

---

---

## Plenary

# Researching and Teaching ESP: Insights from Hong Kong

Winnie Cheng

*The Hong Kong Polytechnic University*

The Research Centre for Professional Communication in English (RCPCE), Department of English, The Hong Kong Polytechnic University was established with an important goal of promoting ESP teaching and research through collaboration across different disciplines, businesses and professions. This paper describes two corpus studies, namely profession- and genre-specific phraseologies and the speech act of <Opine> markers, and discusses some main findings to underscore the value and importance of collaboration between ESP researchers and professionals.

In Hong Kong, the Department of English of The Hong Kong Polytechnic University (PolyU), driven by the mission of “Linguistic Excellence in Professional Contexts”, has been conducting research and scholarship with a focus on applied linguistics, sociolinguistics, and professional communication. In 2005, the department established the Research Centre for Professional Communication in English (RCPCE) with two goals:

1. To investigate the purposes, nature and patterns of communication through the medium of English in business and professional contexts.
2. To promote ESP teaching and research through collaboration across different disciplines, businesses and professions.

To attain these goals, the RCPCE has been involved in many collaborative and interdisciplinary research projects with the pillar industries in the Hong Kong economy, resulting in the compilation of a group of profession-specific corpora. RCPCE has also adopted a range of research methods in language and communication, including corpus methods.

This paper describes research findings of a couple of ESP research studies that

have analysed RCPCE profession-specific corpora, accessible on <http://rcpce.engl.polyu.edu.hk/index.html>, namely a study of the phraseological variation of discipline- and profession-specific English language and a study of the speech act of <Opine> marker in Hong Kong English textbooks. The paper concludes with ESP pedagogical implications.

## **Corpus approach to ESP research**

A corpus is defined as “a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research” (Sinclair, 2005, p. 16). The RCPCE has, over the years, created a number of profession-specific corpora, as follows:

- Hong Kong Corpus of Spoken English (HKCSE) (0.9 million words)
- Hong Kong Corpus of Surveying and Construction Engineering (HKCSCE) (0.57 million words)
- Hong Kong Engineering Corpus (HKEC) (0.92 million words)
- Hong Kong Financial Services Corpus (HKFSC) (0.73 million words)
- Hong Kong Corpus of Corporate Governance Reports (HKCCGR) (1 million words)
- Hong Kong Corpus of Corruption Prevention (HKCCP) (0.55 million words)

To date, ESP as well as language and communication studies have examined the RCPCE profession-specific corpora in terms of discourse markers, cohesive devices, vocabulary, grammars, lexico-grammar, metaphors, multi-word units, (critical) discourse analysis, pragmatics, (critical) genre analysis, and multimodality. Specifically, corpus linguistic research has investigated different kinds of recurrent patterns of the English language, including the behaviour of words and multi-word phrases; grammatical patterns; semantic features; discursal, generic, interactional, and pragmatic features; and textual properties.

## Discipline- and profession-specific English language: phraseological variation

The study compared two kinds of phraseologies, namely n-grams and congrams, in a corpus of engineering ordinances (139,176 words), extracted from the Hong Kong Engineering Corpus (HKEC). The aim of this study was to examine and compare trigrams and 3-word congrams to show the extent to which they are helpful in revealing the aboutness of the specialised professional corpus examined. According to Phillips (1989), the phraseology of the language contained in a text or a corpus can reveal the ‘aboutness’ of the discipline or profession.

The notion of “congram”, proposed by Cheng, Warren, & Greaves (2006) is defined as sets of words that co-occur regardless of constituency variation (e.g., AB and A\*B), positional variation (e.g., AB and BA), or both.

Table 1 shows the most frequent ten trigrams and ten 3-word congrams in the Engineering Ordinances Corpus, with frequencies of occurrence (Table 1).

Table 1  
*Most Frequent Trigrams and 3-Word Congrams in the Engineering Ordinances Corpus*

Trigrams (Frequencies)	3-Word Congrams (Frequencies)
the water authority (186)	of/the/to (2,452)
no person shall (139)	and/of/the (2,263)
the contractor shall (127)	a/of/the (2,186)
in respect of (113)	in/of/the (2,080)
in relation to (108)	of/shall/the (2,070)
project division emsd (108)	be/shall/the (1,815)
referred to in (107)	be/of/the (1,703)
in the case (100)	of/or/the (1,695)
liquefied petroleum gas (92)	for/of/the (1,311)
in accordance with (91)	a/in/the (1,152)

Table 1 shows that some of the trigrams, comprised of lexical words, or lexical and grammatical words, can rather readily be associated with the discipline of engineering, and thus indicating ‘aboutness’, such as ‘the water authority’, ‘the contractor shall’, ‘project division emsd’ (Electrical and Mechanical Services Department), and ‘liquefied petroleum gas’. Other trigrams are neither readily associated with engineering nor the genre of ordinances, despite ‘shall’ in ‘no person shall’ and ‘the contractor shall’. These and the remaining trigrams need to be studied in the respective concordances to identify general meanings or specific meanings, be they engineering-specific, ordinances-specific, or engineering ordinances-specific.

The group of three-word concgrams is equally, if not more, interesting and useful for understanding the genre of engineering ordinances. All are made of grammatical words. In the following, the concordance of the most frequent three-word concgram ‘of/the/to’ ( $N = 2,452$ ) was examined. Three kinds of variations have been identified, and they are:

1. **the** current status **of** each module **to** be monitored  
**the** supply **of** gas **to** any premises  
**the** prevention **of** annoyance **to** persons travelling
2. **The** Ordinance applies **to** certain types **of** projects
3. **to** prevent **the** further escape **of** gas

This small-scale study confirms the value of both n-grams (in this case, trigrams) and concgrams (in this case, 3-word concgrams) in the study of the aboutness of a discipline, a profession, and even a genre. The findings related to trigrams show that the n-gram list could, to various degrees, reveal the “aboutness” of a corpus, especially a specialized corpus such as the English Engineering Ordinances Corpus. Examination of the concordances can further reveal the contextual meanings associated with the n-grams under study. However, although concgrams, made up of grammatical words, appear to be rather meaningless, concordance analyses of concgrams can much better and much more clearly reveal the structural components of the complex nominalisations or clauses constructed by the concgram.

---

---

## Hong Kong English textbooks: The <Opine> marker

Since 2005, a series of research studies has compared the use of different speech acts in a self-compiled specialized Hong Kong English language textbook data set and the business sub-corpus (0.23 million words) of the Hong Kong Corpus of Spoken English (HKCSE). The textbook corpus consisted of fifteen textbooks endorsed by the Education and Manpower Bureau of the Hong Kong Government (equivalent to a Ministry or Department of Education) for use in Hong Kong's upper secondary schools. The speech acts examined are namely correcting others and self-correction, checking understanding, interruption, giving an opinion, and disagreement.

The comparative study of giving an opinion (Cheng & Warren, 2006) involved comparing the speech act of <Opine>, defined by Stenström (1994) as: “<Opine> expresses the speaker's personal opinion, his/her feelings and attitudes” (pp. 90–91), which are “generally realised by a declarative and provided with both lexical and prosodic <Opine> markers” (pp. 90–91). Examples of lexical <Opine> markers are “I feel, I think, it seems, it's a pity that, it's surprising that, it's ...” (p. 91).

Table 2 shows the frequencies of different <Opine> markers in the business sub-corpus of HKCSE.

The following shows some linguistic realisations of giving an opinion in three of the fifteen textbooks examined:

*I think ..., I believe ..., Personally, I find this ..., In fact, I ...* (Esser, 1999, p. 32)

*Personally, I think that ..., In my opinion ..., As I see it ..., In my view ..., Actually I don't agree that ..., I think it's interesting that ..., From my point of view ..., I suppose that ..., It seems to me that ...* (Duncan, 2001, p. 10)

*In my opinion ..., I think ..., I feel that ..., It seems to me that ...* (Lee & Holzer, 1999, p. 77)

Major comparative findings in the textbooks are described, as follows:

- The number of <Opine> markers across the fifteen textbooks varies from four to nine.

Table 2

*Opine Markers: Frequencies of Use in the Business Sub-Corpus of Hong Kong Corpus of Spoken English (HKCSE) (0.25 million words)*

Opine markers	Frequencies
I + think	602
It's	386
I + know	185
I + say	133
I + like	112
I + sure	55
the / one + thing + is	37
I + guess	30
I + feel	27
It / everything + seems	26
I / find	24
I / believe	22

- 82.7% of the 1,714 speaker opinions are realised by just 6 <Opine> markers (*I + think*, *It's*, *I + know*, *I + say*, *I + like* and *I + sure*).
- *I + think* (35.1%) is the most frequently occurring <Opine> marker, and only this <Opine> marker is found in all of the fifteen textbooks.
- The textbooks contain only four (*I + think*, *I + sure*, *I + feel*, and *It/ everything + seems*) of the top ten <Opine> markers in the business sub-corpus (Table 1), and so exclude some of the most frequent forms (*It's*, *I + know*, *I + say*, *I + like*, *the/one + thing + is*, and *I + guess*) (see Table 1).
- At the same time, the textbooks emphasise <Opine> markers which are rarely used by speakers in the data, such as *I + believe* (1.3%), *I + see* (0.6%), *my + view* (0.5%), *my + opinion* (0.5%) and *I suppose* (0.35%).

- None of the textbooks provides and teaches any examples of negative constructions.
- None of the textbooks gives any indication as to whether or not the <Opine> markers are in rank order based on frequency of use.

## **Conclusion and implications**

The first study reported in the paper is a corpus-driven one, meaning that the researcher acts as a “language detective”, seeking answers to questions that can be found by means of corpus queries. Their task is to identify the recurrent patterns found in the corpus output lines (concordance lines) and make generalizations. The implication of this study is that the corpus linguistic information generated electronically, which is not available in a dictionary or a thesaurus but, with the right training provided to students and professionals, is available through searching a profession-specific corpus.

The findings of the first and related studies have been disseminated at continuing professional development (CPD) seminars to members of engineering associations to share with them phraseological corpus findings from the specialised corpora they own. In addition, professional communication training programmes have been conducted for members of relevant professions.

The second study is a corpus-based one, with the <Opine> markers identified in the textbooks used as search items when examining the business sub-corpus in the Hong Kong Corpus of Spoken English (HKCSE). Main findings show that the writing of the English textbooks examined tends to rely on the introspections of the textbook writers rather than real-world language use. Based on the conclusion of the study, it is suggested that textbook writers need to pay greater attention to real world language use when exemplifying speech acts. Teachers and students should be aware that ‘textbook rules’ are not necessarily valid as generalizations. If space is limited, the most frequent <Opine> markers should feature prominently in the textbooks.

Language corpora can clearly make important and useful pedagogical contributions to English language learning and teaching. Corpora, such as the business discourse corpus in the second study, provide a better point of reference

for what is said and how than the intuitive notions of textbook writers.

## References

- Cheng, W., & Warren, M. (2006). I would say be very careful of...: <Opine> markers in an intercultural Business corpus of spoken English. In J. Bamford & M. Bondi (Eds.), *Managing interaction in professional discourse. Intercultural and interdiscursual perspectives* (pp. 46–58). Officina Edizioni.
- Duncan, J. (2001). *Expression: oral practice for AS-level*. Precise Publications.
- Esser, D. (1999). *Teach & practice: AS-level oral English for Form 7*. Pilot Publications.
- Lee, I., & Holzer, V. (1999). *Skills building for AS Use of English*. Longman
- Phillips, M. (1989). *Lexical structure of text*. English Language Research.
- Sinclair, J. (2005). Corpus and text - Basic principles. In M. Wynne (Ed.). *Developing linguistic corpora: a guide to good practice* (pp. 1–6). Oxbow Books.
- Stenström, A.-B. (1994). *An introduction to spoken interaction*. Longman.

## Author Bio

*Winnie Cheng is a retired Professor, and former Director of Research Centre for Professional Communication in English (RCPCE), of the Department of English at The Hong Kong Polytechnic University. She is a founding fellow of the Hong Kong Academy of Humanities. Her research interests include corpus linguistics, conversational analysis, critical discourse analysis, discourse intonation, EAP, ESP, intercultural communication, pragmatics, and writing across the curriculum. <winnie.cheng@polyu.edu.hk>*

**Received:** October 20, 2019

**Accepted:** September 9, 2020