English Language Teaching: Teaching of Hedges

Charles Ko[*]
Music Management Marketing (Hong Kong), Hong Kong

Abstract

A hedge is a mitigating word or sound used to lessen the impact of an utterance. It can be an adjective, for example, 'Small potato me is not as strong as you'; or an adverb: 'I maybe can swim faster than you', while it can also consist of clauses, that it could be regarded as a form of euphemism which should be taught as a main topic in English class of schools around the world. For instance, in Hong Kong schools, based on my observation while teaching in a number of primary and secondary English courses as a tutor, students report that their school teachers usually emphasize the teaching of all cohesive devices in terms of skills of writing while they neglect to explain the importance of the use of hedges in order to show euphemism. In this study, I would adopt Corpus Linguistics, a division of applied linguistics, as methodology to discover a great deal of hedges employed by so-called native speakers of English, for promoting the idiomatic usage of hedges in writing, nevertheless in speaking, so as to help teachers gain resources and inspiration in teaching to students the appropriate English hedges as a consequence of the author's hard effort while revealing from the selected corpora of this paper.

Keywords: *Hedge, English language teaching, corpus*

[*] Charles Ko, Music Management Marketing (Hong Kong), Hong Kong
E-mail: *Charlesko3@gapps.cityu.edu.hk*

Introduction

Background

First of all, some background of corpus linguistics is introduced here. There was once a landmark in modern corpus linguistics: Kucera & Francis (1967), based on the analysis of corpus, subjected it to a variety of computational analyses, from which they compiled the rich and variegated opus, combining elements of linguistics, language teaching, psychology, statistics, and sociology. A further key publication was the "Towards a description of English Usage" (Quirk, 1960:40-61) in which he introduced The Survey of English Usage[1].

Brown corpus

The Brown Corpus is a corpus that is a carefully compiled selection of current American English, totaling about a million words drawn from a wide variety of sources, that the corpus has generated a number of similarly structured corpora: the LOB Corpus (1960s British English), the Kolhapur (Indian English), the Wellington (New Zealand English), the Australian Corpus of English (Australian English), the Frown Corpus (early 1990s American English), and the FLOB Corpus (1990s British English), etc. Besides, the first computerized corpus of transcribed spoken language was constructed in 1971 by the Montreal French Project (Sankoff & Sankoff, 1973).

Survey corpus

The original **Survey Corpus** was recorded on reel-to-reel tapes, transcribed on paper, filed in filing cabinets, and indexed on paper cards. Transcriptions were annotated with a detailed prosodic and paralinguistic annotation developed by Crystal & Quirk (1964).

Corpus linguistics for study of language

It is once mentioned the follows:

> When someone is referred to as a "corpus linguist," it is tempting to think of this individual as **studying language** within a particular linguistic paradigm, corpus linguistics, on par with other paradigms within linguistics, such as sociolinguistics or psycholinguistics. (Meyer, 2004) Whether scholars attribute the status of theory to [corpus linguistics] often somewhat coincides with where they are on the continuum of *corpus-driven* and *corpus-based* linguistics. (Gries, 2010)

The answer of the question of why corpus linguistics sometimes is regarded as a paradigm for study of language can be solved by an inspiration from Gries (2010), as shown as follows:

> Whether scholars attribute the status of theory to [corpus linguistics] often somewhat coincides with where they are on the continuum of *corpus-driven* and *corpus-based* linguistics. (Gries, 2010)

Corpus-driven linguistics

Corpus-driven linguists aim to build theory from scratch, completely free from **pre-corpus** theoretical premises, base theories exclusively on *corpus data* (consider the Brown Corpus and the Survey Corpus mentioned in §1.0.1 and §1.0.2 respectively) and often reject corpus annotation (as a pre-corpus theoretical commitment) (ibid.): corpus-driven linguistics, in essence, means 'bottom-up'; consider Teubert (2005:4):

> While corpus linguistics may make use of the categories of traditional linguistics, it *does not take them for granted*. It is the discourse itself, and not a language external taxonomy of linguistic entities, which will have to provide the categories and classifications that are needed to answer a given research question. (Teubert, 2005:4; my emphasis.)

Corpus-based linguistics

Corpus-based linguists approach corpus data with **moderate** *corpus-external* premises, with the aim of testing and improving such theories, and *often* use corpus annotation (Gries, 2010).

Transition from study to method

Leech (1992) mentions that if the types of linguistic analyses that corpus linguists conduct are

---

[1] The Survey of English Usage was the first research centre in Europe to **carry out research with corpora**. The Survey is based in the Department of English Language and Literature at University College London. In parallel with the parsing of natural language data, the Survey team has carried out research and development of software tools to help linguists use these corpora.

examined, it becomes quite evident that corpus linguistics is more a way of *doing* linguistics than a separate paradigm within linguistics: in other words, it will become a methodological basis for pursuing linguistic research. In §3.1, it is provided an example introducing how corpus linguistics generates the research methodology.

Corpus Linguistics as linguistic studies

In applied linguistics, there are fields of studies inevitably relating to corpus linguistics as a methodology in the research, which the fields include translation studies, and language teaching:

> […] corpus linguistics is a whole system of methods and principles of how to apply **corpora** in language studies and teaching/learning, it certainly has a theoretical status. Yet theoretical status is not theory in itself… (McEnery et al., 2006:7f.)

What is corpus? (*Corpora – an intro for the ELT student or applied linguist at Essex*, retrieved from http://privatewww.essex.ac.uk/~scholp/corpintro.htm#def)
A corpus is a collection of language material, made in some principled way (not haphazardly), either on tape or written in hard copy (e.g. books, student essays) or in electronic form. Such collections are used in many different ways by different people.

To perform any electronic corpus-based task directly we need two things - a **corpus** and a **search engine**.

What are examples of corpora?
There are kinds of corpora including the Corpus of Contemporary American English (COCA) which is a 450 million word corpus of American English, 1990-2012. There are also the British National Corpus (BNC) and American National Corpus (ANC). BNC is the corpus including a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of current British English, both spoken and written (University of Oxford, 2011); while ANC, a project creating a massive electronic collection of American English, is the corpus including texts of all genres and transcripts of spoken data produced from 1990 onward (American National Corpus Project, 2002-2009). In 2013, there is a new corpus released in April called Global Web-Based English (GloWbE). GloWbE is a 1.9 billion word corpus of English dialects of 20 countries, 2012-2013 (see more on http://corpus2.byu.edu/glowbe/.)

Corpus tools and resources
There are the AntConc (http://www.antlab.sci.waseda.ac.jp/software.html), and the WordSmith (http://www.lexically.net/wordsmith/) mainly used by linguists for corpus studies and conducting applied linguistic research.
In addition to the regular corpus interface, there are a wide range of other corpus-based resources available on the website created by Mark Davies, Professor of Linguistics at Brigham Young University (http://corpus.byu.edu/), some of which allow us to download large amounts of data for offline use.

Corpus Linguistics as method in Applied Linguistics

Nowadays, there are a number of applied linguistic activists attempting to trace a path from data to theory by developing research methods generated from corpus linguistics (e.g. Wallis & Nelson (2001) first introduced the *3A perspective*[2] for shaping the corpus linguistic methodology). And according to Barlow (2002), three realms in which corpus linguistics can be applied to teaching are mentioned, namely syllabus design, materials development, and classroom activities (see §3.1).
3.0. Advantages of corpus linguistics as method

---

[2]**Annotation** consists of the application of a scheme to texts. Annotations may include structural markup, part-of-speech tagging, parsing, and numerous other representations. **Abstraction** consists of the translation (mapping) of terms in the scheme to terms in a theoretically motivated model or dataset. Abstraction typically includes linguist-directed search but may include e.g., rule-learning for parsers. **Analysis** consists of statistically probing, manipulating and generalizing from the dataset. Analysis might include statistical evaluations, optimization of rule-bases or knowledge discovery methods.

Sinclair (1998) pointed out that corpus linguistics provides a more objective view of language than that of introspection, intuition and anecdotes because speakers do not have access to the subliminal patterns which run through a language.

Corpus linguistics in teaching

To peek in the collocate "break" with the word *habit*, Ko (2013) making use of the Keyword in context (KWIC) concluded that the KWIC displays at least the following six ways of collocation based on COCA corpus:

Table 1. 6 ways of collocation

1	2	3	4	5	6
(to) break this (…) *habit*	a (…) *habit* to break	(to) break the (…) *habit*	(to) break a (…) *habit*	(to) break his (…) *habit*	the (…) *habit* of (…) to break
(*to +*) *break +* demonstrative pronoun + (… +) *habit*	*a + (… +) habit + to + break*	(*to +*) *break + the + (… +) habit*	(*to +*) *break + a + (… +) habit*	(*to +*) *break +* possessive determiner + (… +) *habit*	*the + (… +) habit + of + (… +) to + break*

(Ko, 2013)

Barlow (2002) states that frequency and register information in corpus could be quite helpful in course planning choices: by conducting an analysis of a corpus which is relevant to the purpose of a particular class, the teacher can determine what language items are linked to the target register. Moreover, a materials developer could create exercises based on real examples which provide students with an opportunity to discover features of language use with the help of a corpus. Nevertheless, corpus can be a useful tool for teachers to teach the collocation by a concordancing programme (ibid.), for example language teachers can make use of a Chinese corpus to teach the collocates with a specific phrase, see Figure 1.
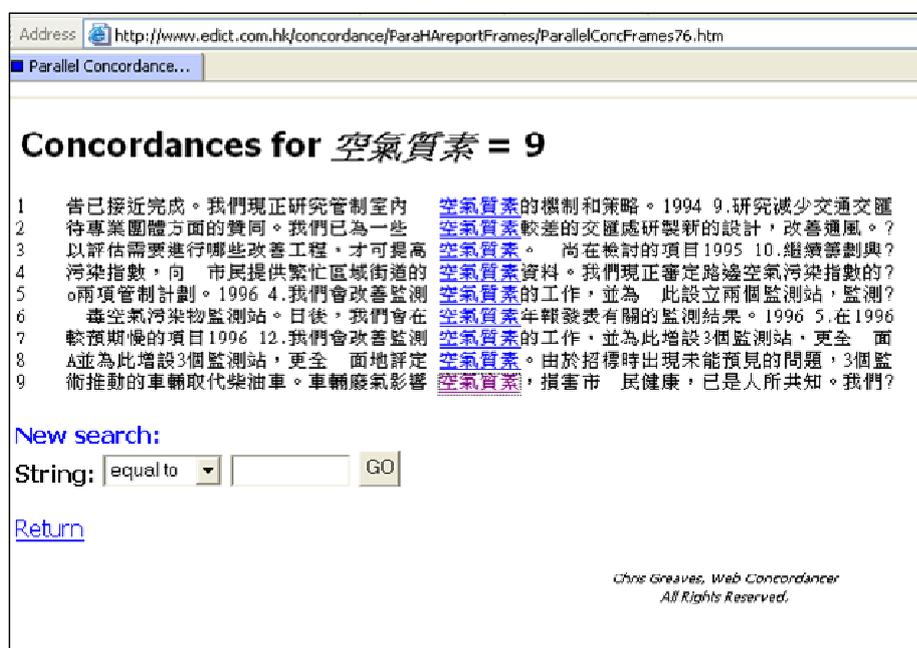

Figure 1. Chinese language teaching with a corpus

In the §4.0 and §4.1, I will adopt Corpus Linguistics, a division of applied linguistics, as methodology to discover a great deal of hedges employed by so-called native speakers of English, for promoting the

idiomatic usage of hedges in writing, nevertheless in speaking, so as to help teachers gain resources and inspiration in teaching to students the appropriate English hedges as a consequence of the author's hard effort while revealing from the selected corpora of this paper.

Methodology, Data analysis, Results and discussion

With reference to the data from the COCA and secondarily GloWbE, Global Web-Based English a 1.9 billion word corpus of English dialects of 20 countries, 2012-2013 (to compare the Global Web-Based English (GloWbE) to other corpora in §1), I will, first of all, in the §4.0, attempt to find out how actual texts are employed by native speakers with clause hedges and hedge clauses; afterwards, there would be more specific analysis in §4.1.

Clause hedges, and hedge clauses (*from* http://m.investorwords.com/2294/hedge_clause.html)
A hedge is a mitigating word or sound used to lessen the impact of an utterance. Firstly, I need to clearly compare and contrast the phrasal terms *clause hedge* (or *clausal hedge*) and *hedge clause*: while *clause hedge*s and *clausal hedges* are obviously defined as hedges consisting of clauses, that it could be regarded as a form of euphemism which should be taught as a main topic in English class of schools around the world, it may be surprising that the similar term *hedge clauses* refer to statements made to free oneself from responsibility, simply speaking, *disclaimers* in daily terminology which are usually adopted in business context especially the one in investment world.

Clause hedge – I am not an expert
Based on the COCA data (see figure 2 below), the frequency of the clause hedge *I am not an expert* is 20. To investigate one of its usage in COCA, figure 2 extracted from the corpus shows that even a scholar should use a hedge to minimize the impact of his or her statement.

Clause Hedge	FREQ
I AM NOT AN EXPERT	20

Figure 2. Frequency of *I am not an expert*

Source	information:
Date	2010
Publication information	Jul2010, Vol. 18 Issue 4, p18-21, 4p
Title	Making clinical decisions in emergency situations.
Author	Parsonage, Ian
Source	Emergency Nurse

Expanded context:

knowledge gained from my past experiences, therefore, I considered that James's condition was not life threatening, and that I had time to take a comprehensive history and set of baseline observations to back up my initial diagnosis of severe allergic reaction. # Benner (1984) explains how an intuitive approach distinguishes experts from novices, with experts no longer relying on analytic principles to connect their understanding of clinical situations to the appropriate action. # I am an adult trained nurse with no recognised academic paediatric training, so it could be argued that **I am not an expert** in this case and would, therefore, struggle to use my intuition when making decisions. # One of the key findings of a recent report by the Royal College of Paediatrics and Child Health (RCPCH) (2008) is that the failure of healthcare professionals to recognise serious illness in children was a cause of preventable death. # One of the report's recommendations is that all healthcare professionals who treat sick children should have appropriate training and supervision to develop their key skills and competencies, maintain standards and

Figure 3. Expanded context with source information

More specific analysis
Disclaimer of warranties
Secondly, I will study another linguistic phenomenon, researching on the authentic usages of disclaimers by businessmen or non-businessmen. Nowadays, based on my observation while teaching in a number of primary and secondary English courses as a tutor, students report that their school teachers usually emphasize the teaching of all cohesive devices in terms of skills of writing while they neglect to explain the importance of the use of hedges in order to show euphemism. In my opinion, euphemism is not only for indirectness that may be for being humble, it can nevertheless be used for saving one selves, which should not only be realized by hedges, but also the disclaimers (of warranties). In this subsection, I attempt to peek into the so-called native American stylistic disclaimers, studying how American make a disclaimer by finding out in corpus of COCA. For this purpose, I type the word "disclaimer" in the box of WORD(S), then the frequency appears to be 470 that I only select the following seven of them to study, in Figure 4.

1	love Steven Tyler, OK. So I'm going to -- there's a **disclaimer** right off the bat, all right. I like him, he's a
2	Exemption, which gives some seniors a discount on their house taxes. First a **disclaimer**: My wife and I qualify for this exemption. We received it once.
3	complications and can help surgeons discuss these occurrences more objectively with their patients. # **DISCLAIMER**: No funding was received for the preparation of this article. # TABLE 1
4	are? Prof-REDNISS: Yeah. And those, they require researchers to sign a **disclaimer**. In the Curie Institute in Paris, you can hold up a Geiger counter
5	now? GEORGE-WILL-1-ABC# (Off-camera) He's getting at - and here I give my weekly **disclaimer** that my wife is an adviser to Mr Perry. He's getting at the
6	3831 AUTHOR'S NOTEThe customary **disclaimer** is correct: The characters and events herein are fictitious. Any similarities to anyone
7	liner to protect your underwear from any wetness. # One (very important) **disclaimer**: Ditch the condoms only if you are in a monogamous relationship and know for

Figure 4. "Disclaimers"

The result shows that they usually (1) make a disclaimer first before proposing their second utterance, consider item 2 in figure 4 above; (2) make a disclaimer before making a promise or what they will do, consider item 1. To conclude, disclaimer can be regarded as a must if someone is not sure whether he or she can have such an ability to continue his or her commitment, especially consider items 3, 6, and 7 that used with a colon to explicitly say that there is a disclaimer.

Adjective hedges - Small potato & poor me
Moreover, I type the words *small potato* on COCA for the investigation of the use of adjective hedge. While the frequency is 13, I select the first five in figure 5 for investigation, shows that the phrase *small potato* can describe a small-scale business, in first item. In other words, the words *small potato* should not only modify human, but probably and mostly selected for hedging some belongings of oneselves.

1	(Photo-of-J.R. -Simp OSGOOD: (Voiceover) J.R. Simplot turned his **small potato** business into billions. He gave us the frozen french fry. (Photos-of-Robert-M) OSGOOD:
2	Laden, I can understand. But why Manson? He's a very **small potato**. " # " That one's for a selfish reason. " I admitted
3	the life cycles of scale insects, the animal looks a bit like a **small potato** bug. The typical crawler, after hatching from its egg, walks a short
4	beats a minute and each of his lungs shrink to the size of a **small potato**. He is pushing his body to the limit. Free divers say there's
5	can't make latkes without potatoes. " She opened a cupboard. One **small potato** lay at the bottom of a big bag. # " You didn't buy

Figure 5. Small potatoes

Nevertheless, the hedge items involving the hedge *poor me* are investigated as Figure 6.

1	no! ye have not a whole body, Sir; do ye but use **poor me** for your one lost leg; only tread upon **me**, Sir; I
2	to living a happy life, and it helps keep you from' Oh, **poor me**, I don't have a date Saturday night!' " # In
3	who look as if they've borrowed their fathers' waistcoats? -- instead of **poor me**. You're the one who fears regrets, not I! " "
4	do have to find a way to laugh through pain. The reason those " **poor me** " jokes work is that they have a grain of truth. # SHAPE
5	said he'll stop in on your nine-thirty. " Henry rolls his eyes. **Poor me**. Poor us. Meredith looks away, turns the page. The ironic

Figure 6. Poor me

It is concluded that, the hedge *poor me* can be used as an interjection to mean I am not at a high position in the specified conversation, consider item 2.

Adverb hedges - Maybe…
Furthermore, I find that *maybe* is quite a frequent hedge among American that when comparing to others, such as British people and Hongkongers, American use the hedge much more than the usage by others: the results are based on the ones in GloWbE, see figure 7 below.

Adverb Hedge	ALL	US	CA	GB	SG	MY	PH	HK
MAYBE	413758	124128	31517	94393	10581	8907	9399	5746

Index – US For American, GB For British people, HK For Hongkongers 0.672 seconds

Figure 7. Maybe

Conclusion

The paper promotes that Corpus Linguistics (CL), a division of applied linguistics, should be a methodology to teach in class. In this study, the author attempts to discover a great deal of hedges employed by so-called native speakers of English, for promoting the idiomatic usage of hedges in writing, nevertheless in speaking.

CL as a science
For instance, a research paper states that translators sometimes adopt the non-collocate word *poor* with the word *habit* in the translation of 陋習 ("poor habit") (Shei, 2001). Shei (2001) mentions there will be a serious problem in translation if translators only translate the source text without paying attention to the aspect of collocates, consider:

   Original text: 要革除一項陋習,常常需要下最大的決心.

   Target text 1: To **remove** a poor habit, you often need to make the biggest determination.
   (Literal translation)
   Target text 2: To **break** a bad habit, you often need to show the greatest determination.
   (Translation concerning the collocates) (Source:
   http://geocities.ws/ccshei/publication/collocation_translation.htm)

To improve the clause "To remove a poor habit", we first need to find out the **verb** collocate with the word *habit*: is the word *remove* appropriately employed in the translation upon the concern of collocation? To answer this question, Ko (2014) consulted the COCA, and the answer is there is probably a problem in the translation as there are few people using the language in this way, i.e. only 2 in frequency.

Summary
In the current trends in corpus linguistics (e.g. Kruger, Wallmach & Munday, 2011), it seems that

English Language Teaching: Teaching of Hedges

linguists have already taken it for granted that the corpus linguistics should be an essential method in any applied linguistic research. So as to help teachers gain resources and inspiration in teaching to students the appropriate English hedges, the author reveal different kinds of hedges from corpora as a consequence of his hard effort.

References

American National Corpus Project. (2002-2009). *American National Corpus*. Retrieved May 17, 2013, from http://www.anc.org/index.html

Barlow, M. (2002). Corpora, concordancing, and language teaching. Proceedings of the 2002 KAMALL International Conference. Daejon, Korea.

Bowker, L., & Pearson, J. (2002). *Working with specialized language: a practical guide to using corpora*. London: Routledge.

Crystal, D., & Quirk, R. (1964). *Systems of prosodic and paralinguistic features in English*. The Hague: Mouton.

Gries, St.Th. (2010). "Corpus linguistics and theoretical linguistics: a love–hate relationship? Not necessarily…*". *International Journal of Corpus Linguistics*, 15: 3.

Ko, C. "Analysis of the collocates with the word *habit* with the COCA corpus". *International Journal Of Pure & Applied Research In Engineering & Technology*. Vol. 2, Issue 4/ November 2013.

Ko, C. K. S. (2014). "Analysis of the collocates with the word habit with the COCA corpus". Second Asia Pacific Corpus Linguistics Conference (APCLC 2014). 7-9 March 2014. The Hong Kong Polytechnic University, Hunghom, Kowloon, Hong Kong.

Kruger, A., Wallmach, K., & Munday, J. (ed.) (2011). *Corpus-based translation studies: research and applications* (1st edition). London: Continuum International Pub. Group.

Kucera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence: Brown University Press.

Leech, G.N. (1992). "Corpora and theories of linguistic performance". In Jan Svartvik (Ed.), Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4–8 August. Berlin/New York: Mouton de Gruyter, p.106.

McEnery, T., & Wilson, A. (1996). *Corpus linguistics*. Edinburgh: Edinburgh University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: an advanced resource book*. London/New York: Routledge, p.7f.

Meyer, C.F. (2002). *English corpus linguistics: an introduction*. Cambridge: Cambridge University Press.

Meyer, C.F. (2004). *English corpus linguistics: An introduction.* Cambridge: Cambridge University Press.

Quirk, R. (1960). "Towards a description of English usage", *Transactions of the Philological Society*, pp. 40–61.

Sankoff, D., & Sankoff, G. (1973). "Sample survey methods and computer-assisted analysis in the study of grammatical variation". In Darnell R. (ed.) *Canadian languages in their social context.* Edmonton: Linguistic Research Incorporated, 7–64.

Shei, C.-C. (2001). *The problem of lexical collocation in Chinese-to-English translation.* 長榮大學學報 5 (2)，頁135-149。

Sinclair, J. (1998). "Corpus evidence in language description". In Gerry Knowles, Tony Mcenery, Stephen Fligelstone, Anne Wichman, (Eds.) *Teaching and language corpora.* London: Longman, pp. 27-39.

Stubbs, M. (1993). "British traditions in text analysis: From Firth to Sinclair". In M. Baker, F. Francis & E. Tognini-Bonelli (Eds.) *Text and technology: In honour of John Sinclair.* Amsterdam/Philadelphia: John Benjamins, p.2f.

Teubert, W. (2005). "My version of corpus linguistics". *International Journal of Corpus Linguistics*, 10(1): 4.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work.* Amsterdam/Philadelphia: John Benjamins, p.1.

University of Essex. (January, 2005). *Corpora – an intro for the ELT student or applied linguist at Essex.* University of Essex. Retrieved May 29, 2013, from http://privatewww.essex.ac.uk/~scholp/corpintro.htm#def

University of Oxford. (2011). *About the BNC.* Retrieved May 17, 2013, from http://www.natcorp.ox.ac.uk/

Wallis, S., & Nelson G. (2001). "Knowledge discovery in grammatically analysed corpora". *Data Mining and Knowledge Discovery*, 5: 307–340.