

# A Corpus-based Syllabus for Teaching the Lexical Pattern: [*The-N-is that*]\*

Hijeon Kim  
(Korean Bible University)

**Kim, Hijeon. (2004). A corpus-based syllabus for teaching the lexical pattern: [*The- N-is that*]. *Language Research* 40(1), 279-299.**

This study attempts to design a corpus-based syllabus for teaching the lexical pattern, [*The-N-is-that*], which is [*the thing is that*], [*the fact is that*], [*the truth is that*], and [*the trouble is that*]. We developed teaching materials in line with recent approaches of consciousness-raising (C-R) activities and data-driven learning (DDL). When designing the sequential material and implementing C-R tasks, we carefully choose activities that raised learners' awareness of these usage patterns within the systematic three-stage task framework (Pre-task, Task, Post-task). Concerning [*the-N-is-that*] pattern, Kim (2001, 2003a) found that it constitutes an independent formal idiom or grammatical construction due to its unique properties, that is the fixed determiner *the*, *is* and *that* complementizer, expressing epistemic modality. These linguistic findings were applied to the English teaching process for this paper. This paper adopts the corpus-based approach via UKSPOK corpus (1999), proposing that teaching these patterns provide an effective input to acquiring lexicons and grammar. A learner's awareness of these formulaic expressions can be used to increase their strategic competence in asserting claims, especially in English for specific purposes, –academic discussions, debates, and public speeches.

**Key words:** corpus-based teaching, usage-based grammar, grammaticalization, data-driven learning, Consciousness-Raising (C-R) activities

## 1. Introduction

This study aims to design a corpus-based syllabus for teaching the lexical pattern, [*The-N-is-that*], such as [*the thing is that*], [*the fact is that*], [*the truth is that*], and [*the trouble is that*]. This is called *The Fact*

---

\*An earlier version of this paper was presented at The First International Conference on ASIA TEFL held in Pusan, Korea. Special thanks go to Professor Jeongwoon Park and Seongha Rhee and three anonymous reviewers for their valuable comments.

is *That* construction (FIT), since the noun *fact* inherently denotes the epistemic certainty meaning shared by these members (Kim, 2003a). This attempt is in line with the need that language learning is also based on naturally occurring genuine data rather than made-up examples.

The recent usage-based grammar using large machine readable corpus data has stressed the importance of semi-fixed phrases, lexical patterns or phraseology e.g., [verb *one's way to*] (Hunston, 2002), and [*a sort of*] (Barlow, 2000). However, for teachers who wish to use corpus data, the major challenge is how the data-driven learning (Johns, 1997) from the corpus, can be integrated into a plan for an ordinary lesson. With this interest, we focus only on [*the thing is that*], [*the fact is that*], [*the truth is that*], and [*the trouble is that*], henceforth FIT, construction, instead of encompassing all members, in seeking to develop a teaching material for teaching its collocational structures, rhetorical functions, and syntactic, phonological, and semantic restrictions. For this study, we adopted the Corpus-based approach via UKSPOK (United Kingdom of Spoken English, 1999) corpus data. In doing so, the recent language teaching approaches such as the Lexical approach (Willis, 1996, forthcoming), the Consciousness-Raising (C-R) activities, and the data-driven learning (DDL) (Johns, 1997) methodologies were exploited.

Concerning the [*the-N-is-that*] pattern, Kim (2001, 2003a, 2003b, 2003c) found that it constitutes an independent formal idiom or grammatical construction due to its unique properties. The fixed determiner *the*, the almost obligatory present tense verb, *is*, and the complementizer, *that*, rarely allow negations or any modals before the verb, which serve to express epistemic modality in emphasizing the value of truth in a speaker's assertion. These linguistic findings were applied to the English teaching process in this paper. The basic argument for learning within usage-based grammar is derived from the need to focus on the whole pattern or paradigm rather than to teach an individual word or the sum of individual words. In this regard, the collocational patterns fit best with a corpus-based approach since native speakers' intuition is a poor guide at least for the areas of collocations, hidden meanings and frequency (Cook, 2001). So, confronted by students' questions like "Which is the more frequent determiner, *the* or *my* before N in [\_\_\_N is that]?", (*The trouble is that* vs. *My trouble is that*) or "Why does this pattern strongly prefer *the* instead of a possessive?," or "What is the difference between [*The thing is that*] and [*The fact is that*] in the use of adjectives before N?", most teachers (both native and non-native English teachers) might have some trouble in answering the questions,

unless they learned these patterns.

This paper first introduces the basic concepts or emphases when using corpora in English teaching. The next section will concentrate on developing a teaching material in order to teach the FIT pattern by using corpus. It is assumed that for the intermediate-advanced level students, the current FIT pattern can be a useful unit for grammar and vocabulary because the pattern: [*The-N-is-that*] itself has meaning (objective epistemicity) that is like a grammatical pattern and it enjoys some degree of productivity. In addition to this, the members of the FIT pattern can also be a unit of vocabulary teaching since they can function as a multi-word single lexeme.

## 2. Issues in Language Teaching Using Corpora

### 2.1. New emphases in Application of Corpora

In applying corpora to teaching FIT constructions, we need to take account of basic concepts, or new emphases, in the recent teaching approaches based on Usage-based Grammar (Barlow, 2000). These are summarized by Barlow (2000), Cook (2001), Johns (1997), and Willis (1996). The general emphasises are listed below; each emphasis will be discussed in more detail in the next section.

- (1) An emphasis on **Data-Driven Learning: DDL** (Johns, 1997) is due to the importance of phraseology (sequence preference) for example, complementation patterns, verbs in the passive, and association between lexical items, such as *shed+tears* or grammatical environment.
- (2) An emphasis on a **Lexical Syllabus** (Willis, 1996) is due to the view that the lexical pattern as an aspect of grammar. For example, the pattern [V N to N] as *conceded victory to the ruling party* has something to do with GIVING (e.g. *allot, assign, accord, bequeath, bring, cede, deliver*), and the pattern [N+Be+...ing] can be interpreted as the present continuous tense.
- (3) An emphasis on **Hidden Meaning** or **Semantic Prosody** (Louw, 1997) is derived from an extended context. For example, speakers utter *I must admit* when saying something that is a contradiction with previous remarks.

## 2.2. The Importance of Frequency and Data-Driven Learning

In view of the importance of usage in language description and change, directions for analyses were determined by the relative frequency of grammatical and lexical expressions in the corpus (Sheibman, 2002). In English learning, the more frequent patterns should be taught earlier than the infrequent ones because the proportion of frequencies in a corpus will be mirrored in real language. Therefore, frequent meanings or collocates of words in a corpus are becoming significant in teaching. However, there are some limitations in applying the frequency-based analysis. One of them is that it can suggest that very infrequent uses can legitimately be ignored. For example, Mindt (2000) reports that 98% of verbs in the past refer to past time and a hypothetical future, this is very rare indeed. It means that hypothetical meanings are unimportant in terms of frequency, although for a student learning hypothetical meanings, the past tense cannot be ignored. In this sense, the frequency-based analysis might bring out a problem.

Then how can we apply the frequency-based approach to English teaching in an effective way? In fact, the increasing interest in genuine language use and frequency has motivated the emergence of some new teaching methodologies. One of them is Data-Driven Learning (DDL), which uses corpora for teaching material. DDL does not teach a language feature, but presents learners with evidence, and then asks them to make hypotheses and draw conclusions. The theory behind DDL is that students act as language detectives (Johns, 1997), discovering facts about the language from authentic examples, and in doing so, a student may well notice something that no textbook covers. Because corpus data can reveal previously unnoticed patterns, by studying the corpus data in the form of concordance lines, students answer questions about language.

In accordance with the DDL approach based on corpus data, there is a need to develop a syllabus encouraging consciousness-raising activities (Willis, 1996). Consciousness-raising activities fit best with DDL, since C-R activities are designed to draw learners' attention to the language use rather than teaching rules. Such activities as "What is the difference between X and Y?" or "What is the frequent determiner in X?" or "Is it possible to alter a determiner in X?" are good examples. When applied to FIT constructions, the possible tasks would include "What is the difference between [*The thing is that*] and [*The fact is that*] in their use of determiners and

adjectives?” or “Can you list all the nouns that precedes the verb *Is*?” or “Do you find any common semantic characteristics among the nouns: *Thing, Fact, Truth, Reality, Trouble*? DDL also provides learners with corpus evidence and encourages them to research and make rules. Therefore, it suits best with a lesson that has such learner-centered activities built into it. The framework for the task-based learning by Skehan (1996) and Willis (1996) can be used effectively in an attempt to implement tasks in DDL. Especially, faced with the problems of how tasks can be sequenced, Skehan suggests three stages in task implementation: 1) Pre-task: Introduction to the topic and task, 2) Task: Activities / Post task, and 3) (public) Performance.

For the Task stage, Consciousness-Raising activities should be used to draw learners' attention to some language features in the texts. Consciousness-Raising activities include i) identifying a particular usage, ii) classifying similarities and differences, iii) building hypothesis and iv) exploring cross-linguistic differences.

### 2.3. Lexical Syllabus and the Importance of Phraseology

[*The fact is that*], [*The thing is that*], [*The truth is that*], and [*The reality is that*] (FIT) have a lexical pattern [*The-N-is-that*]. This pattern has its own meaning, that is, objective epistemic modality (see the details in Kim 2003a, 2003b) just like a single lexicon, but the pattern enjoys productivity to some degree with the variation of a head noun (N), just like a grammatical pattern. In this sense, the study of FIT might gain useful insight from the recent Usage-Based Grammar (Barlow, 2000) approach, which stresses the importance of lexical patterns or phraseology (preferred sequences of words) such as *VERB one's way to* (Hunston, 2000), *sort of thing, thing to do* (Barlow, 2000).

The phraseologies including collocation, phrases, verb patterns, and complementation patterns are at the heart of language use and accordingly crucial input to language learning (Sinclair, 1991). The reason comes from the hypothesis that most of language use is not creative in the Chomskyan sense, but is based mainly on the use of lexical phrases or semi-prefabricated chunks. In other words, the creativity of language comes in large part from modification of prefabricated structures, rather than the novel combination of lexical categories (Barlow, 2000). For example, the choice between the definite article *the* or the possessive *his* might not be an optional alternative but is rather constrained by the pattern they belong

to. The possessive (*his*) is typically found in the patterns of [*break/lose/hurt his leg*] while the definite article (*the*) in the patterns of [*shoot/hit/stab someone in the leg*] (Hunston, 2002). In phraseology, the emphasis on lexis derives from the assumption that patterns have both lexical and grammatical meanings and lexis cannot be distinguished from grammar (Hunston, 2001). As an example, the pattern of [NP+be+ing] indicates the grammar of the present continuous tense.

In the lexical approach, the lexical items with patterns are resources for language creativity and language change. As Barlow (1996) comments, learners can create schemata or frames for grammatical features when they experience authentic examples of the language. He suggests that learners are not seen as passive pattern extractors, but rather as cognizers to make cognitive distinctions.

Given this, in English teaching, there arises an increasing need for grammar that is lexically based and that can be integrated into a lexical syllabus by using a computer concordance. The lexical syllabus stresses i) the most common word forms, ii) the typical patterns of usage, and iii) the combinations they usually form (Willis, 1996).

Then the next question might be what to teach first. The simplest argument for learning lexicons is to teach the typical uses in a language first. Similarly, Willis (1996) argues that English is a lexical language and the most productive way to interpret grammar in the classroom is through lexical patterning. All patterns involving frequent lexical items are important in the classroom because learning the lexis means learning the patterns and therefore the grammar. When applied to the FIT pattern, teaching the typical uses of [*The fact is*] and [*The truth is*], should be taught first rather than [*The simple fact is*], [*The plain truth is*] because the zero adjective pattern far outnumbers the pattern with adjectives in the corpus (Kim, 2003a).

Accordingly, with all the arguments, we assume that if we can develop a carefully designed syllabus to teach the FIT pattern, learners can successfully induce their schemata for using the phrases: [*The thing is*], [*The fact is*], [*The truth is*], [*The trouble is*], and/or [*The reality is*].

#### 2.4. Hidden Meaning and Semantic Prosody

Usually a set of concordance lines are used to identify collocates or lexical patterns. However sometimes, a larger co-text or an extended context is required to understand the connotative message or rhetorical use. The

term *semantic prosody* is used to describe a word that is typically used in a particular environment, such that the word takes on connotations from that environment (Channell, 2000). As an example, the phrase *sit through* conveys the connotations of boredom, indicating something lengthy and boring (Louw, 1997). The features of semantic prosody can be summarized as follows, based on Hunston (2002):

- (1) The semantic prosody of a lexical item is a consequence of the more general observation that meaning can belong to a whole phrase rather than single words.
- (2) It accounts for connotation in addition to its real meaning.
- (3) The semantic prosody of a word is often not accessible from a native speaker's conscious knowledge.

Especially, the semantic prosody bears significance in the area of language teaching because pragmatic meanings are beyond the reach of intuition, and therefore need to be taught. For example, Channell (2000) notes that the phrase *par for the course* is used not only to denote something, which frequently happens, but also to evaluate that event negatively. Native speakers of English often react with surprise to information of this kind: "Of course that's true, why didn't I think of it before?" Likewise, concerning the use of FIT constructions, when native speakers are asked about how and why the phrases of [*The thing is that*], [*The fact is that*], [*The truth is that*], [*The reality is that*], [*The trouble is that*] are used by people, their typical responses are "I don't know. So, vocabulary teaching needs to take account of the pragmatic usage: how and why speakers use them and what connotations are attached to them because not only linguistic knowledge but also pragmatic awareness increases learners' fluency of the target language.

Now let us go back to the issue of the semantic prosody of FIT that has the following connotative meaning: the following statement after [*the fact/ truth/reality/trouble/thing is that*] is certainly true and is strongly asserted by the speaker. It also follows the rhetorical pattern: False Claim (refuse other previous claims) + disconnector (*but*) + Correct Claim (assert the speaker's claim as true). The rhetorical strategy often exploited in the use of FIT is related with CONTRAST or ANTITHESIS (Kim, 2003a, 2003c).

### 3. Teaching Materials for FIT

While admitting these concepts or advantages of corpus-based teaching, teachers have some reservations in applying a corpus material in the real classroom. The major challenge for the teacher who wishes to use corpus data is to devise a task or a syllabus in such a way that students will obtain maximum benefits from it. More specifically, many teachers have reservations when considering how the Data-Driven Learning (DDL) from corpus data can be integrated into the plan for an ordinary lesson.

In this sense, we propose teaching materials for FIT, which reinforce Consciousness-Raising activities and Data-Driven Learning. This will entail syllabus design procedures in line with the activities. Also, students will be encouraged to find similarities and differences between patterning in Korean and in English: '*sasileun,*' '*keukei,*' vs *The fact is, The thing is.*

#### 3.1. Design of Teaching Material for FIT

For teaching English for specific purposes, the issue of what to teach is of particular significance. If students are required to discuss in academic settings, debates, or public speeches, the FIT patterns will provide a useful strategic formula in advancing their arguments more effectively and objectively. The reason comes from the fact that the patterns of [*The thing is that*], [*The truth is that*], [*The reality is that*] and [*The fact is that*] can be exploited to make people's assertions more impersonal and powerful. However, it is only recently that corpus studies have become assumably significant in designing a syllabus. Therefore, the current work here will be one of the few experimental reports.

Here we suggest a way of how FIT data in corpus can be made available for Consciousness-Raising activities to teach the FIT patterns. Particularly, we address how a series of C-R activities can be sequenced according to the various aims and stages in order to raise students' awareness of the usage of the FIT patterns and finally to let them make hypotheses about the usage patterns. Teaching the FIT pattern in the DDL approach needs to take into account the following emphases:

- (1) To organize pre-and post-task activities
- (2) To draw the data, which illustrate the targeted feature (FIT) from genuine data-corpus (UKSPOK in COBUILD)

- (3) To isolate specific features in FIT patterns for a focused attention: the typical use of the determiner (*the*), the frequent linking verb (*is*), the use of adjectives before nouns, the types of nouns, the lexical pattern, the hidden rhetorical structure, and the pragmatic connotation of FIT constructions.

### 3.2. Consciousness-Raising Activities for FIT Based on a Spoken Text

The following is the possible three-stage Consciousness-Raising (CR) activities we designed for providing a useful exposure for students to recognize the usages of [*The thing is that*], [*The fact is that*], [*The truth is that*], [*The reality is that*], [*The trouble is that*].

#### 3.2.1. The Pre-task

The pre-task phrase ensures a smooth and natural transition because this phase helps students to understand the task (e.g., by brainstorming ideas, playing a recording or reading a text). Students can be given some preparation time to think how to do the task, and this seems to result in better quality language use at the task stage (Willis, 1996).

#### Activity 1: Brainstorming ideas

Write down at least five expressions used to emphasize speakers' claims.

Teachers first start a lesson with the questions like “Do you know how speakers emphasize their claims?” or “Do you know what kind of expressions are used to make the assertion powerful and effective in debates, public speeches, or discussion sessions?” The teachers ask students to discuss their personal experience as a group and write down the expressions (a possible answer would be *in fact, it is true, truly*).

#### Activity 2: Find the expression in a recording of conversation

Find the expression in a dialogue served to make claims powerful. If you find it, tell what stress pattern was used.

Once students are aware of the general topic of the lesson, the next activity is to introduce more clearly the specific topic, that is, the FIT patterns: [*The thing is that*], [*The fact is that*], [*The truth is that*], [*The reality is that*], [*The trouble is that*]. To raise the awareness of the FIT

patterns, students listen to the recording of a short dialogue in which [*The truth is that*] is recoded. If they find the expression, let them find the phonological stress pattern of [*The truth is that*] (Kim, 2003a, 2003c) by asking “Where does the stress fall in the truth is that?” If they cannot, let them listen to the tape several times so that they find the stressed noun (*truth*).

The following is the transcript of the recording of the dialogue used as an illustration. This is how one conversation goes.

He was always up there boasting about the sexual prowess he never had encountered it because of this thing that happened when he was twelve years old. The truth is er by the time he was twenty-four er it scared him. <tc text=pause> Sex scared him rigid. He still had bad memories of school was unable to tell anybody and hadn't got round to the business of getting someone to talk to him or giving himself a

<Conversation including [*truth is*]>

### 3.2.2. The Task

Once learners understand the objective and topic of the lesson through the pre-task activities, they are ready to get exposed to written data of computer concordance lines searched by the key node [*fact is*], [*truth is*], [*reality is*], [*thing is*].

Through the computer concordance lines, students find the linguistic, grammatical collocations, lexical patterns, hidden pragmatic meaning, and rhetorical structure in the usage of FIT constructions. This process belongs to identification and consolidation task types.

#### Activity 3: Identifying collocations

##### Activity 3-1:

Look at the concordances below and find the determiner preceding [*fact is*], [*truth is*], [*reality is*], [*trouble is*] what is the frequent determiner?

The activity 3-1 uses the determiner as a starting point to highlight the fixed determiner (*the*). In order to teach the idiomatic usage of the definite article (*the*), teachers can ask students “Do you think it is better to say [*This fact is that*] or [*The fact is that*], [*My trouble is that*] or [*The*

*trouble is that*], [*My reality is that*] or [*The reality is that*], [*A truth is that*] or [*The truth is that*]?"

it have made then? <M04> Well the fact is that the <M01> I mean it wouldn't about it but <ZF1> the the <ZF0> the fact is that well again it comes back to <M01> markets. Now the plain fact is that they couldn't even with are pathetic but <ZF1> the <ZF0> the fact is that if you are writing about the say glorifying at certain points the fact is that you portray an image of Ds are cheaper in the States but the fact is that of course everything is course in the transitional period the fact is that it was demonstrably and side B starts 076 <ZZ0> <M02> But the fact is that childbirth is much more That' right. Yeah <F01> I mean the fact is that <M02> <ZF1> of the of the it ends up more polished er but the fact is that you can open up cans of er er but er er <ZF1> the <ZF0> the fact is that the country now although

< Concordances of [*fact is*]>

I won't spend it on drugs. But the reality is someone who compulsively spends and mixed I think <F02> Hm <M01> the reality is it's that it is an area that's these people are infertile. But the reality is that infertility is a major the findings. <F02> I think the reality is that for many many people not horrific the incident on Friday the reality is that we don't want it to get and impose one. <M01> Mhm. <M02> The reality is that people do these things the television <ZF1> the <ZF0> the reality is that they cannot in fact behave people <M01> Yes <M02> Well the reality is that it wasn't. And <M01> Mm and all that kind of thing. The reality is that adults don't do it. <F01>

<Concordances of [*reality is*]>

to write it <F02> Yes <F04> And the trouble is by the time they've got <FOX> t. I liked her but <F04> Yeah. The trouble is of course they had different too. She rehabilitates them The trouble is she doesn't actually like she one. Or was it some tables. But the trouble is I've probably just put down t paid. <F01> But I think Yes. But the trouble is they say as soon as something <ZGY> <M03> Uh huh. Mm <M04> The trouble is I had an accident twenty-five a sample urine sample. Er but the trouble is that I mean for a start she Don't you <F02> actually. And the trouble is that when we have tea it's a tea one. She's got <F02> Yeah. But the trouble is is now you're going to have to go whether <ZGY> questions <ZGY> Well trouble is I'm not There's there are that might happen <FOX> But the trouble is I mean <ZZ1> place name <ZZ0>

<Concordances of [*trouble is*]>

<tc text=sighs> I mean the truth is it's a pittance is given out to pensioners she is dyslexic she still doesn't know why. The truth is I suspect reading between the lines that Now she was in a caring profession but the truth is it was an unfortunate atmosphere at that when he was twelve years old. The truth is er by the time he was twenty-four er it that may be a contributory factor The truth is we don't know do we <ZF1> and <ZF0> and lit really is a good case with Nature Cos the truth is er I don't think it reflects much on Yeah. <M02> Er you know the truth is er if you want to learn how cyclogenesis music <ZZO> chances are not but the truth is that MX is still loved desperately by this finds it out Because we all know the r the truth is is that g ministers are saying in towards the last couple of weeks. Erm the truth is that the Sunday Express was in free fall or the potential user end but the I suppose the truth is that if every one really knew er that

<Concordances of [truth is]>

#### Activity 3-2:

Look at the concordances and find the most frequent complementation after [fact is], [truth is], [reality is], [trouble is].

This activity focuses on the most frequent collocation after [N is,] (i.e., a *that* clause). In order to make it clear, teachers can ask students “Do you think it is better to say [*The fact is to go there*] or [*The fact is that I go there*], [*The trouble is to meet her tomorrow*] or [*The trouble is that I'm gonna meet her tomorrow*]?”

#### Activity 3-3:

Look at the concordances and list all the adjectives before *Fact, Reality, Truth,* and *Trouble.*

Ask students “Is it frequent to add adjectives before N?” Then observing students' respond, say “If not, why do you think this pattern strongly rules out the use of pre-modifiers before N?”

#### Activity 4: Classifying the semantic and grammatical category

After identifying collocates, students are required to classify them according to similarities and differences based on formal and semantic criteria.

## Activity 4-1:

Look at the concordances and list all the nouns preceding *Be*.

Look up the dictionary and define the lexical meaning of these nouns (*fact, truth, reality, trouble*).

What do these nouns have in common in terms of their meaning? How are they classified in grammar? Are they different from common nouns such as *book, pen, and girl*?

This activity draws attention to the shared meaning and use of the nouns that fit the [*The \_\_\_ is that*] pattern. When asked these questions, advanced learners might look up in the dictionary and recognize that the nouns: *fact, truth, reality, and trouble* share the meaning associated with factuality or truth. The following definition follows Oxford Advanced Learner's Dictionary (1980).

- Fact: something known to be true or accepted as true
- Truth: something that is true, belief accepted as true
- Reality: something real
- Trouble: something that causes worry

Only the meaning of the nouns *trouble* differ from other nouns, but it still belongs to the group of factuality nouns since the noun *trouble* presupposes something real and just adds a worrying attitude. In this activity, students can recognize [*The-N-is-that*] pattern has constraints that allow for possible Ns, which have factual meaning. The common nouns such as *book, pen, and girl* cannot occur in this pattern.

## Activity 4-2:

Look at the below concordances of [*feeling is*], [*belief is*], [*assumption is*]. Compare the nouns: *fact, truth, reality, trouble* vs. *feeling, guess, belief, assumption* in their use of the possessive determiner.

Do you find it natural to say [*My fact is that*] [*My reality is that*] [*My trouble is that*]?

Do you think it is natural to say [*My feeling is that*] [*My guess is that*] [*My assumption is that*]?

These questions focus on the semantic difference between the two groups of nouns: factual nouns (*fact, reality, truth, trouble*) and hypothetical nouns

(*feeling, guess, assumption*). Through this activity, students become aware of the fact that the [factual N + *Is* + *That*] pattern has the constraint that only the definite article (*the*) is allowed, while the [hypothetical N + *is*] pattern freely allows the possessive determiners such as *my* or *your*.

out of hand now. <M01> Mm. <F01> My feeling is it's <M01> Yeah. <F01> ,got out erm she said er you know your feeling about it and I said My feeling is I'm willing to pay half <ZZ1> or not. <M03> Yeah. Well I my feeling is if it does in the States it require a licence in the States. My feeling is that it would require a licence been making of violence. And my own feeling is that the vast majority of A W B from <M01> I mean I mean I my feeling is <ZF1> I <ZF0> I approve of the must always say No. And er my own feeling is that generally speaking <F03> Well you see the thing that my feeling is that <ZF1> you <ZF0> you have call and very few minutes. My feeling is that you have decided elected him <M02> Oh yes. <F01> Your feeling is how long has he been using this

<Concordances of [*feeling is*]>

their backs because the automatic assumption is Oh you're having sex. What <MOX> Yeah please. <MOX> I mean my assumption is that you are collectively things <tc text=pause> I mean the assumption is that because you can't you're <M01> Yeah. <F01> I mean the assumption is you have <M01> <ZG1> That's it <M01> Right. <F01> The assumption is we don't want 'em on our patch I still get misunderstood and the assumption is whether they say it or not isn't <MOX> It's an unchallenged assumption isn't it. We never <ZF1> we <XF> <ZGY> <M01> Yeah. Yeah. <M02> The assumption is <ZF1> the ocean <ZF0> <M01> God was there the place shook. My assumption is if the place shook the people faithfully <F01> Mhm <F02> So my assumption is she's got the car she's

<Concordances of [*assumption is*]>

things are thought about er and my guess is that sooner or later it'll begin to a tree. <F01> That's right er your guess is as good as anyone else's as to fo the legitimate history plays. My guess is it would be King John which is a <M01> <tc text=clears throat> <M02> guess is that that's broadly how they the papers <M01> Mhm. <M02> and my guess is that there was consultation also the er sea surface temperature but my guess is if you looked around a latitude purchasers. Erm <ZF1> w <ZF0> we our guess is that a lot of the purchasers a while <F01> Mhm <F02> You know my guess is erm it'd be a stretch to let Can we get on with it please <MOX> My guess is MX <ZGY> <MOX> So

<Concordances of [*guess is*]>

## Activity 5: Making rules or hypothesis (patterns of FIT)

Students were asked to generalize about the FIT building on the preceding work. In addition, if possible, they were asked to find the latent rhetorical patterns of FIT.

## Activity 5-1

Can you find the general pattern of [*The fact is that*], [*The truth is that*], [*The reality is that*], [*The trouble is that*]?

Based on the tasks so far, students may easily draw the rule of [*The+factual N+is+that*-clause]. If they seem to make the formula with no difficulty, they can go one step further.

## Activity 5-2

Compare the pattern [*The+factual N+is+that*] with the following data of the Noun THING.

Thank you indeed for those. The only thing is that of course it completely ruins the game, might be a vacancy somewhere you see and the thing is that I got some experience somewhere and <tc text=laughs> however he said erm The thing is that you read the book on Roman law trapped <M01> Mm. <F01> Yes. Erm the thing is that you've got to be two jumps ahead <M01> <tc text=laughs> <M02> Amazing thing is that I still got a degree. <tc text=laughs> nothing but erm I <tc text=pause> the main thing is that gerontology then I'm in touch with Er b <ZF1> the other thi <ZF0> the other thing is that there was er military service ert for a lark. <M01> The funny thing is that I was talking to somebody from Edmund precision. But was it worth it. The only thing is that it makes you so disillusioned with the join or not is not the issue. But the thing is that I've come here to say to you you <M01> handed down <M02> Yeah. But <ZGY> thing is that though that it hasn't because <M01> Mmt more ill <F02> well the only other thing is that if I if it came to it was I'd have to go I don't know. Well I mean the only thing is that erm I'd have to well I don't know whether The only thing is that <ZF1> you you <ZF0> you run into problems. FOX> The amazing thing is that though you can get everything.

<Concordance lines of [*thing is*]>

In this activity, students are encouraged to recognize the different patterns between [*The fact is that*] and [*The thing is that*]. The sequence of [*The thing is that*] enjoys a high proportion of pre-modifiers before THING, therefore yielding the pattern [*The + adjective + thing + is + that*], whereas

[*the fact is that*] sequence dominantly excludes adjectives, thus yielding [*The + fact + is + that*].

Activity 6. Hidden meaning or rhetorical structure

- Students need to learn the context in order to learn how and why the speaker uses FIT.
- This task is beyond the lines of concordance requiring more extended context.

Activity 6-1:

Compare the following pair of sentences: one with the preface [*the fact is that*] and the other without this phrase.

- It is the most comprehensive and recent book.
- The fact is that it is the most comprehensive and recent book.

By comparing the pair of sentences, students might think [*the fact is that*] is used by the speaker to emphasize his or her assertion or to make claims more powerfully. Following questions would be:

- What do you think of the function of the addition of [*the fact is that*]?
- Do you find any rhetorical pattern in the speaker's use of FIT? Please read again the whole context where [*The truth is that*] is situated.

alone in the house is tough for him because he hasn't really got anyone to turn to FX has her parents at least and that's why the song is suitable. Whether they get together or not in the future I don't know <ZZ1> music <ZZ0> chances are not but the truth is that MX is still loved desperately by this lady and that makes the words so appropriate. <ZZ1> music 450-510 <ZZ0> It's pretty hard to take in but FX's beginning to realize that her future won't be with MX <ZZ1> music <ZZ0> as they'd planned all

<extended context of [*The truth is that*]>

In the above context, [*the truth is that*] occurs within the relation of contrast, in which a correct claim is contrasted with a dubious or false claim. Usually a speaker's statement comprises the correct claim following the phrase [*The truth is that*]. By using, [*the truth is that*], the speaker refutes a previous view as false and by doing so, highlights the truth validity of his or her own viewpoint. The rhetorical pattern of the phrase [*The truth is*

*that*] usually has a stressed N (*truth*) marking contrasted meaning. In other words, the stressed noun *fact, truth, reality* imply the contractiveness between false (apparent) truth and the truth (Kim, 2003a). The high proportion of the contrastive connective *but* also hints at the antithesis relation. The argumentation pattern of [*the truth is that*] in contrastive relation, can be schematized as:

False claim + contrastive + *the truth is that* + Correct claim  
 connective (*the fact is that*)  
 (*but*)

<Argumentation pattern of FIT>

He's alone in the house is tough for him because he hasn't really got anyone to turn to FX has her parents at least and that's why the song is suitable. Whether they get together or not in the future I don't know <ZZ1> music <ZZ0> chances are not but the truth is that MX is still loved desperately by this lady and that makes the words so appropriate. <ZZ1> music 450-510 <ZZ0> It's pretty hard to take in but FX's beginning to realize that her future won't be with MX <ZZ1> music <ZZ0> as they'd planned all.

The appointments were a waste of time that the treatment that the psychiatric nurse was giving her were a waste of time She continued to have a lot of ups and downs but she managed to return to work. Now she was in a caring profession but the truth is it was an unfortunate atmosphere at that time er fairly gossipy and rather bitchy. Er she didn't have an awful lot of support at work. In May FX asked her to leave her house because she simply couldn't cope with FX's depression.

<Conversations including [*the truth is*]>

The above conversations where the [*the truth is that*] occurs, can be analyzed according to the argumentation pattern of FIT.

False claim	Contrastive Connective	The truth is	Correct claim
<i>He hasn't really got anyone to turn to</i>	<i>but</i>	<i>The truth is</i>	<i>FX is still loved desperately by this lady</i>
<i>She was in a caring profession</i>	<i>but</i>	<i>The truth is</i>	<i>It was an unfortunate atmosphere</i>

< [*The truth is that*] in a CONTRAST relationship>

Activity 6-2: Cross-linguistic exploration: Compare Korean with English. Students are encouraged to find similarities and differences between patterning in Korean and English.

How [The fact is that], [The thing is that], [The truth is that], [The reality is that] can be translated into Korean?

Interestingly, the patterns of [The truth is that], [The fact is that], and [The reality is that] can be translated into one Korean adverbial 'sasileun' meaning *in fact*. This might provide additional cross-linguistic evidence that the patterns of [The truth is that], [The fact is that], and [The reality is that] can comprise one functional category in English.

Just like English, [The thing is that] shows a somewhat different usage pattern in Korean. The Korean counterpart of [The thing is that], 'keukei' (the thing) implies a very weak lexical meaning of *the thing* 'keukeot' because in both Korean and English the noun *thing* is extremely bleached and both [The thing is that] and 'keukei' cannot convey strong factual meaning as opposed to [The fact is that] or [The truth is that]. This may be the reason why in both English and Korean, they often combine it with various adjectives such as *the interesting thing is*: 'jaemissneunkei', *the funny thing is*: 'usskineunkei', *the strange thing is*: 'isanghankei'.

### 3.2.3. Post-task

At the last stage of the lesson for FIT, building on the students' findings in the previous tasks, they become engaged in a similar task. They make a similar short dialogue in the security of a small group. Then, they are asked to volunteer to do the performance in front of the audience.

Activity 7:

Try to make similar dialogue on the issue of giving money as a gift, by using the expressions: [The fact is that], [The truth is that], [The reality is that] and the argumentation pattern of [False claim+but +The N is that + Correct claim].

## 4. Conclusion

Although corpora are widely acknowledged as a valuable resource in describing language, many teachers have reservations when considering

the use of corpus-based materials in language teaching. Responding to this, this paper offers the means by which corpus-based DDL can be integrated into the plan for an ordinary lesson. For teaching FIT, in particular, we have developed a teaching material in line with recent approaches: consciousness-raising (C-R) activities and data-driven learning (DDL). When sequencing and implementing C-R tasks, we carefully choose activities that present learners with corpus evidence and ask them to make hypotheses and draw conclusions on the syntactic, semantic, rhetorical and collocational patterns of FIT. In short, corpus-based materials, and a systematic framework consisting of three stages (Pre-task, Task, Post-task) have been proposed.

With regard to FIT constructions, the emphasis has been given to the whole pattern or paradigm rather than an individual word or the sum of individual words, in that the meaning associated with the FIT patterns is determined by the whole pattern. In other words, the collocational patterns in FIT constructions fit best with a corpus-based approach since native learn speakers' intuition is a poor guide at least for the areas of collocation, hidden meanings, and frequency (Cook, 2001).

It is hoped that the application of FIT patterns to learning and teaching English provide an effective input to acquire lexicons and grammar, especially for the intermediate-advanced learners. We assume that a learner's awareness of the FIT formula can be used to increase their fluency, more specifically, learn strategic competence in asserting claims, especially in English for specific purposes, such as academic discussions, debates and public speeches.

## References

- Barlow, M. (2000). Usage, blends and grammar. In M. Barlow and S. Kemmer, eds., *Usage-Based Models of Language*. Stanford: CSLI Publications.
- Channel, J. (2000). Corpus-based analysis of evaluative lexis. In S. Hunston and G. Thompson, eds., *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.
- Cook, G. (2001). The philosopher pulled the lower jaw of the hen: ludicrous invented sentences in language teaching. *Applied Linguistics* 22, 366-387.
- Hunston, S. (2000). Phraseology and the modal verb: a study of pattern and meaning. In C. Heffer and S. Hunston, eds., *Words in Context: A*

- Tribute to John Sinclair on His Retirement*. University of Birmingham CD-ROM.
- Hunston, S. (2001). Colligation, lexis, pattern and text. In G. Thompson and M. Scott, eds., *Patterns of Text: in Honour of Michael Hoey*. Amsterdam: John Benjamins.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Johns, T. (1991). Should you be persuaded: two samples of data-driven learning materials. In T. Johns and P. King, eds., *Classroom Concordancing*. University of Birmingham. CD-ROM.
- Johns, T. (1997). Contexts: the background development and trialling of a concordance-based CALL Program. In A. Wichmann et al eds., *Teaching and Language Corpora* (pp. 79-99). London: Longman.
- Kim, H.-J. (2001). Grammaticalization of the *The Fact Is That*-Clause construction. *Proceedings of the First Seoul International Conference on Discourse and Cognitive Linguistics*. (pp. 438-451). Seoul, Korea: The Discourse and Cognitive Linguistics Society of Korea.
- Kim, H.-J. (2003a). *The Fact- Is- That and Related Constructions: A Usage-Based Approach and Implications in English Education*. Ph. D. Dissertation. Hankook University of Foreign Studies.
- Kim, H.-J. (2003b). A corpus-based syllabus for teaching the lexical pattern: [The -N -is -that]. *Proceedings of The First International Conference on Asia TEFL* (pp. 284-286). Pusan, Korea: Asia TEFL.
- Kim, H.-J. (2003c). A usage-based phonology approach to the recent changes of [in fact] > [Fact]. *Proceedings of The Second International Conference on Discourse and Cognitive Linguistics* (pp. 550-559). Seoul, Korea: The Discourse and Cognitive Linguistics Society of Korea.
- Louw, T. (1997). Contexts: the background developed and trialling of a concordance-based CALL program. In A. Wichman et al eds., *Teaching and Language Corpora*. London: Longman.
- Mindt, D. (2000). *An Empirical Grammar of the English Verb System*. Berlin: Cornelsen.
- Scheibman, J. (2002). *Point of View and Grammar*. Amsterdam/Philadelphia: John Benjamin Publishing Company.
- Sinclair, J. M. (1991). *Corpus Concordance Collocations*. Oxford: Oxford University Press.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics* 17, 38-62.

Willis, J. (1996). A flexible framework for task-based learning. In D. Willis and J. Willis, eds., *Challenge and Change in Language Teaching*. London: Heinemann.

Willis, D. forthcoming. *The Lexical Syllabus*.

Hijean Kim  
Department of English  
Korean Bible University  
205 Sangye 7-dong, Nowongu  
Seoul 139-791, Korea  
E-mail: upennkim@hanmail.net

Received: Dec. 1, 2003

Revised version received: Feb. 16, 2004

Accepted: Feb. 21, 2004