus that contains 1.9 billion words from 20 different countries (Davies and Fuchs 2015:2). The texts in GloWbe are taken from various web pages and blogs that presumably represent both currency and socio-cultural salience. The selected lexical items for discussion include the following: *Mid-levels*, *dim sum*, *jetso*, *iron rice bowl*, *yum cha*, *wushu*, *kowtow*, *qigong*, *laissee* and *Legco*.

**Keywords:** *lexicographic evidence, · corpus, frequencies, GloWbe Corpus, Hong Kong English, triangulation*

## 1 Introduction

Meyer and Nelson (2006) detail the types of linguistic data – introspection, experimentation, and corpus – and some methodological issues that these types of linguistic evidence bring about in support of linguistic claims. Such types of linguistic data may similarly count as lexicographic evidence for the codification of lexemes in the dictionary. However, the proportion of these types of lexicographic evidence still varies among lexicographers, and there is no necessary uniform practice for the corroboration of such varying sources of lexicographic evidence in order to achieve the valued social sciences notion of ‘triangulation’.

In this paper, some of these triangulation issues are discussed in terms of the codification of 10 lexemes in *The Dictionary of Hong Kong English* (Cummings and Wolf 2011) and their corresponding (non-) occurrences in the Hong Kong English component of a recently completed mega corpus, *The GloWbe Corpus* (Davies and Fuchs 2015). While it is obvious that one is a dictionary and the other a corpus, this paper hopes to show that both enterprises may benefit from and complement each other in the quality of their linguistic claims.

## 2 The Dictionary of Hong Kong English (DHKE)

Why single out *The Dictionary of Hong Kong English* (henceforth ‘DHKE’) for our discussion on triangulation? Reasons include its methodology and various reviews by well-known linguists and lexicographers.

In its introduction, the DHKE lists itself as the first reference dictionary of Hong Kong English and that “only words and word senses that are particular to HKE (Hong Kong English) or have a specific reference to Hong Kong are listed.” (p. xii) However, ‘particularity’ does not mean ‘exclusivity’, and Cummings and Wolf acknowledge that these

terms may be similarly used in another (Asian) variety of English. The selection of entries is made “from a variety of sources”, including English-language newspapers based in Hong Kong, cartoons and literary works by local authors, government information, local student essays, spoken language heard on TV, and internet websites. The authors mention having consulted various literature on Hong Kong English, including “short lists” of Hong Kong English words (encapsulated in Bolton 2003). Also, while the authors mention the use of frequency information from the Hong Kong component of the International Corpus of English (Greenbaum and Nelson 1996), they seem to have mainly used “brief Google.hk surveys...at different times in the year 2010” (p. xxvi). The top band includes “more than 10,000 occurrences on Google.hk” and the second next most frequent band includes between 1000 and less than 10,000 occurrences on Google.hk. While giving readers an idea of the frequency of the items, Cummings and Wolf (2011) also acknowledge the limitations of such an approach: the data from Google.hk are not exclusively curated for Hong Kong usage alone, and the salience of items may vary over time in the media. Nonetheless, readers should still be able to “fathom the importance of widespreadness” of the items included in the dictionary.

In other words, what has been employed is a mix of introspection, citational evidence (casual occurrences in various print and spoken sources), a curated dataset (the ICE corpus), and a non-curated dataset (“brief Google.hk” surveys). While the sources of citational evidence are listed in DHKE, the proportions of these sources of evidence are not detailed. At the end of the day, it would very much depend on the lexicographers’ overall experience and judgment on the inclusion and treatment of these entries in the dictionary. In this regard, Salazar (2014: 103) cites an article of mine (Ooi 2001) for necessitating both etic (outsider) and emic (insider) perspectives in lexicography. Salazar concurs with me that such perspectives are necessary for “the accurate description of the lexicon of world Englishes, including those of the Southeast Asian region.” The distinction between ‘etic’ (from ‘phonetic’) and ‘emic’ (from ‘phonemic’), first introduced by the late linguist Kenneth Pike, is to offer a methodological solution to the problem of objectivity in the social sciences. Applying this distinction to lexicography, knowing the lexicographer’s etic and emic perspectives, i.e. where s/he is ‘coming from’, would provide greater understanding regarding the inclusion, exclusion, and treatment of entries in the dictionary. Applying this distinction to Cummings, it would seem that the etic experience is somewhat complemented by a somewhat emic one of having been in Hong Kong for 10 years; for Wolf, who is introduced as living in Germany, it would very much seem to be mainly an etic experience of Hong Kong. And, in my case, it is one of knowing Malaysian Cantonese natively, learning Malay (and English) in school in Malaysia and Singapore, being exposed to both Mandarin and Cantonese Chinese in Singapore for more than half my life, being a frequent visitor to Hong Kong during the past 10 years or so, and interacting with cousins in Hong Kong.

So, in reading through the approximately 460 entries selected for DHKE, one can safely say that a number of them are certainly shared by both Malaysian and Singapore(an) English. Such non-unique entries in DHKE include *iron rice bowl* (a term quite popular in Singapore at one time), *pajamas* (U.S. spelling) or *pyjamas* (British spelling), *bird’s nest*, *Hakka*, *cheongsam*, and *bubble tea* (which, anyway, is a term borrowed from Taiwan English). And, while DHKE asserts that *ketchup* originates from either Hong Kong or Chinese English, both *The Merriam-Webster Dictionary* and *Dictionary.com* cite Malay (‘kechap’ – fish sauce) as the origin.

Notwithstanding such friendly ‘ownership’ disputes, the 10 lexemes for examination in this paper will be based on the top two frequency bands used by the DHKE. This approach is meant to offer a more up-to-date basis for comparison with the Hong Kong component of the GloWbe corpus which was completed in 2013.

### 3 Global Web-based English Corpus (GloWbe)

The Global Web-based English Corpus (henceforth, ‘GloWbe’) contains approximately “1.9 billion words of text from 20 different countries, which includes six Inner Circle and 14 Outer Circle countries” (Davies and Fuchs 2015:2), based on Braj Kachru’s well-known distinction between countries in which English functions as a Native Language (hence, ‘Inner Circle’ countries) and those that have English as a Second Language (hence ‘Outer Circle’ countries). These 20 countries include the United States, United Kingdom, Canada, Australia, Ireland, New Zealand, India, Sri Lanka, Pakistan, Bangladesh, Singapore, Malaysia, the Philippines, Hong Kong, South Africa, Nigeria, Ghana, Kenya, Tanzania, and Jamaica.

In order to achieve the current size of the corpus, the only possible source is web pages (see also Kilgarriiff and Grefenstette 2003). Davies and Fuchs (2015) also explain that 60 per cent of the corpus comes from blogs, so that informal language is represented “fairly well”. The remaining 40 per cent consists of “more formal genres and text types.” Also, the curation of this corpus goes beyond what has been achieved in Cummings and Wolf’s brief surveys of google.com.hk. Davies and Fuchs (2015: 4) say that they use Google’s Advanced Search and “Limiting by Region” features in order to ensure that the web pages are correctly associated for the respective countries. They also remove “boilerplate” material (recurring headers, footers, sidebars etc. ), tag the entire corpus, and then import the texts into a database such that these texts from the various countries can be compared side-by-side in a normalized manner i.e. frequency per million words.

#### 4 Discussion of 10 lexemes from DHKE

As mentioned in Section 2, lexemes that will be discussed in this section are first filtered from the top two frequency bands (reflected at the bottom right of each DHKE figure in this section). In addition, the terms are selected based on what is felt about their distinctiveness relative to other Asian varieties of English, notably both Malaysian and Singapore(an) English that I have first-hand knowledge of. This ‘feeling’ will then be gauged more objectively by means of the GloWbe corpus.

##### 4.1 Mid-levels

The DHKE lists the term as follows (see Figure 1):

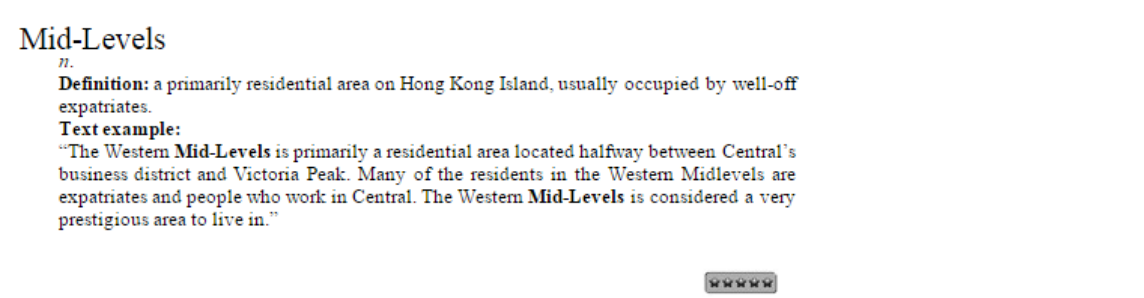


Figure 1 *Mid-levels* (DHKE)

Let us compare the corresponding evidence from GloWbe (see Figures 2 and 3):

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ	JM
MID-LEVELS	271	8	1		2	5	2	3	1		1	3	3	1	240		1				

Figure 2 *Mid-levels* (GloWbe)

the line you'll also get to see the jungle of housing skyscrapers in **Mid-Levels**, the preferred living quarters for expats. There are several junctions along the route , Hong Kong as the venue for the exhibition. Central, Sheung Wan and **Mid-Levels** are prime locations in the commercial heartland of Hong Kong Island and the accessibility of option for families because homes typically have more space. # Popular expatriate neighborhoods include **Mid-Levels** (the area about halfway up Victoria Peak from the harbor, so some short hikes starting from Aberdeen Reservoir to the Peak, then down to the **Mid-Levels** or the Western District. Two of these hikes from the picnic place down the west." # Flat by Anthony Hindmarsh 660 square feet Caine Road, **Mid-Levels** Coffee Assembly # Tucked into a nondescript building in SoHo, Coffee Assembly is Many newly developed residential buildings in Hong Kong, especially those in premium areas like **Mid-Levels**, Pokfulam and Kowloon Tong, are quickly taken up by those Chinese. ? Is it a 3 bedroom house in mid-level or a 3 bedroom apartment **mid-levels** or is it a place in Stanley? I mean these are the kind of you can see from another thread on Asiapat, there are a group of us **Mid-Levels** mums meeting up this Friday afternoon at Pacific Coffee if you fancy coming along flip the property for a quick buck. I'm predicting Chelsea prices soon for **Mid-levels**. # Yes. And Chelsea, I think, is still more expensive for

Figure 3 Selected concordance listing of *Mid-levels* (GloWbe)

GloWbe's *Mid-levels* (ranked the highest in C&W) occurs 240 times for Hong Kong (out of the 271 occurrences world-wide) – hence, there is a close fit between GloWbe and google.hk. In other words, we can say that this term is very much particular/distinctive to Hong Kong English. In addition, the concordance listing indicates that the term indicates richer associations of 'premium areas', 'preferred living quarters for expats', 'prime locations' and even 'mums' – beyond the mere collocate of 'Western' in DHKE.

## 4.2 Dim sum

In the GloWbe corpus, another popular term *dim sum* occurs only 529 times (for Hong Kong) out of 1584 times world-wide, showing that the term has spread to other countries such as the UK, U.S., Canada, Malaysia and Singapore (see Figure 4)

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ	JM
DIM SUM	1584	102	109	116	19	39	11	18	7	3	14	297	213	56	529	27	2	6		4	12

Figure 4 *dim sum* (GloWbe)

DHKE has a section entitled "Hong Kong English words now used internationally", and the term *dim sum* is not in this section; the evidence from GloWbe shows otherwise. The term is also found in both *Dictionary.com* and *The Merriam-Webster Dictionary*, with the origin listed as "Chinese (Guangdong)". The Guangdong part is corroborated in the corresponding DHKE entry (see Figure 5), in which the source language is indeed Cantonese:

## dim sum

/dim sam/ n.

**Source language:** Cantonese (點心).

**Definition:** a variety of different traditional Cantonese foods usually of a size to be picked up by chopsticks. Examples include *har gow* and *shaomai*.

**Text example:**

"They ordered a tea, and a steamer basket of tripe, the only **dim sum** the kitchen had left."

**Note:** The spelling *tim sum* is obscure in HKE, but more common in Singapore.

\*\*\*\*\*

Figure 5 *dim sum* (DHKE)

DHKE further claims that the alternative spelling *tim sum* is "more common in Singapore." On checking GloWbe, there are only 2 and 4 instances of this term from Singapore and Malaysia respectively. Instead, the preferred term in Mandarin-speaking Singapore is *dian xin* (which occurs 5 times in Singapore, and 3 times in Hong Kong respectively).

## 4.3 jetso

At the other end of the spectrum, there is an apparent conflict between google.hk and GloWbe, with the term *jetso* ('a bargain or special discount', Cummings and Wolf 2011: 86) ranked the highest but not represented at all in the GloWbe corpus (see Figure 6):

**jetso**  
 /dʒetsəʊ/ *n.*  
**Source language:** Cantonese (著數).  
**Definition:** a bargain or special discount.  
**Text example:**  
 “Jetso Spending Offer, Promotional Offer will be coming soon.”



Figure 6 *jetso* (DHKE)

It would seem that Hong Kong bloggers (at least on the surface of it) do not tend to use the term *jetso* to talk about bargains or special discounts. However, the term *discount* in GloWbe shows that Hong Kongers do love their discounts, comparing favourably with a number of other countries, i.e. Hong Kong (1528), Singapore (1367), Malaysia (1308), the Philippines (1110) and South Africa (762).

#### 4.4 *iron rice bowl*

DHKE codifies this term as follows (see Figure 7):

**iron rice bowl**  
*idiomatic expression.*  
**Definition:** a secure job from which it would be difficult to be dismissed.  
**Text example:**  
 “The government recognises that the ‘iron rice bowl’ is not conducive to economic modernisation and has also sought to reform the welfare system so that the burden falls on its own shoulders.”  
**Underlying conceptualization (generic):** A JOB IS A FOOD CONTAINER [TARGET DOMAIN Δ OCCUPATION] [SOURCE DOMAIN Δ FOOD CONTAINER].  
 See: golden rice bowl.



Figure 7 *iron rice bowl* (DHKE)

In Figure 7, the given example is not quite salient, as it is contained within quotation marks. Beyond this, there is another entry section entitled “underlying conceptualization (generic)”. As a reference dictionary for the lay reader, it would certainly be strange to include Lakoffian conceptual metaphors for which the reader could be confused regarding the ways in which the underlying meaning is to be decomposed. In another entry (not included here because of its low frequency), the term *mafu* (essentially a pimp) is said to be related to *mafoo* (a stable boy) and so the underlying conceptualization “sex is riding” makes one do a double take on the way in which this sentence is to be read and understood. Therefore, one must agree with Béjoint’s (2011:478) comment that such notes are expressed “in a language that is too formulaic and abstract for the users of such a dictionary, even though the information to be conveyed is relatively simple: basically, they say that ‘X is like Y’.”

Turning to GloWbe, my belief that the term is found in Singapore English is borne out (see Figure 8):

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ	JM
BOWL	40	3		9		1		1		1		10		1	13						1

Figure 8 *iron rice bowl* (GloWbe)

More interestingly, the term occurs 9 times in the UK component (see Figure 9):



1	GB G	ondeviceresearch.com	A	B	C	elsewhere for opportunities. # A recent BBC Radio 4 show " China's New <b>Iron Rice Bowl</b> " is a fascinating view into the problems a
2	GB G	guardian.co.uk	A	B	C	# First, and perhaps most fundamentally, economic reform has effectively fractured the ' <b>iron rice bowl</b> ' of cradle-to-grave provisio
3	GB G	dcsience.net	A	B	C	IMF. They are bound to get in at some stage. I think the <b>iron rice bowl</b> of the public sector is about to run out. Just wondering how
4	GB G	dcsience.net	A	B	C	the sillier bits of gimmickry introduced by new Labour. # The public sector " <b>Iron Rice Bowl</b> ", as DangerCon puts it, is hardly the lan
5	GB G	guardian.co.uk	A	B	C	and 1978. Though wages were low, the welfare system - the famous " <b>iron rice bowl</b> " - guaranteed lifetime employment, pensions
6	GB G	guardian.co.uk	A	B	C	, economic reforms, geared to creating wealth in urban areas, have smashed the <b>iron rice bowl</b> and caused severe inflation. The d
7	GB G	guardian.co.uk	A	B	C	? # " Though wages were low, the welfare system - the famous " <b>iron rice bowl</b> " - guaranteed lifetime employment, pensions, hea
8	GB G	open.edu	A	B	C	. It didn't deliver anything good to the people. " # Under the <b>Iron Rice Bowl</b> system of guaranteed employment in the state sector,
9	GB B	in2up.com	A	B	C	" Really? " I was shocked, in China, it is an " <b>Iron Rice Bowl</b> " job, secure with a steady income. " It was a

Figure 9 *iron rice bowl* (concordance, Great Britain section, GloWbe)

From this, one can glean that the term is not unknown to readers of *The Guardian* broadsheet.

#### 4.5 *yum cha*

The lexeme *yum cha* (literally 'drink tea') is codified in the DHKE as follows (see Figure 10):

**yum cha**  
 /jʌm tʃɑː/ v.  
**Source language:** Cantonese (飲茶).  
**Definition:** 1 to have teatime in a Chinese style; 2 to drink tea.  
**Text example:**  
 1 "Although **yum cha** and dim sum are used synonymously in Hong Kong. There is, properly, a distinction—when you go to a restaurant to **yum cha**, you eat dim sum or, when you **yum cha** you drink tea and eat dim sum."  
 2 "Yum cha, a term in Cantonese, literally meaning 'drinking tea,' refers to the custom of eating small servings of different foods while sipping Chinese tea ..."

www

Figure 10 *yum cha* (DHKE)

Before commenting on Figure 10, let us also take a look at the corresponding frequencies in GloWbe (see Figure 11):

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ
CHA	187	3	1	6		74	15					11	16	10	49	1				1

Figure 11 *yum cha* (GloWbe)

The combination of both Figures 10 and 11 indicates the term *yum cha* is indeed a well-loved activity in Hong Kong. Among Singaporean and Malaysian Chinese, this activity is no less loved but expressed in terms of going to the tea house and eating *dim sum* or *dian xin*. The accompaniment of hot tea is usually a given, hence the apparent distinction between *yum cha* and *dim sum* in the DHKE textual examples do not quite apply in Southeast Asia.

More surprising, though, the term *yum cha* occurs 74 times in Australia compared with 49 times in Hong Kong: it is certainly the preferred term to *dim sum* in the Antipodean country. An explanation for this is not only the high numbers of Asian migrants (from China, Hong Kong, Taiwan, Singapore and Malaysia) to Australia but the activity being more blogged about there (as a cursory concordance inspection of the 74 occurrences there shows).

#### 4.6 *wushu*

The term *wushu* is generally understood as the Mandarin equivalent of *kung fu*, traditionally its more codified counterpart (whose origin is Cantonese) in dictionaries. In Ooi (2001), I noticed that the term *kung fu* was codified in a number of English learner dictionaries then, compared with *wushu* which was not. Nowadays, both *The Merriam-Webster Dictionary* (for U.S. English) and *Dictionary.com* (a popular dictionary found also in smart phones) indicate that both terms are embraced in such native English contexts. In the DHKE, however, the term *kung fu* is relegated to the list of "Hong Kong English words now used internationally" (which, actually, is also conflated with "Chinese English" words) while the term *wushu* is found in the main text (see Figure 12):

**wushu**  
 /wuːʃuː/ *n.*  
**Source language:** Mandarin (武術).  
**Definition:** a martial art similar to Kung fu.  
**Text example:**  
 “If his **wu shu** medals at one level authenticate him, there are purists who question whether **wu shu** (designed for performance, not combat) is a ‘proper’ martial art.”



Figure 12 *wushu* (DHKE)

While this definition somewhat differs from the one found in the other two dictionaries (with “wushu” meaning Chinese martial arts collectively), the spelling as two words in the textual example is out of sync with the lexical entry which indicates its preferred spelling as one word (not two). Also, turning to GloWbe (Figure 13), we can see that the term *wushu* is by no means unique to Hong Kong English:

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ	JM
WUSHU	376	15	4	27	2	13	48	12	108	2	1	48	9	30	49	3			2	2	1

Figure 13 *wushu* (GloWbe)

GloWbe therefore indicates that the term is known in countries as far as Sri Lanka (abbreviated as LK) and New Zealand (abbreviated as NZ)

#### 4.7 kowtow

In checking out the entry for *kowtow* and its accompanying illustration in DHKE (see Figure 14), it is surprising to learn that the meaning also includes the first sense, i.e. the idea of ‘knocking [one’s] bent index and middle fingers on the table to express gratitude to the person who served the tea.’

**kowtow**  
 /kəʊtau/ *n., v.*  
**Source language:** Cantonese (叩頭) or Mandarin.  
**Definition:** 1 a finger-based method of saying “thank you” to the person serving tea; 2 a form of Chinese bowing where the forehead is knocked on the floor to (a) God or an altar; (b) a political leader (archaic); 3 a metaphor for political submission.  
**Text example:**  
 1 “After a person’s cup is filled, that person may knock his bent index and middle fingers (or some similar variety of finger tapping) on the table to express gratitude to the person who served the tea.”  
 2 “The attendance of the elders and the gentry was compulsory, while those over sixty were invited as guests of honour. The **kowtow** and the three prostrations were in the order of the government officials first, then the gentry, then the elders, then whoever happened to be there.”  
 3 “Relations between Seoul and Washington have been tense at times under Roh, a progressive elected on a wave of anti-Americanism in December 2002 after he vowed never to **kowtow** to the United States.”



Figure 14 *kowtow* (DHKE)

This practice of tapping one’s index and middle fingers on the table as an expression of polite gratitude to the person refilling one’s teacup is commonly practised among Chinese Malaysians too. The practice is said to have begun with the Qianlong Emperor who used to travel incognito and, when he poured tea for his retinue, would receive their tacit gratitude through the use of this gesture.

A detailed examination of the entire Hong Kong concordance in GloWbe fails to turn up this first textual meaning (see Figure 15):

1	HK	G	...07.tibetmagazine.net	A	B	C	But before collection, they must pray and burn aromatic plants. Some herders will <b>kowtow</b> and worship holy mountains every evening. # In Tibet, holy palaces to worship
2	HK	G	biglychee.com	A	B	C	bean spillers in Hong Kong. It's pathetic the way all the news organizations <b>kowtow</b> to this little prick. Giving him a platform to whinge -- how pathetic.
3	HK	G	eg.anhsiangchan.org	A	B	C	to hold onto "the routines of reciting how many sutras a day, to <b>kowtow</b> how many times, to eat what sorts of vegetarian diets and how many precepts
4	HK	G	english.kaiwind.com	A	B	C	of Li Hongzhi: # Li once said, "Do you know why they <b>kowtow</b> to me when they see me? They know that I am greater than any
5	HK	G	chinausfocus.com	A	B	C	in key international economic institutions. Not even a Europe in disarray was willing to <b>kowtow</b> to such a brazen demand and China should have realized this before it asked and
6	HK	G	hrichina.org	A	B	C	and especially after the CPC came to power, even though our countrymen no longer <b>kowtow</b> physically like the people of old, they kneel in their souls even more so
7	HK	G	paidalain.com	A	B	C	with arms rising and falling slowly and the intensity falling, just like the lazy <b>kowtow</b> machine in an oilfield. From time to time, I had to ask others
8	HK	G	zhangshunhong.com	A	B	C	an edict on 3 August which said that if Macartney and his suite refused to <b>kowtow</b> when they met Viceroy Liang or other mandarins, they should not be urged to
9	HK	G	zhangshunhong.com	A	B	C	as important and superior, and said that even if the envoys did perform the <b>kowtow</b> to Zheng Rui, he would not be honoured by it himself, and if
10	HK	G	adgo.com.hk	A	B	C	gigantic boulder in the path of progress. The farmers might smile and nod and <b>kowtow</b> but from long experience they knew better than to believe a word you said.
11	HK	G	sanyapark.com	A	B	C	in his hand. Disconcerted, men on the ship fell on their knees to <b>kowtow</b> . The sea deity slowly passed by the ship and steong light was shining for
12	HK	B	biglychee.com	A	B	C	's secretion was tears, given the relative shortness of his nose, depth of <b>kowtow</b> and that other substances would leave unseemly tyre tracks? # It would be interesting
13	HK	B	global-history.org	A	B	C	of Linyi also donated to the emperor an elephant, which "knows how to <b>kowtow</b> and pray." In the reign of Houliang of the Southern Dynasty, two
14	HK	B	maxdesign.com.hk	A	B	C	ah, you many bless us! We often down river, to give you <b>kowtow</b> , you why again we crashed a ship? Gold dogs, silver lion,
15	HK	B	zhongnanhaiblog.com	A	B	C	see business leaders and politicians who grew up in a sophisticated, international financial centre <b>kowtow</b> to Beijing's Communist ruler of the day -- but they must. It's
16	HK	B	biglychee.com	A	B	C	servants who joined in the smearing of CY earlier this year -- Beijing's last-minute <b>kowtow</b> to public opinion and abandonment of Henry was a humiliating slap in the face,
17	HK	B	distxiaoa.com	A	B	C	America prostate cancer number. I most worried about this thing. Is this a <b>kowtow</b> , foreplay to insufficient lack of lubrication, no ground for blame; but one

Figure 15 *kowtow* (Concordance, Hong Kong component, GloWbe)

Instead, its metaphorical usage (in the sense of “political submission”) seems to be used more in a number of countries, including the United States (Figure 16) for which *kowtow* occurs 7 times more than Hong Kong:

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ	JM
KOWTOW	476	120	11	79	10	37	49	13	18	8	6	22	24	15	17	5	7	25	6		4

Figure 16 *kowtow* (GloWbe)

With its well-known usage in Inner Circle countries, the term *kowtow* is well-codified in various native English dictionaries and so (like *typhoon*) is no longer a mere Hong Kong English feature. Following the DHKE structure, it should have been placed in the section entitled “Hong Kong English words now used internationally.” Indeed, this section is unnecessary – the words there should have been merged with the ones in the main text, with the indication that such words have now become standard English ones.

#### 4.8 qigong

The preceding remarks regarding the standardness of *kowtow* also applies to *qigong*, which is listed in DHKE as follows (see Figure 17):

**qigong**  
 /tʃiːˈɡɔŋ/ n.  
**Source language:** Mandarin (氣功).  
**Definition:** a set of breathing and movement exercises.  
**Text example:**  
 “First you indulge in a **Qigong** session (the ancient Chinese art of energy work, \$800) followed by a spot of Past Life Therapy (\$3,000), in which your therapist will take you back to a previous life and help you complete any unfinished business.”



Figure 17 *qigong* (DHKE)

The standardness of this term is well-attested in GloWbe, with an almost equal number of entries (to the Hong Kong ones) in the North American and British contexts (see Figure 18):

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ	JM
QIGONG	686	135	105	97	8	18	38	4	3		14	23	108	4	114	2	3			7	3

Figure 18 *qigong* (GloWbe)

#### 4.9 laisee/ red packet

This Cantonese term for *red packet* (a cash gift placed in a red or pink envelope, during auspicious occasions such as Chinese New Year) is particularly Hong Kong English, whose main substrate influence is Cantonese. One can predict its occurrence in other Asian contexts which are heavily influenced by Cantonese, e.g. the Cantonese-speaking cities of Ipoh and Kuala Lumpur. In Singapore, which is influenced more by Mandarin and Hokkien Chinese,

the preferred terms are *hong bao* and *ang pow* respectively. Both Figures 19 and 20 show the entries for *lai see* and *red packet* respectively in DHKE:

**lai see**  
 /lai si:/ n.  
**Source language:** Cantonese (利是).  
**Definition:** 1 money given as a gift in a red envelope; 2 a gift of money intended as a bribe.  
**Text example:**  
 1 “Children and young adults will find more money in their **lai see** packets this Lunar New Year, according to a survey.”  
 2 “Staff must not solicit ‘Lai See’ from any external business associates in any circumstances.”  
**Underlying conceptualization:** A BRIBE IS A GIFT [TARGET DOMAIN Δ CORRUPTION] [SOURCE DOMAIN Δ GIFT].  
**See:** lucky money, red packet.

www

Figure 19 *laissee* (DHKE)

**red packet/pocket (money)**  
 fixed expression.  
**Definition:** see *lai see*.  
**Text example:**  
 1 “Students are given some **red packets** (only one **red packet** contains a 10-dollar bank note).”  
 2 “Anybody is [sic] still single loves the Chinese New Year because they are entitled to receive the **red pocket money** from the married people.”

www

Figure 20 *red packet* (DHKE)

While *red packet* lists *lai see* as its definition, its textual examples are wholly positive. However, *lai see* is also seen to have a negative nuance by being associated with its underlying conceptualization of bribery and corruption (“A bribe is a gift”).

Turning to the GloWbe corpus, the distribution of *lai see* is as follows (see Figure 21):

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG
LAI SEE	56		2				1					11	6		35		1

Figure 21 *lai see* (GloWbe)

Without reproducing the attendant concordances here, the negative sense is indeed found in the Hong Kong context (e.g. “in corruption and ‘lai see’ taking”); in Singapore and Malaysia, though, there is no evidence of the bribe/corruption sense but instead there is the association with a token sum of money (e.g. “He did not take any salary, not even lai see”) or “lucky money” (“It is customary for married couples to give lai see”).

#### 4.10 Legco

For me, this term – referring to the Legislative Council of Hong Kong -- is quintessentially Hong Kong English. Evidence that this term has not found its way to standard English is that it is not found in either *Dictionary.com* or *The Merriam-Webster Dictionary* (although one can encounter it in *freedictionary.com*). Strangely, though, the term is not found in DHKE at all, despite occurring overwhelmingly 602 times for Hong Kong English in the GloWbe corpus (see Figure 22):

CONTEXT	ALL	US	CA	GB	IE	AU	NZ	IN	LK	PK	BD	SG	MY	PH	HK	ZA	NG	GH	KE	TZ
LEGCO	677	2		4		1		37				4	1		602			3	20	3

Figure 22 *Legco* (GloWbe)

Without again reproducing the attendant concordances from the various contexts, one can see that the media in various countries which use the term *Legco* mainly report the occurrences in Hong Kong (e.g. for India, there is the report on what “the President of the Legislative Council (Legco), Mr Jasper Tsang Yok-Sing, said..”) Of course, there are local uses in other countries such as Ghana, e.g. in the sentence “By 1960, LEGCO had an African majority.” Focusing on blogs and web pages in the GloWbe corpus would mean that, in addition to local news and events, global matters are also written about.

## 5 Conclusion

In this paper, we have shown that triangulation of lexicographic evidence exercised for a project such as the DHKE can be more rigorous in the application of both emic and etic perspectives – and the recourse to a curated dataset, i.e. the GloWbe corpus. However, in its defence, the GloWbe corpus was not available in 2010 when the dictionary was being compiled. And, in itself, the GloWbe corpus does have its limitations (see *English World-Wide* 36). Davies and Fuchs (2015:26) acknowledge that the corpus does not include spoken material, and that “there may be speakers from other countries who may have posted to [the website of a particular country].” GloWbe is best used as one of the relevant tools for the discovery, comparison and codification of the various Inner and Outer circle varieties that it purports to represent. Finally, the DHKE should be congratulated as a fine – though first – systematic effort at dictionary-making that showcases Hong English to the rest of the world.

## References

- Béjoint, Henri. 2011. Review of Patrick J Cummings and Hans-Georg Wolf *A dictionary of Hong Kong English: words from the fragrant harbour*. *International Journal of Lexicography*, 476-480.
- Bolton, Kingsley. 2003. *Chinese Englishes: a sociolinguistic history*. Cambridge: Cambridge University Press.
- Cummings, Patrick, and Hans-Georg Wolf. 2011. *A dictionary of Hong Kong English: words from the fragrant harbor*. Hong Kong: Hong Kong University Press.
- 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). *English World-Wide* 36:1, 1-29.
- Greenbaum, Sidney, and Gerald Nelson. 1996. The International Corpus of English project. *World Englishes* 15:1, 3-15.
- Kilgarriff, Adam and Gregory Grefenstette. 2003. Web as corpus.  
<http://www.kilgarriff.co.uk/Publications/2003-KilgGrefenstette-WACIntro.pdf>  
 Accessed 20 May 2015
- Meyer, Charles and Gerald Nelson. 2006. Data collection. In *Handbook of English Linguistics*, eds. Bas Aarts, 93-116. Malden, MA: Blackwell.
- Salazar, Danica. 2014. Towards improved coverage of Southeast Asian Englishes in the *Oxford English Dictionary*. *Lexicography: Journal of Asialex* 1, 95-108.
- Ooi, Vincent. 2001. Globalising Singapore-Malaysian English in an inclusive learner’s dictionary. In *Who’s centric now?: The present state of post-colonial Englishes*, ed. Bruce Moore, 95-121. Oxford: Oxford University Press.