

行政院國家科學委員會補助專題研究計畫成果報告

學術論文之詞彙慣用語分析

Lexical Phrases in Academic Research Papers

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC89 - 2422 - H - 009 - 001

執行期間：88 年 8 月 1 日至 90 年 1 月 31 日

計畫主持人：郭志華

共同主持人：

計畫參與人員：

執行單位：國立交通大學語言教學與研究中心

中華民國九十年五月七日

行政院國家科學委員會專題研究計畫成果報告

學術論文之詞彙慣用語分析

Lexical Phrases in Academic Research Papers

計畫編號：NSC 89-2422-H-009-001

執行期限：88年8月1日至90年1月31日

主持人：郭志華 國立交通大學語言教學與研究中心

一、 中文摘要

詞彙慣用語是指具有高頻率及成為習慣用法的字詞文法單元，並且常與某些言談情境和溝通功能相關聯。針對某一文類的詞彙慣用語所做的研究可以披露使用該文類的言談團體的一些特色及主要的言談策略。於寫作教學上，研究指出中等程度學習者常需要大量的不同情境中表達各種溝通功能的常用詞彙片語，而缺乏某一文類的詞彙慣用語常導致學習者無法符合其言談團體之期望。

科技研究論文長期以來已形成一種高度約定俗成的文類，並且在詞彙及結構上有其特殊用法。此領域之研究指出，科技研究論文具有字詞文法上的特色，這些特色並與其結構息息相關。本計畫從詞彙慣用語之觀點來探討科技研究論文，並尋找這個文類中的詞彙慣用語。

研究分兩部分。首先我們提出一個期刊論文各章節的溝通功能表，我們再分析三十六篇科技研究論文，找出各章節的詞彙慣用語，並把它們與溝通功能相連結。第二部分是字詞搭配分析。我們從十二篇科技研究論文中找出高頻率的字，再分析它們的字詞搭配，並與一般英語搭配字典這些字的字詞搭配相比較。

以研究方法而言，借助主要字詞索引軟體，我們結合了文類分析與語料庫分析。此研究方法使我們可以用更實證的方式探討科技研究論文這個文類，是一個很有潛力的研究方法。

研究結果顯示，在科技研究論文的每一章節中，相對於每一重要溝通功能，都有一些詞彙慣用語，報告中列了一些例子。字詞搭配分析則顯示科技學術英語和一般英語在字詞搭配上有一些有趣的差異，而且差異反應了科技研究論文這個文類的字詞文法特色。最後，報告中亦討論研究結果於教學上之意義與應用。

結果於教學上之意義與應用。

關鍵詞：詞彙慣用語、字詞文法單元、字詞搭配分析、主要字詞索引軟體

Abstract

Prefabricated language, or formulaic language, is lexicogrammatical units characterized by high frequency and conventionality in use and associated with certain discourse contexts and communicative functions. Investigation on the formulaic language of a genre can reveal the defining characteristics of the discourse community and the dominant discourse strategies. Pedagogically, classroom writing research has indicated that intermediate-level learners are in need of a sufficiently large repertoire of lexicogrammatical expressions associated with a wide range of discourse functions in various contexts. A lack of phraseological competence usually leads to failure to conform to the expectation of the discourse community.

Scientific research papers have been established as a highly conventionalized genre with idiosyncratic language usages. Research has found that scientific research papers have distinct lexicogrammatical features closely related to their macrostructures. This study takes a phraseological approach to scientific research articles in hope of identifying lexical phrases and collocations recurring in this genre. It consists of two parts. First, a corpus of thirty-six scientific research papers was analyzed and recurrent lexicogrammatical patterns were identified in relation to a proposed framework of communicative functions of the major sections in scientific research papers. The second part is an analysis of the collocational patterns of high-

frequency words in a corpus of twelve articles in comparison to the patterns of the same words occurring in general English. Methodologically, with the help of concordancing software, combining genre analysis with corpus analysis proves a promising attempt that can demonstrate how phraseology may characterize a genre.

Results show that in each major section of scientific research papers, a number of recurrent lexical phrases can be identified in correspondence to particular communicative functions of the section. Collocational analysis of high-frequency words from sample research papers reveals significant differences in collocational patterns between scientific-academic English and general English. The differences characterize the idiosyncratic lexico-grammatical features of this particular genre.

Finally, pedagogical implications as well as applications are discussed.

Keywords: lexical phrases, prefabricated language, formulaic language, lexico-grammatical units, collocational analysis, concordancing software

二、緣由與目的 (Introduction)

Language production and language learning behavior have long been regarded as generative, rule-governed processes. Recent studies in first and second language acquisition, however, have revealed that prefabricated language, or formulaic language, may play a role in language development. It was indicated that language learners usually pass through a stage at which they use a large number of unanalyzed chunks which are basic to creative rule-forming processes (Wong-Fillmore 1976; Huang & Hatch 1978; Vihman 1982; Pawley & Syder 1983; Peters 1983; Nattinger & DeCarrico 1992; Wray 1999). For example, Wong-Fillmore (1976), which was one of the most extensive studies on child second language acquisition, claimed that learners analyzed formulaic sequences and derived rules from them, which they subsequently produced creatively. More

recently, Myles, Hooper & Mitchell (1998), a longitudinal study, found that learners would “unpack” their early chunks and use part of them productively in the generation of new utterances.

Prefabricated language, or formulaic language, is lexico-grammatical units characterized by high frequency and conventionality associated with certain discourse contexts and functions. There has been a great diversity in the terms used by linguistic scholars to describe these lexico-grammatical units, such as phraseology (Howarth 1996; Cowie 1998), lexical phrases (Nattinger & DeCarrico 1992), formulaic language (Myles, Hooper & Mitchell 1998; Weinert 1995; Wray 1999), prefabricated patterns or prefabs (Howarth 1996; Granger 1998), word combinations (Howarth 1996; Benson 1986), collocations (Sinclair 1991; Gledhill 2000), fixed expressions (Moon 1998), and formulas (Vihman 1982). Despite limited consensus in the use of terms, interest in the analysis of what can be broadly referred to as phraseology has been growing and there is a general recognition of the crucial part it plays in language acquisition and production.

A lot of research on phraseology has focused on its role in first or second language acquisition. For instance, Wray (1999) indicated that formulaic language offers processing benefits by providing a short cut to production and comprehension. In addition, formulaic language can be associated with socio-interactive functions, such as ritual speech acts (Aijmer 1996; Altenberg 1998). Weinert (1995), reviewing research on the role of formulaic language, identified three different functions of formulaic language: communicative, production, and learning strategy.

Another focus of research is the phraseological competence of non-native speakers in comparison to that of native speakers and to what extent phraseology presents difficulties for the former (Allerton 1984; Pawley & Syder 1983; Cowie 1992; Granger 1998; Howarth 1998). For instance, the phenomenon of phraseological incompetence has been described as follows: “so often the patient language learner is told by the native speaker

that a particular sentence is perfectly good English ... but that native speakers would never use it" (Allerton 1984: 39). Granger (1998) analyzed NS and NNS corpora and compared their written performance. Results showed that L1 plays an important role in the acquisition and use of prefabricated language in L2. Howarth (1998) analyzed the phraseology of learners' academic writing, and indicated that even advanced non-native writers may fail to communicate effectively as a result of a lack of phraseological competence. Gledhill (2000) investigated the discourse functions of collocation in research article introductions. He found that collocations of high frequency words in medical research abstracts and articles are useful indicators of the prototypical phraseology of the genre.

This study, from a pedagogical perspective, analyzed and identified the recurrent, conventionalized lexical phrases as well as collocational patterns in the genre of scientific journal articles. It was believed that the phraseology of a genre should be explicitly addressed in the writing course of this genre, particularly to non-native learners.

Methodologically, as Gledhill (2000) indicated, the attraction of a combined approach to both genre and corpus analysis lies in the potential for a corpus to reveal recurrent patterns across a representative sample of texts. In this study, we took such a combined approach. We first examined a corpus of thirty-six articles from three journals in different scientific research fields. The lexico-grammatical patterns were then identified on the basis of the following criteria:

1. The pattern occurs more than once in the corpus.
2. The pattern performs a particular communicative function in the section where it occurs.
3. The pattern has syntagmatic simplicity and paradigmatic flexibility; that is, it permits wide variation of lexical content in relatively simple syntactic frames (Nattinger & De-Carrico 1992).
4. The pattern is used in a specialized sense in scientific research papers; that is, it is conventionalized and

idiosyncratic in this genre.

After the recurrent lexical phrases were identified, they were linked to a framework of communicative functions of the sections (that is, Introduction-Materials and Methods-Results-Discussions and Conclusions) we proposed.

The second stage was a collocational analysis. Collocation provides very basic contextual information about words and their usages; in addition, high frequency words often have a restricted and idiosyncratic syntax and embedded in set phrases which have rhetorical force (Sinclair 1991; Moon 1992). Therefore, we intended to analyze the collocational patterns of high frequency words in a corpus of twelve scientific journal articles in order to compare the collocational patterns between scientific-academic English and general English. Computer software LMC was used for frequency analysis and concordancing. We selected nodes from a list of the top 50 most frequent words in our corpus. The nodes were limited to subtechnical verbs and nouns. The collocates of each node, with a span of 4 words on each side, were then identified from the concordance list the computer provided.

三、結果與討論 (Results and Discussions)

Lexico-Grammatical Patterns

Results from the analysis show that in each major section of scientific research papers, a number of lexico-grammatical patterns can be identified in correspondence to particular communicative functions of the section. Following are some examples from the results (as a result of the limitation of space, only one example of the lexico-grammatical patterns is provided for each communicative function).

A. Introduction

1. General setting
...have long been used as
2. Purpose of research
The purpose of the present research was to
3. Literature review
(citation) demonstrated that

4. Indication of a gap or extension of previous knowledge

However, few researchers have reported

5. Rationale

A reason for this study is to

6. Major findings and results

The results show that

7. Main features of research

In this paper, a ... is developed

8. Value of research

would greatly improve

9. Content and organization to follow

in the following sections

B. Materials and Methods

1. Overview or design of the investigation

Experiments were undertaken to

2. Methods, approaches, models, techniques, or systems

...techniques were used to

3. Population, samples, or subjects

Samples of ... were collected from

4. Location, time, or figures and tables

This is illustrated in Fig. (number), where

5. Experimental conditions, criteria, assumptions, or hypotheses

It is assumed that

6. Procedure

Several steps were taken to

7. Materials or equipment

...be measured using

8. Variables, equations, or measurements

The equation for ...is as follows: ... where

9. Tests

Tests were performed

10. Comparison with other studies

The same ... has been reported in (citation)

C. Results

1. Overall description of the investigation

Analyses were conducted using

2. Location of figures and tables

...be presented in Fig. (number)

3. Presentation of important findings and results

As can be seen in Fig.(number)

4. Comments on the results

These results suggest that

D. Discussions and Conclusions

1. Background information

An effort has been made to

2. Statement of results

It has been shown that

3. Expected or unexpected outcome

Unfortunately, we were not able to

4. Reference to previous research

This is in good agreement with (citation)

5. Explanation

In view of the fact that

6. Deductions, generalizations, and applications

It can be concluded that

7. Recommendations

...be still needed for future...

Collocational Patterns of Use: an Example

Following the procedure of collocational analysis described in the previous section, eleven words were selected as the nodes for which collocational patterns were analyzed: *use, show, result, model, electron, current, device, region, temperature, voltage, figure*. The collocational patterns of *use* is discussed here to show how they are different from the collocational patterns of the same word in general English, as shown in *The BBI Dictionary of English Word Combinations* (Benson, Benson & Ilson 1997).

1. as noun: *by use (of); make use of; (the) use (of)*

2. as verb: *to use (a model/variable, etc.)*

3. as past participle (in the pattern of noun + (commonly/widely) + used + as/to/for/in + noun/verb): *the models used in the simulation; the devices used for comparison; the theories used to describe; used as parameters*

4. as present participle (in an adverbial clause at pre- or post-modifying position): *using a process/method/model/ theory/eqn(17)/ measurements*

The underlined patterns above were not found in the *BBI Dictionary*, while a familiar

pattern in general English and listed in the *BBI*, that is, *be/get used to* was not found in our sample corpus.

In comparison to the collocational patterns in the *BBI*, the word *use* in scientific research papers, when used as a noun, has a much more limited number of collocates than in general English; in addition, it occurs frequently in the form of present or past participles. The many occurrences of *using...* clause in our corpus, functionally providing a “contingency” to what is described in the main clause (Quirk et al. 1972), are used for the rhetorical function of indicating the “means” of an action taken in the main clause. This pattern is not included in the *BBI*. The high frequency of this pattern characterizes one of the idiosyncratic lexico-grammatical features of scientific research papers, in which writers often need to indicate the method or equipment by which data are collected, measurement is taken, or theorems are proved. The collocational patterns of other high frequency words, such as *figure*, *show*, and *model*, also show significant differences that reveal the “scientific-academic usages” of these words.

四、計畫成果自評 (Self-evaluation)

This project investigates the phraseology of scientific research papers. The results show that there are a number of recurrent lexical phrases in each section of the research paper and they can be linked to the communicative functions of the section. The collocational patterns reveal the contextual dynamics of high frequency words in scientific research papers.

It should be admitted, on the other hand, that the sample size is not large enough and the analysis is, in a sense, subjective. The criteria established for the identification of the lexical phrases may not be stringent enough in terms of statistical significance. However, the main purpose and aim is pedagogical. Addressing the common problem of non-native writers of research papers from a practical perspective, we hope to collect conventionalized lexico-grammatical patterns to provide them with useful genre-specific models or choices of language use.

五、參考文獻 (References)

- Aijmer, K. (1996). *Conversational Routines in English*. London and New York: Longman.
- Allerton, D. (1984). “Three (or Four) Levels of Word Co-occurrence Restriction,” *Lingua*, 63, 17-40.
- Altenberg, B. (1998). “On the Phraseology of Spoken English: the Evidence of Recurrent Word-combinations,” in A. P. Cowie (ed.), *Phraseology: Theory, Analysis and Applications*. Oxford: Clarendon Press.
- Arnaud, P., & Bejoint, H. (eds.) (1992). *Vocabulary and Applied Linguistics*. London: Macmillan.
- Benson, M., Benson, E., & Ilson, R. (eds.) (1997). *The BBI Dictionary of English Word Combinations*. (revised edition) Amsterdam/Philadelphia: John Benjamins.
- Cowie, A. P. (1992). “Multi-word Lexical Units and Communicative Language Teaching,” in P. Arnaud and H. Bejoint (eds.), 1-12.
- Cowie, A. P. (1998). *Phraseology: Theory, Analysis, and Applications*. Oxford: Clarendon Press.
- Gledhill, C. (2000). “The Discourse Function of Collocation in Research Article Introductions,” *English for Specific Purposes*, 19, 115-135.
- Granger, S. (1998). “Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae,” in A. P. Cowie (ed.), *Phraseology: Theory, Analysis and Applications*. Oxford: Clarendon Press.
- Howarth, P. (1996). *Phraseology In English Academic Writing*. Tübingen: Max Niemeyer Verlag.
- Howarth, P. (1998). “The Phraseology of Learners’ Academic writing,” in A. P. Cowie (ed.), *Phraseology: Theory, Analysis and Applications*. Oxford: Clarendon Press.
- Huang, J., & Hatch, E. (1978). “A Chinese Child’s Acquisition of English,” in E. Hatch (ed.), *Second Language Acquisition: a Book of Readings*. Rowley, MA: Newbury House.

- Moon, R. (1992). "Textual Aspects of Fixed Expressions in Learners' Dictionaries," in P. Arnaud and H. Bejoint (eds.), 13-27.
- Moon, R. (1998). *Fixed Expressions and Idioms in English*. Oxford: Clarendon Press.
- Myles, F., Hooper, J., & Mitchell, R. (1998). "Rote or Rule? Exploring the Role of Formulaic Language in Classroom Foreign Language Learning," *Language Learning*, 48, 3, 323-363.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Pawley, A., & Syder, F. H. (1983). "Two Puzzles for Linguistic Theory: Native-like Selection and Native-like Fluency," in J. C. Richards and R. W. Schmidt (eds.), *Language and Communication*. New York: Longman.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A Grammar of Contemporary English*. London: Longman.
- Sinclair, John (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Vihman, M. M. (1982). "Formulas in First and Second Language Acquisition," in L. Menn and L. Obler (eds.), *Exceptional Language and Linguistics*. New York: New York Academic Press.
- Weinert, Regina (1995). "The Role of Formulaic Language in Second Language Acquisition: A Review," *Applied Linguistics*, 16, 2, 180-205.
- Wong-Fillmore, L. (1976). *The Second Time Around: Cognitive and Social Strategies in Second Language Acquisition*. Ph.D. dissertation, Stanford University.
- Wray, A. (1999). "Formulaic Language in Learners and Native Speakers," *Language Teaching*, 32, 4, 213-231.

