EFL Learners’ Consciousness-Raising through a Corpus-Based Approach*

Shin Chul Hong
(Dong-A University)


Recent corpus-based approaches have emphasised autonomous and process-oriented methodologies focused on developing learners’ cognitive skills. This cognitive development is based on consciousness-raising being considered a necessary condition for language learning (Schmidt, 1990). From this viewpoint, the study was aimed at how this kind of corpus-based approach contributes to EFL learners’ interlanguage development. To this end, the study examined advanced Korean learners’ acquisition of English determiners via a corpus-based methodology designed to enhance their cognitive ability and skills. To achieve this, the study contrastively analysed two methodologies: traditional structuralist and corpus-based. The former focused on a product-oriented perspective which is commonly used in Korean grammar classes, while the latter focused on a process-oriented perspective which was based on learners’ exploration of corpus data. The study discovers that both methodologies lead learners to raise their consciousness of knowledge about the grammatical rules of language, but the corpus-based methodology also raises learners’ consciousness of the knowledge of how language is used at the level of noticing. This result suggests that this level of consciousness, raised via a process-oriented corpus methodology, can contribute to EFL learners’ interlanguage development.

I. INTRODUCTION

For the last thirty years or so, the use of corpus data in language learning has inspired new ideas and theories (Stewart, Bernardini & Aston, 2004). Even though there are arguable limitations (see Owen, 1993; Widdowson, 2000), corpus data have the potential

* This is a part of Ph.D thesis (Hong Shin Chul, 2009).
to offer improvements in language education (Biber, Conrad & Reppen, 1998; Conrad, 2000; Granger, 2002; Hunston, 2002; Kennedy, 1998; McCarthy, 2001; McEnery & Wilson, 2001, pp. 119-122; O’Keeffe, McCarthy & Carter, 2007; Sinclair, 1997). In this respect, one of the distinctive features is that learners and teachers can make use of authentic language which can be extracted easily from corpora. However, this has not been found to be easily applicable to the language classroom (Owen, 1993; Seidlhofer, 2003, pp. 77-82; Widdowson, 1998). For example, early corpus-based approaches to language pedagogy (e.g., the Cobuild project: see Sinclair, 1987) have been focused on a theoretical-descriptive perspective. Their fundamental assumption was that corpus-based descriptions (e.g., collocation, concordance, and semantic prosody) of native speakers’ language use could guarantee that learners would use native-like language. However, according to Gavioli (2005), exposing learners to descriptions of real language does not imply that they can learn its features. Furthermore, the simplistic presentation of corpus data, based on theoretical descriptions, cannot promise learners’ achieving native-like language use. In this regard, Kennedy (1998) insists that corpus data have to be used sensibly for pedagogical purposes. ‘Sensibly’ in this context may refer to a balanced attitude for towards applying corpus data to language pedagogy, which can take into account of both theory-based descriptions from a perspective of corpus linguistics perspective and pedagogical considerations. From this viewpoint, the word ‘sensibly’ can be a key word in describing how corpus data should be applied to language learning and teaching, on the basis of various pedagogical considerations. With regard to this sensible use of corpus data, theories of cognitive psychology have focused on social and psychological assumptions which have emphasised language process rather than language product, i.e. regarding language learning as a psychological activity which can move learners’ interlanguage closer to their target language system (Bernardini, 2000; Hymes, 1979; Manolopoulou-Sergi, 2004; Robinson, 2001; Skehan, 1996, 1998). This assumption may be closely related to what a corpus-based approach should be. In fact, teachers and applied linguists have doubts about the use of corpus data in the classroom because of the corpus-based methodology, based on a structuralist perspective in a simplistic manner, which is often used in EFL language classes. For this reason, if corpus-based methodology can explain how corpus data can be used to develop the learners’ cognitive processes or mechanisms involved in interlanguage development, it could be a sensible application for language learning. From this point of view, the purpose of the study is to investigate this kind of contribution in terms of the sensible use of corpus data for language learning and teaching.
II. THEORETICAL BACKGROUND

1. Self-discovery Learning

Self-discovery learning emphasises learners’ own discovery of the linguistic features which they should learn and want to know, from analysing corpus data (Aston, 2001; Bernardini, 2001; Gavioli, 2005; Johns, 1994). In particular, Johns (1994) suggests a model of data-driven-learning (DDL), which highlights learners’ developmental ability through the sequence: identify-classify-generalise. The distinctive feature of DDL is that learners are supposed to construct their own language, culture, and world knowledge from their own personal experiences (Tan, 2002). For the last ten years, DDL, as one of the typical corpus-based approaches, has been adopted in language learning, focusing on the relationship between corpus consultation and interlanguage development (Bernardini, 2002; Chambers, 2005; Cheng, Warren & Xun-feng, 2003; Cobb, 2007; Gaskell & Cobb, 2004; Kennedy & Miceli, 2001; Thompson, 2006). Chambers (2005) investigates direct corpus consultation, based on DDL, as an alternative to conventional consultation (e.g., dictionaries, course books or grammar) in undergraduate courses. The study reveals that learners benefit from corpus consultation rather than conventional consultation in their subsequent language learning. In particular, Kennedy and Miceli (2001) investigate specific learning patterns with regard to corpus consultation. In other words, their study focuses on how learners use corpus consultation for their language learning. They examine intermediate students at Griffith University in Australia. In this study, Kennedy and Miceli reveal that corpus research skills, based on DDL, contribute to developing appropriate habits as researchers who incorporate logical reasoning and techniques for language learning. In the above studies, DDL shows how scientifically-interesting corpus data can be applied to language pedagogy. However, even though they find that corpus consultation based on DDL contributes to language development, the DDL-based approaches still have limitations with regard to the problematic relation between scientific interest and pedagogical usefulness, as Kennedy (1992) pointed out. For example, Willis, Shortall and Johns (1995) argue that learners who are not motivated or intelligent in their language learning have great difficulty in handling corpus data under the DDL approach. In particular, Hadley (2001, 2002) points out the problems with EFL learners’ difficulty of in analysing corpus data, since they are not familiar with process-oriented DDL approaches. Even though the empirical studies of corpus consultation discussed above (Chambers, 2005; Kennedy & Miceli, 2001) demonstrate the contribution of DDL to learner’s subsequent language development, they focus on European educational settings. For this reason, the application of DDL may not imply that it works well with beginning or intermediate EFL learners, such as Koreans or Japanese who have quite different
educational backgrounds compared to learners in European countries. In this regard, some studies have been conducted in Korean educational settings (Kim & Chun, 2008; Ko, 2005; Yoon, 2005). For example, Kim and Chun (2008) investigate 48 intermediate and advanced Korean university students in an EFL setting. The study examines Korean students’ vocabulary learning on the basis of classroom experiments focusing on concordance activities. The results from the study suggest that DDL-based corpus consultation can foster students’ vocabulary learning. Even though the DDL-based methodologies adopted by the above studies offer a positive signal to language learning, in both European and Asian educational settings, they need to be carefully applied so as not to overload or demotivate learners. In other words, corpus-based methodologies should be sensibly applied to ESL/EFL classrooms on the basis of a balanced perspective between both theoretically-described corpus data and pedagogical considerations.

2. Consciousness-raising

One of the distinctive features of DDL approaches is to focus on process-oriented methodology which encourages learners to develop their cognitive skills and ability through analysing data from corpora (Gaskell & Cobb, 2004; Sun, 2003). The theoretical assumption of DDL is that learners’ cognitive development can be achieved through process-oriented methodology which can raise their consciousness of language use. The notion of consciousness is very useful for describing the learners’ cognitive processes involved in interlanguage development (Schmidt, 1990, 1992, 1993, 1994, 1995, 2001; Schmidt & Frota, 1986). According to Schmidt (1990), consciousness, in relation to language, can be defined as individuals’ subjective experience of language use in their mind, and has three dimensions: awareness, knowledge and intention. Among these three dimensions, the researcher will focus on consciousness as awareness, which is relevant to this study. More specifically, his operationalisation of consciousness as awareness also has three degrees: perception, noticing and understanding. The first level, perception, is to create an internal representation of external events (Baars, 1986; Schmidt, 1990). At this level, perception is not necessarily conscious, so subliminal perception is possible. Since learners’ degree of consciousness is too low, they cannot verbalise their internal representation. The second level, noticing, refers to intentional attention to internally-realized subjective experience which may or may not be verbalised (Schmidt, 1990). Therefore, verbal reports resulting from subjective experience are considered to be noticing. However, noticed experience is not always verbalised. It is possible, according to Schmidt, for individuals to notice something that they cannot properly verbalise. Noticing as the second level of consciousness, i.e. awareness, is rather different from perception in that noticing requires having more conscious cognitive loading than perception. For
example, when we are reading a book, we are normally concentrating on the content of the book rather than on music from the next room. However, we can perceive the music even though we do not recognise what kind of music it is or its words (Bowers, 1984). The third level, understanding, is related to reflection on what is noticed. In other words, noticed experience can be analysed and compared with other occasions. In general, this kind of mental activity, thinking, is considered as understanding. According to Schmidt, noticing plays an important role in controlling the cognitive process between input and intake, and is a necessary condition for language learning. In the literature, Schmidt’s idea of noticing has been applied to various areas of SLA and language pedagogy (Ellis & Schmidt, 1997; Lynch, 2001; Mackey, 2002, 2006; Nitta & Gardner, 2005; Rosa & O’Neill, 1999). The above studies are based on tasks designed to raise learners’ noticing from a pedagogical perspective. In a broad sense, the tasks which these studies used are likely to be similar to consciousness-raising (CR) tasks, which share key characteristics with consciousness as awareness at the level of noticing. For example, Ellis defines CR tasks as:

A pedagogical activity where the learners are provided with L2 data in some form and required to perform some operation on or with it, the purpose of which is to arrive at an explicit understanding of some linguistic property or properties of the TL [target language] (Ellis, 1997, p. 160).

According to the above definition, CR tasks are supposed to raise learners’ consciousness on language to ‘explicit understanding’, which can be equal to Schmidt’s third level of understanding. Even though Ellis does not specifically mention noticing, the second level of consciousness as awareness, it can be loosely included in the boundary of explicit understanding. In this regard, recent empirical studies (Lynch, 2001; Mackey, 2006; Shekary & Tahririan, 2006) show how CR tasks based on learners’ noticing (i.e., explicit understanding) can be applied to language learning in communicative situations. Mackey (2006) investigates 28 ESL learners who joined a university-level intensive programme. She uses tasks designed to raise learners’ noticing of problematic grammatical forms through classroom interactions focusing on meaning negotiation and interactive feedback. The study concludes that noticing can be raised by tasks based on interaction, particularly feedback, and that this raised noticing has a close relationship with learners’ L2 development. Similarly, Lynch (2001) examines six EFL students attending pre-sessional courses in a university. In particular, he investigates noticing, focusing on cognitive process. From this study, he concludes that when learners’ output, externalised through formal correction or semantic precision, is transferred to input, it can enhance their noticing, which may contribute to their language development. Shekary and Tahririan (2006) investigate the effectiveness of noticing through a synchronous form of
computer-mediated chat (SCMC) methodology where learners type their conversations in real time, via the Internet. The main idea of the study is that learners’ noticing can be enhanced by meaning negotiation in a collaborative learning context. Shekary and Tahririan discover that online meaning negotiation promotes learners’ noticing which may contribute to language development.

With regard to the sensible use of corpus data for language pedagogy, the researcher’s hypothesis is that Schmidt’s noticing (i.e. consciousness as awareness as the second level) is a very useful notion to support how corpus data contribute to language development from a cognitive psychology perspective. In other words, if corpus data based on a DDL perspective (i.e., process-oriented methodology) can be used to raise learners’ consciousness at the level of noticing, it can contribute to learners’ subsequent language learning. Few empirical studies have focused on how learners develop their cognitive processes on the basis of process-oriented DDL approaches from the perspective of cognitive psychology, in other words, consciousness. Even through the above empirical studies on corpus consultation (Chambers, 2005; Kennedy & Miceli, 2001) demonstrated how corpus data contribute to language learning, they did not describe how or what kind of cognitive mechanisms are involved in this contribution. For this reason, the aim of this study is to investigate how corpus data, based on process-oriented DDL approaches, can contribute to learners’ interlanguage development in terms of cognitive psychology. For this, Korean learners were empirically examined on the basis of the following research questions:

1. Which works best in terms of performance (i.e. grammatical accuracy) between corpus-based and traditional structuralist methodologies?
2. Which works best in terms of consciousness-raising between traditional structuralist and corpus-based methodologies?

III. METHOD

1. Participants

Thirty advanced Korean learners, with at least ten years of previous English study in EFL contexts, participated in the case study. In particular, they did not have specific knowledge of what corpus data are. They were 25 female and 5 male students in the department of English language and literature at a university in Korea. All the participants joined the study as a seminar, which took place between August and November, 2006. The seminar was a compulsory course where they were graded, according to their level of
achievement. In fact, 35 students registered for the seminar, but 5 students failed because they did not fulfil the requirements (e.g., the assignments, attendance, and so on), so they were excluded from this study.

2. Material Design

1) The lectures

The lectures were designed to deliver grammatical knowledge about the use of determiners, focusing on seven categories: countability, definiteness, indefiniteness, some and any, common expressions, some rules, and others. To guide them, three grammar books were used: *Practical English Usage* (Swan, 1995), *A Communicative Grammar of English* (Leech, 1994), and *A Student's Introduction to English Grammar* (Huddleston & Pullum, 2005). In particular, the contents of the lectures included not only grammatical knowledge and usage, but also the use of determiners that Korean learners find problematic, from a remedial perspective.

For this, three types of content were introduced: terminology, issue and methodology. First, learners were informed about some terminology (e.g., the zero article) used in the lecture in order to present the rules for determiners systematically. Second, learners were supposed to discuss some issues related to Korean learners’ typical errors or problems with determiner use (e.g., the notion of countability). In this regard, they were able to get useful advice on how to deal with these problem areas in terms of the linguistic differences between Korean and English. Third, some practical methodologies (e.g., an Internet search tool) for learning determiner rules were introduced. For example, learners could search for ‘British Museum’ on the Google web site, and then discover how native speakers use the article in this case. The rationale behind the content was that learners need to think about their problematic use of determiners.

2) The exercises

The exercises consisted of two different types, one for each group (Traditional group and Corpus group). The Traditional-group exercise was designed from a traditional perspective. In this context, ‘traditional’ means a more or less structuralist approach (see Johnson, 2001), which emphasises systematic patterns of language structure rather than social contexts. Unfortunately, this kind of approach is used all too often in grammar classes in Korea (Flattery, 2007). All examples in the exercises were based on invented language and focused on delivering grammatical knowledge and use from a product-oriented perspective. For this reason, the main role of the Traditional-group exercises was for students to be able to practise
what they learnt in the lecture; the exercises were based on two books: *How English Works* (Swan & Walter, 1997) and *English Grammar in Use* (Murphy, 1985). Some exercise items were created by the researcher. Meanwhile, the Corpus-group exercises were designed from a process-oriented perspective. In this regard, two aspects were emphasised: authenticity and cognitive process. In terms of authenticity, all examples were based on corpus data, extracted mainly from the BNC (British National Corpus) and presented in the following three formats: KWIC (Key Word in Context), sentence view and context view. In terms of cognitive process, the Corpus-group exercises were based on the DDL sequence (identify-classify-generalize), as in the following example (see Figure 1). The rationale behind these exercises was to enhance learners’ cognitive ability and skills through analysing corpus data in an inductive manner. Thus, learners were supposed to raise their consciousness of both what to learn and how to learn.

**FIGURE 1**  
An Example of a Self-Discovery Item in the Corpus-Group Exercises

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a variety of churches providing practical help to</td>
<td>a number of</td>
<td>guests already referred.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Assistant Manager, Maurice Adams, to identify</td>
<td>a number of</td>
<td>church-based projects for the coming year. These</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>you are agreeing to make regular payments every</td>
<td>a number of</td>
<td>years; it is probably easier for you to pay by</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A new <em>Anti-Thermon Law</em>, which combines</td>
<td>a number of</td>
<td>new provisions for the investigation of serious</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>prisoners convicted of criminal offences, but</td>
<td>a number of</td>
<td>prisoners of conscience will also be released.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>of this book is nevertheless very impressive and</td>
<td>a number of</td>
<td>the issues he discusses impinge on my concerns as an</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>a February 1999 Dr. Manwaring argued</td>
<td>a number of</td>
<td>Animer's meeting last autumn.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>but link management as well. You get</td>
<td>a number of</td>
<td>experience and shop stewards who are Adam West.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>and the whole character has changed in</td>
<td>a number of</td>
<td>places. Finally, completely unscrambled, the calculations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>what was has been taken by the catalogue on</td>
<td>a number of</td>
<td>questions. The auction catalogue can also quote from.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>the historical sources can be followed up in</td>
<td>a number of</td>
<td>publications, both about individuals and institutions.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>these brief generalised statements can be found in</td>
<td>a number of</td>
<td>books, some by investigative journalists, some in museums</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>note about the sculptor. There are</td>
<td>a number of</td>
<td>other ways in which the work could.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Adirondacks, In London, where he spent</td>
<td>a number of</td>
<td>years, he had gained a reputation as a crook and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>the novels of cosmic-chaos, and so on</td>
<td>a number of</td>
<td>ways a double book. It consists of two</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1) Try to find some patterns in the above box.

2) In Line 6, the definite article is used with ‘a number of’. Guess why, on the basis of the following context:

If there are aspects of the science of art that Kemp fails to engage adequately, the scope of this book is nevertheless very impressive and a number of the issues he discusses impinge on my concerns as an historian of science. Let me allude to two of these issues.

3) The personal journals

The personal journals were designed to investigate learners’ cognitive processes of language learning on the basis of their own descriptions of what they had learnt from each class (Oxford, 1990; Schmidt & Frota, 1986). In fact, think-aloud methodology is often
used for this in SLA (Bachman & Palmer, 1996; Johnson, 2001; McDonough & McDonough, 1997; O’Malley & Chamot, 1990; Raimes, 1985). However, such methodology, focusing on simultaneous verbal reports during a task, was unlikely to be effective in this study. According to the pilot study, which focused on Korean learners’ acquisition of English articles, they had great difficulty in reporting concurrently what they were thinking about. For this reason, it was desirable for learners to have enough time to describe their thinking, in vivid detail, by keeping a personal journal. Additionally, they were allowed to write their personal journals in Korean or English.

4) The essays

The essays focused on measuring the learners’ determiner use in a naturalistic situation and consisted of three parts: instruction, essay topic, and notes. In the case of instruction, the essay length and the use of reference books were introduced. With regard to the essay length, 500 words were considered reasonable for Korean learners. Since there had been no specific research on the length of ESL/EFL learners’ essays associated with their grammar use, the essay length was determined by practical constraints. For example, Hong (2004) analysed the errors of advanced Korean learners on the basis of argumentative essays which were 500 words long. According this research, 500-word essays are sensible for examining Korean learners’ grammar use. With regard to the essay topic, learners were supposed to write two essays (Pre/Post-essays) on the same topic (let’s talk about the key stages in your life so far) for contrastive analysis. The Pre-essay was focused on learners’ use of determiners before the seminar, and the Post-essay after the seminar. Since there was at most a three-month gap between the two essays, the same essay topic would be very useful for investigating how learners differently used the determiners during this period. Furthermore, a different essay topic might influence on learners’ language use (Weigle, 2002, pp. 91-94). In this regard, a limitation was that learners might use the same contents and grammar with just minor changes in their essays. To address this, the researcher strongly recommended rewriting essays and informed them how to write a Post-essay with the same topic. Furthermore, learners were not allowed to use additional reference books (e.g., grammar books or textbooks) or use proofreading by native speakers.

1 The pilot study was carried out in 2005 and focused on Korean learners’ acquisition of English articles from the perspective of cognitive psychology. Briefly, two methodologies were adopted for data collection: personal journals and concurrent verbalisation. In the former, six advanced Korean learners were supposed to describe what they had learnt from each lecture and exercise. The latter was to record concurrent verbalisation of their written test.

2 The researcher’s recommendation was limited to two cases: essay length and copied essays. For example, learners who wrote the essays less 500 words and copied essays based on their previous were recommended to rewrite them.
5) The interview

The interview was designed to elicit learners’ strategies for using determiners to investigate their cognitive processes. These consisted of pre-prepared and elaborate questions in two stages: an interview test and an interview about the interview test. First, the interview test was designed using the same test questions as the Pre-test. Since there was a three-month gap between the two tests, learners might not remember specific test items. Second, the interview about the interview test hinged on the following question: how did you choose your answers in the test? Even though the interview was guided by the above specific question, its format was open-ended and interviewees were encouraged to describe what they were thinking of. According to Dörnyei (2007), this is called a ‘semi-structured interview’. A distinctive feature is that learners were supposed to take the Pre-test again, as the interview test, and then they had to answer the above question. The rationale behind this procedure was to analyse learners’ cognitive strategies or processes through learners’ own explanations, and to investigate any particular patterns in the two groups in the use of English determiners from the perspective of cognitive psychology.

3. Procedure

The case study was based on a ten-week seminar series for Korean advanced learners. The procedure of for the seminars consisted of six steps: orientation, grouping, classes, personal journals, final tests, and interview. The first step, of orientation, comprised four components: introduction, Oxford placement test (Allen, 1992), Pre-test, and Pre-essay. In the introduction, learners had all many kinds of information about the seminar (e.g., the procedure, assignment, lectures, and so on) for their voluntary contribution. In the placement tests, learners were supposed to take two types of tests: Oxford placement test (100 items/ 50 min.) for measuring their general ability of English grammar, and a Pre-test (50 items/ 30 min.) for measuring their ability in determiner use. In particular, the contents of the Pre-test were designed on the basis of the seven categories adopted in the lectures (see Section 3.2.1). In the component of the Pre-essay component, learners were informed how to write a 500-word essay within a two-week deadline. In the second step, the Korean learners were divided into two groups (Corpus group: 15 students and Traditional group: 15 students). Learners could not select their own groups. Instead, they were selected by the researcher according to the results of the two placement tests (Oxford test and Pre-test) in order to maintain a homogeneous level of learners’ grammatical ability (see Table 1: the groups have similar means in the test scores). For example, a list was made on the basis of all participants’ test results, and then it was used to construct the two groups according to higher the test scores. In the third step, classes consisted of two parts: lecture (20-30 min.)
and exercise (20-30 min.). In the first part, the two groups had eight lectures together, which these focused on specific grammatical topics regarding to English determiners. In the second part, learners were divided into the two groups (Traditional and Corpus) after the lectures, and were supposed to complete their own type of exercises (see Section 3.2.2). In the fourth step, learners were supposed to submit their personal journals which described what they had learnt in each class. These personal journals had to be written as homework, after each lecture and exercise, because learners needed sufficient time to think about what they had learnt from each class. Each journal was then collected at the start of the next class. In the Final-test step, there were two components: Post-test (50 items/30 min.) and Post-essay. The Post-test was designed using the same criteria as the Pre-test, to measure learners’ achievement. In the case of the Post-essay, learners were informed of the essay topic two weeks before the deadline. In the last step, there were two components: interview test (20 min.) and interview (30 min.). Learners took the interview test, which was the same as the Pre-test, and were interviewed on the basis of their interview test. The procedure in this interview was to ask them how they had chosen particular answers when answering the questions in the test.

4. Data Analysis

The data analysis of the case study had two key elements: triangulation and contrastive analysis. Triangulation was based on three types of data (test, essay and cognitive) which were analysed using the SPSS (Statistical Package for the Social Sciences) program. In this, various statistical scores and information were used. In particular, two types of t-test were performed: a paired-sample test and an independent-sample test. The former was used to compare differences between or similarities in each group, and the latter between the two groups (Traditional and Corpus). With regard to the second element, the three types of data were contrastively analysed between the two groups.

1) Test-based analysis

The test-based analysis focused on the accurate use of determiners in the tests: Oxford placement test and Pre/Post-tests. The purpose of this analysis was to measure learners’ grammatical knowledge of determiners from the tests. The purpose of the Oxford placement test was to assess learners’ general ability in grammar use, and that of the Pre-test their specific ability with determiners. The total score of the three tests was 100. In particular, Cronbach’s alpha was calculated to measure reliability of the tests. The reliability of the three tests was acceptable according to the Cronbach’s alpha (see Table 1).
2) Essay-based analysis

Unlike the test-based analysis focusing on learners’ grammatical knowledge – in other words, accuracy – the essay-based analysis was intended to investigate learners’ productive use of determiners. For this reason, the essays were analysed from two perspectives: accuracy and frequency. From the accuracy perspective, the analysis focused on learners’ errors in the use of determiners. Meanwhile, from the frequency perspective, the analysis focused on two factors: how many lexical determiners (both errors and non-errors) were used, and how many different types of lexical determiners were used in the two essays. The essay-based analysis was based on a manual coding scheme, consisting of two procedures: error finding and error/non-error coding. The first procedure comprised two steps. The first step was performed by a native speaker who was working as an English teacher in the department of Linguistics and English Language at Lancaster University. The second step was performed by the researcher. The second procedure also comprised two steps: the first step by the researcher, the second by the native English teacher. The coding had seven categories (e.g., countability, definiteness, indefiniteness, some and any, common expressions, some rules, and others) based on the content of the lectures. When there were disagreements about the encoding between the researcher and the native English teacher, we discussed and solved them in order to achieve inter-reliability of the coding scheme.

3) Cognitive analysis

The cognitive analysis was based on two types of data: personal journals and interviews. The first analysis focused on learners’ own journals, where they were supposed to describe what they had learnt in each class. For this, the learners’ journals were encoded via the following two steps: determiner categories and learner consciousness. The first step was based on the seven categories used in the essay-based analysis. However, the second was concerned with learners’ three levels of consciousness of what they were learning in the classes: ‘Don’t know’ (Level 1), ‘Not very clear’ (Level 2), and ‘Clear explanation’ (Level 3). The purpose of the three levels adopted in this study was to investigate learners’ degrees of consciousness as awareness in the two groups. For this reason, the above three levels were classified on the basis of the learners’ own descriptions in their journals. The theoretical background to these three levels hinges on Schmidt’s three levels of consciousness as awareness: perception, noticing and understanding. In this domain, Schmidt and Frota (1986) adopted this classification of learners’ self-descriptions in order to investigate the correlation between levels of consciousness and language learning. In particular, Schmidt (1990) emphasises learners’ verbal reports as one of the most important criteria for classifying their consciousness levels. In other words, if learners can
verbalise what they have learnt through seminars, then their verbalisation can be classified as consciousness at the level of noticing or understanding. If learners cannot verbalise it, this implies that they have failed to raise their consciousness, or that they have raised their consciousness at the level of perception. Since perception does not necessarily equate to consciousness (Baars, 1986; Schmidt, 1990), learners’ failure to verbalise can also be classified under perception. For this reason, the classification adopted in this study was not the same as that of Schmidt.

From this point of view, the study classified the Korean learners’ verbal reports into these simplified categories: Level 1 (‘Don’t know’); Level 2 (‘Not very clear’); and Level 3 (‘Clear explanation’). Level 1 (‘Don’t know’) refers to a lower degree of consciousness. In this level, learners fail to describe what they have learnt. Additionally, they mention this failure (e.g., I do not understand the use of the definite article ‘the’ as a general reference). For this reason, it is different from Schmidt’s perception which includes failure to verbalise in that learners at this level consciously raise what they do not know. Level 3 (‘Clear explanation’) refers to a higher degree of consciousness, similar to Schmidt’s ‘understanding’. At this level, learners can clearly verbalise what they have learnt. Level 2 (‘Not very clear’) refers to a degree of consciousness somewhere between perception and understanding — in other words, noticing. Here, learners’ verbalisation is not as clear as that of Level 3, but much clearer than Level 1, so they have some uncertainty over what they have learnt.

**FIGURE 2**

An Example of a Personal Journal

The most important information that I've come to know from the lecture is that 'any' can be used in an affirmative sentence. No English native speakers would be as surprised as I was because they use it naturally and correctly in affirmative sentences. But I was taught by my English grammar book long ago that 'any' is mostly used in negative sentence with some exceptions. However, many Korean students didn't think such exceptions are significant probably because such grammar point will not be covered at the school exams. Teachers and students altogether just satisfy the clear-cut definition that 'some' must be used in affirmative sentence and 'any' in negative sentence. However, I think the 'exceptions' in grammar book is not really 'exceptions' in real communication situations. Native speakers would use 'any' in so many affirmative sentences. After the lecture, I went home and typed in 'any' on the Internet English newspaper. I found many sentences where the word 'any' was used in affirmative sentences. Following are a couple of examples that I found:

(a) 'Our army and people are fully prepared to counter any provocation or threats of nuclear warfare.'

In (a), 'any' means 'all' or 'every', which is contrast with my preoccupied of the word meaning 'nothing' or 'not'. Now I think I am conscious raised whenever I encounter 'any' and 'some' are used in English texts.

For example, Figure 2 presents an extract of a personal journal in the Corpus group and focuses on the contents of 'some and any'. A learner mentions the use of 'any in a
non-affirmative statement (Subcategory 2) at the level of understanding (Level 3: ‘Clear understanding’). With regard to the coding scheme for the personal journals, it had two procedures, as with the essay coding schemes.

The second analysis of the interview was based on learners’ own explanations of the interview test. The interviews were analysed on the basis of the following levels of encoding:

1. ‘No explanation’ (Level A): Learners cannot explain why they chose a particular answer in the test.
2. ‘Unclear explanation’ (Level B): Learners seem to know the reason why they chose a particular answer in the test, but cannot clearly verbalise it.
3. ‘Obvious explanation’ (Level C): Learners can clearly explain why they chose a particular answer in the interview test.

For example, in Figure 3, the first item was encoded as Level C (Clear explanation) which means that a learner clearly explains how s/he chose an indefinite article in the first item. In order to reliably code procedures when learners’ explanations were not clear enough to be classified into the three levels, the researcher asked learners to explain again why they chose a particular answer in the test, and all interviews were recorded as MP3 files as a double check. The inter-reliability of the interview test was acceptable (Cronbach’s alpha: Traditional group=.643 and Corpus group=.655).

**FIGURE 3**
An Example of Interview Coding

<table>
<thead>
<tr>
<th>Tuesday 17 July</th>
</tr>
</thead>
<tbody>
<tr>
<td>We did a brilliant experiment during chemistry today. We put a substance (I forgot its name of it) into the water, and it gave off a smell like bed eggs! Denise couldn’t stand it and ran out of the lab. She said she went to the girls’ toilet and was sick three times. Gaining knowledge can be a painful experience.</td>
</tr>
<tr>
<td><strong>O</strong>: Level C (Obvious explanation)</td>
</tr>
<tr>
<td><strong>A</strong>: Level B (Unclear explanation)</td>
</tr>
<tr>
<td><strong>X</strong>: Level A (No explanation)</td>
</tr>
</tbody>
</table>
IV. RESULTS

1. Test-based Results

The test-based analysis focuses on the three test results: Oxford placement test, Pre-test, and Post-test. First, in the Oxford placement test, the two groups have similar test results (Traditional group: Mean= 53.60, SD= 9.92; Corpus group: Mean= 56.40, SD= 12.11) (see Table 1).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Tests</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxford placement test</td>
<td>Traditional group</td>
<td>15</td>
<td>9.92</td>
<td>53.60</td>
<td>0.785</td>
</tr>
<tr>
<td></td>
<td>Corpus group</td>
<td>15</td>
<td>12.11</td>
<td>56.40</td>
<td>0.774</td>
</tr>
<tr>
<td>Pre-test</td>
<td>Traditional group</td>
<td>15</td>
<td>10.23</td>
<td>57.60</td>
<td>0.684</td>
</tr>
<tr>
<td></td>
<td>Corpus group</td>
<td>15</td>
<td>8.96</td>
<td>57.26</td>
<td>0.645</td>
</tr>
<tr>
<td>Post-test</td>
<td>Traditional group</td>
<td>15</td>
<td>12.20</td>
<td>52.66</td>
<td>0.634</td>
</tr>
<tr>
<td></td>
<td>Corpus group</td>
<td>15</td>
<td>9.58</td>
<td>68.80</td>
<td>0.652</td>
</tr>
</tbody>
</table>

Even in the t-test, there is no significant difference between the two groups (t(28)= -0.69, p>.05). Second, in the Pre-test, the two groups do not show different patterns with the Oxford placement test (Traditional group: Mean= 57.60, SD= 10.23 and Corpus group: Mean= 57.26, SD= 8.96). The t-test also shows no significant difference (t(28)= -0.09, p>.05). For this reason, the two groups can, for practical purposes, be considered equal in their use of both general grammar and determiners.

However, in the Post-test, the two groups have different test results. The Corpus group shows improvement between the Pre-test (Mean= 57.26, SD= 8.96) and the Post-test (Mean= 68.80, SD= 9.58). In the t-test, the Corpus group achieves meaningful improvement (t(28)= -3.40, p<.05), but the Traditional group does not (t(28)= 1.20, p>.05). Nevertheless, the difference between the two groups is not enough to conclude that the Corpus group is better in terms of determiner use, since the test-based results focus only on accuracy in the tests. It is necessary also to investigate Korean learners’ use of determiners in a naturalistic situation, such as essays.

2. Essay-based Results

1) Accuracy

In the Pre-essays, the accuracy of the two groups shows a difference (Traditional group:
88% and Corpus group: 84%), but in the t-test, this difference is not significant (t(28)= 1.74, p>.05) (see Table 2). In the Post-essays, the Corpus group (Pre-essay: 84% and Post-essay: 95%) shows significant improvement in the Post-essays unlike the Traditional group (t(28)=2.18, p<.05). Consequently, the Corpus-group learners have achieved a higher level of accuracy in the Post-essays than the Traditional group.

**TABLE 2**

<table>
<thead>
<tr>
<th>Essays</th>
<th>Traditional group</th>
<th>Corpus group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-essay</td>
<td>88.29% (981/1111)*</td>
<td>83.95% (1020/1215)</td>
</tr>
<tr>
<td>Post-essay</td>
<td>91.42% (992/1085)</td>
<td>95.21% (1133/1190)</td>
</tr>
</tbody>
</table>

*(correct use of determiners/ total number of determiners used)

2) Frequency

The frequency is based on two analyses: how many lexical determiners are used and how many different types of lexical determiners are used in the two essays. First, in the analysis of the average percentage of determiners used in the essays, the two groups do not show particularly significant patterns (see Table 3).

**TABLE 3**

<table>
<thead>
<tr>
<th>Essays</th>
<th>Traditional group</th>
<th>Corpus group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-essay</td>
<td>15.57%</td>
<td>15.22%</td>
</tr>
<tr>
<td>Post-essay</td>
<td>15.12%</td>
<td>14.72%</td>
</tr>
</tbody>
</table>

In the t-test, the two groups show no significant difference between the two essays (Traditional group: t(14)= .73, p>.05; Corpus group: t(14)= 1.54, p>.05). Second, in the analysis of the category types used, the averages of the category types used in the two groups do not show significant differences in the two essays (Traditional group: t(14)= -1.08, p>.05; Corpus group: t(14)= -1.72, p>.05) (see Table 4). Therefore, the two essays in each group do not show any significant difference in terms of the number of categories used.
Consequently, in the analysis of frequency, the Korean learners do not show significant differences in terms of how many determiners are used and how different types of categories are used. The Corpus group shows progress in their essays in terms of accuracy, but not in terms of frequency. In the essay-based analysis, the focal point is limited to the investigation of Korean learners’ determiner use, not other grammatical aspects of language use (e.g., use of tenses or aspect). For this reason, the improvement of the Corpus group is judged by only two factors (accuracy and frequency).

3. Results for Cognitive Data

1) The Personal journals

The personal journals are assessed via two analyses: mentioned categories and consciousness level. First, the frequency of the mentioned categories in the groups is not very different in the journals (see Table 5). The t-test also indicates no significant difference between the two groups in terms of frequency of the categories mentioned (t(12) = -.25, p > .05). Second, the frequency of learner-consciousness levels is quite different in the two groups (see Table 5). In the Traditional group, Level 3 is higher than the other two levels. However, in the Corpus group, Level 2 is highest. This phenomenon is not very different in the consciousness levels in each category. In the above analysis, there is a clear difference between the two groups in terms of the consciousness levels of each category. One interesting point is that the Traditional group has a higher frequency of Level 3, referring to clear explanation and understanding of the categories with some degree of confidence. Therefore, it can be inferred from the above result that the Traditional group has relatively clear knowledge of determiners in their journals. Meanwhile, it seems that the Corpus group can explain general notions and rules of determiners though with some questions and doubts. In association with the above results, it needs more investigation to seek any correlation between the patterns of consciousness levels and learner production. This is investigated in two ways: the percentage of consciousness-raising via the lectures, and the percentage of consciousness-raising via the essays.

<table>
<thead>
<tr>
<th>Essays</th>
<th>Traditional group</th>
<th>Corpus group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-essay</td>
<td>23.8%</td>
<td>24.08%</td>
</tr>
<tr>
<td>Post-essay</td>
<td>25.44%</td>
<td>26.8%</td>
</tr>
</tbody>
</table>

TABLE 4
Average Percentages of Category Types Used

Consequently, in the analysis of frequency, the Korean learners do not show significant differences in terms of how many determiners are used and how different types of categories are used. The Corpus group shows progress in their essays in terms of accuracy, but not in terms of frequency. In the essay-based analysis, the focal point is limited to the investigation of Korean learners’ determiner use, not other grammatical aspects of language use (e.g., use of tenses or aspect). For this reason, the improvement of the Corpus group is judged by only two factors (accuracy and frequency).

3. Results for Cognitive Data

1) The Personal journals

The personal journals are assessed via two analyses: mentioned categories and consciousness level. First, the frequency of the mentioned categories in the groups is not very different in the journals (see Table 5). The t-test also indicates no significant difference between the two groups in terms of frequency of the categories mentioned (t(12) = -.25, p > .05). Second, the frequency of learner-consciousness levels is quite different in the two groups (see Table 5). In the Traditional group, Level 3 is higher than the other two levels. However, in the Corpus group, Level 2 is highest. This phenomenon is not very different in the consciousness levels in each category. In the above analysis, there is a clear difference between the two groups in terms of the consciousness levels of each category. One interesting point is that the Traditional group has a higher frequency of Level 3, referring to clear explanation and understanding of the categories with some degree of confidence. Therefore, it can be inferred from the above result that the Traditional group has relatively clear knowledge of determiners in their journals. Meanwhile, it seems that the Corpus group can explain general notions and rules of determiners though with some questions and doubts. In association with the above results, it needs more investigation to seek any correlation between the patterns of consciousness levels and learner production. This is investigated in two ways: the percentage of consciousness-raising via the lectures, and the percentage of consciousness-raising via the essays.
First, with regard to the percentage of consciousness-raising via the lectures, 54.4% (the Traditional group) and 63.9% (the Corpus group) of the determiner categories introduced in the lectures are mentioned in their journals (see Table 6). For example, the Traditional group mentioned 54.4% of the determiner categories in their journals among the total number of the determiner categories introduced in the lectures. The difference between the two groups in this regard is not significant (t(28) = -1.81, p > .05). The P-value in the t-test is 0.08, indicating that the difference is on the boundary between significance and non-significance. For this reason, the Corpus group learners may raise consciousness in more determiner categories in the journals than the Traditional group learners, but it is difficult to conclude that the Corpus group has better achievement in terms of consciousness-raising.

### TABLE 5

<table>
<thead>
<tr>
<th>Categories</th>
<th>Levels of consciousness</th>
<th>Traditional group</th>
<th>Corpus group</th>
<th>Levels of consciousness</th>
<th>Traditional group</th>
<th>Corpus group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Level 1</td>
<td>Level 2</td>
<td>Level 3</td>
<td>Total</td>
<td>Level 1</td>
</tr>
<tr>
<td>Countability</td>
<td>30</td>
<td>7</td>
<td>20</td>
<td>1</td>
<td>28</td>
<td>7</td>
</tr>
<tr>
<td>Definiteness</td>
<td>21</td>
<td>3</td>
<td>15</td>
<td>12</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>Indefiniteness</td>
<td>43</td>
<td>8</td>
<td>27</td>
<td>14</td>
<td>49</td>
<td>8</td>
</tr>
<tr>
<td>‘Some’ and ‘Any’</td>
<td>29</td>
<td>4</td>
<td>9</td>
<td>24</td>
<td>37</td>
<td>4</td>
</tr>
<tr>
<td>Common expressions</td>
<td>76</td>
<td>6</td>
<td>29</td>
<td>41</td>
<td>74</td>
<td>6</td>
</tr>
<tr>
<td>Some rules</td>
<td>43</td>
<td>7</td>
<td>12</td>
<td>19</td>
<td>54</td>
<td>7</td>
</tr>
<tr>
<td>Others</td>
<td>14</td>
<td>0</td>
<td>8</td>
<td>7</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Total (counts)</td>
<td>232</td>
<td>28</td>
<td>120</td>
<td>104</td>
<td>252</td>
<td>28</td>
</tr>
</tbody>
</table>

Second, with regard to the percentages of consciousness-raising over the essays, they show quite different patterns. The percentages of each group are 65.6% (Traditional group) and 81.6% (Corpus group) in the Post-essay. In the t-test, the two groups do show any significant difference (t(28) = -3.51, p < .001). In the Post-essays, the Corpus group uses more categories mentioned in the journals than the Traditional group.

### TABLE 6

<table>
<thead>
<tr>
<th>Groups</th>
<th>Categories introduced in the lectures vis-à-vis Categories mentioned in the journals</th>
<th>Categories used in the Post-essays vis-à-vis. Categories mentioned in the journals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional group</td>
<td>54.4%</td>
<td>65.6%</td>
</tr>
<tr>
<td>Corpus group</td>
<td>63.9%</td>
<td>81.6%</td>
</tr>
</tbody>
</table>
2) Interviews

The analysis of the interviews is based on two kinds of data: interview test and interviews. First, in the interview test, the Traditional group shows similar results in the two tests (t(14) = -0.83, p > .05), but the Corpus group has different results (t(14) = -4.33, p < .001) (see Table 7). Second, the interviews focused on learners’ self-explanation and their cognitive strategies. As for analysing the self-explanations, the researcher analysed them via the following steps:

① The frequencies of the three levels of self-explanation
② The three levels of self-explanation versus correct use

<table>
<thead>
<tr>
<th>TABLE 7</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy of the Two Tests</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>Traditional group</td>
</tr>
<tr>
<td>(432/750)*</td>
</tr>
<tr>
<td>Corpus group</td>
</tr>
<tr>
<td>(428/750)</td>
</tr>
</tbody>
</table>
* (Number of correct answers/ total number of test items)

In the first step, the frequencies of the three levels are different in the two groups. The two groups have ‘Obvious explanation (Level C)’ as the highest frequency in their self-explanation (see Table 8). However, the Corpus group has a relatively higher frequency of ‘Unclear explanation (Level B)’. In the second step, learners’ self-explanations are compared with the answers in the interview test. With regard to correct answers, the two groups have similar patterns in the proportion of the three levels except ‘Unclear explanation (Level B)’ (see Table 9). The Corpus group has a relatively higher proportion of Level B than the Traditional group. This higher proportion indicates that the Corpus group learners have more items which they cannot properly explain in association with why they chose their answers. With regard to wrong answers, the two groups show quite different patterns of the three levels. One of the distinctive differences is that the Traditional group has a higher proportion of Level C. However, this level has the lowest proportion among the other two levels in the Corpus group.

<table>
<thead>
<tr>
<th>TABLE 8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Percentages of Self-Explanation in the Two Groups</strong></td>
</tr>
<tr>
<td>Self-explanation</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Traditional group</td>
</tr>
<tr>
<td>Corpus group</td>
</tr>
</tbody>
</table>
*(Raw frequency/ total number of test items)
**TABLE 9**

<table>
<thead>
<tr>
<th>Interview test</th>
<th>Self-explanation</th>
<th>No explanation (Level A)</th>
<th>Unclear explanation (Level B)</th>
<th>Obvious explanation (Level C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional group</td>
<td>Correct</td>
<td>2% (15/750)*</td>
<td>5.6% (42/750)</td>
<td>52% (390/750)</td>
</tr>
<tr>
<td></td>
<td>Wrong</td>
<td>9.9% (74/750)</td>
<td>3.6% (27/750)</td>
<td>26.9% (202/750)</td>
</tr>
<tr>
<td>Corpus group</td>
<td>Correct</td>
<td>2.7% (20/750)</td>
<td>21.6% (162/750)</td>
<td>46.4% (348/750)</td>
</tr>
<tr>
<td></td>
<td>Wrong</td>
<td>11.5% (86/750)</td>
<td>10.1% (76/750)</td>
<td>7.7% (58/750)</td>
</tr>
</tbody>
</table>

* (Raw frequency/ total number of test items)

Through the above two steps of the self-explanation analysis, there is one distinctive feature of learners’ self-explanation. In general, the Corpus group has a higher proportion of ‘Unclear explanation (Level B)’ than the Traditional group. However, it should be differentiated from ‘No explanation (Level A)’. The higher proportion of Level B indicates that the Corpus group learners know the reason ‘why’ they chose particular answers, but cannot accurately verbalise it.

**V. DISCUSSION**

1. Traditional Structuralist Methodology vs. Corpus-based Methodology

The two groups in the case study are based on two different types of methodology: traditional structuralist and corpus-based methodologies. The corpus-based methodology of the Corpus group achieves better results in the test-based analysis and essay-based analysis than the traditional structuralist methodology of the Traditional group. Since the two groups have the same conditions, except for the exercise types in the seminars, the better performance of the Corpus group must be related to the corpus-based exercises. Thus, it is necessary to discuss what factors are involved in the improvement of the Corpus group. Two factors are considered in each methodology: language and organisation. In terms of language, traditional structuralist methodology uses invented language which fails to develop learners’ communicative competence and capacity. Even though some examples extracted from authentic texts are used for learning grammar, this does not contribute to improving the use of determiners. However, corpus-based methodology, based on corpus language, can encourage learners to discover various aspects of language use. One of the distinctive features of this methodology is that corpus language can contribute to presenting not only how determiners are actually used, but also how learners find out useful information from their own learning strategies.
This is also related to ‘organisation’ as the second feature of the two methodologies. The process-oriented perspective adopted by corpus-based methodology focuses on developing learners’ cognitive processes which emphasise both analytical ability and cognitive skills, while the product-oriented perspective adopted by traditional structuralist methodology only emphasises analytical ability. For this reason, the Corpus-group learners have opportunities to develop their cognitive skills and ability to obtain what they need to know using process-based corpus data. Since corpus-based methodology requires learners to analyse large amount of language data, learners are supposed to select and decide on useful information from the data provided. Therefore, the Corpus-group learners are able to develop their strategies on how to analyse massive amounts of data and this can contribute to their language development. By contrast, the quantity of data (examples) given to the Traditional-group learners is less than that for the Corpus-group learners, so they are likely to concentrate on what they should know from the data given to them. For this reason, the Traditional-group learners may not increase their cognitive ability for grammar learning.

2. Consciousness-raising between the Two Groups

The study is based on two different types of exercises based on two different perspectives: process-oriented and product-oriented. The Corpus group, based on a process-oriented perspective, show consistency in their journal descriptions of determiner learning. A higher proportion describes their uncertainty over using determiners (see Table 5). Meanwhile, the Traditional group, based on a product-oriented perspective, have a higher proportion of ‘Clear explanation’, indicating that they have a clear idea of what the rules and notions are. This pattern should imply that the Traditional group learners will achieve a higher level of accuracy in the tests and the essays, but the converse is true. The Corpus group have a higher level of accuracy in these areas. In Section 2.2, we discussed the three levels of consciousness: perception, noticing and understanding. With regard to these three levels of consciousness, the higher proportion of ‘Not very clear’ in the personal journals of the Corpus group can be interpreted as equating to Schmidt’s noticing, since they mentioned questions and expressed doubts over how the rules and notions can be applied in certain situations. With regard to this interpretation, Schmidt’s three levels are not sufficiently clear to classify the learners’ level of consciousness as awareness, so it is very difficult to equate his noticing with uncertainty in the Corpus group. However, the level of consciousness in the Corpus group is higher than Schmidt’s perception and lower than his understanding. A distinctive feature here is that the Corpus group raised their consciousness at the level of noticing on both knowledge of ‘how’ (procedural knowledge) and ‘what’ (declarative knowledge). However, the Traditional group has a higher proportion of ‘Clear explanation’ which can be interpreted as their level of understanding,
and a lower proportion of ‘Not very clear’ in their journals. The Traditional-group learners may have a clear idea of what the rules and notions of the determiner are, rather than how to use them. This pattern for the Traditional group can be linked to the reason why they did not achieve a higher level of accuracy in the tests and the essays. In terms of consciousness-raising, they might be raising their consciousness of declarative knowledge rather than procedural knowledge. Unfortunately, this pattern of consciousness-raising is not likely to contribute to improving their determiner use. These differing patterns of raising the consciousness of procedural knowledge between the two groups may relate to the use of authentic corpus data based on process-oriented methodology. In fact, the Corpus group may have a lot of difficulty in analysing corpus data to discover what they should know. It is probably not easy to apply the determiner rules which are well organised and simplified for pedagogical purposes to much more complicated authentic language in the corpus-based exercises. On the other hand, it might be relatively easy for the Traditional-group learners to apply the determiner rules to simple and clear invented language in their exercises.

The differing patterns of consciousness-raising between the two groups can be explained in terms of the modification of cognitive process. In other words, the consciousness-raising of procedural knowledge in the Corpus group contributes to developing cognitive skills and strategies for determiner use. For example, in the interviews, the Corpus group has a higher proportion of ‘Unclear explanation’ (see Table 8), indicating that they might know the reason why they chose a particular answer in the test, but cannot clearly verbalise it. However, this pattern cannot be interpreted as evidence of unconscious processes. Since the Corpus group has the highest proportion of ‘Obvious explanation’, their proportion of ‘Unclear explanation’ can be explained as their unverbalised conscious processes. If the higher proportion of ‘Unclear explanation’ becomes evidence of unconscious cognitive processing of determiners, one question raised is how to explain the highest proportion of ‘Obvious explanation’ in the Corpus group. The reason for the above phenomenon may relate to learners’ process of changing their cognitive strategy which has been acquired from previous learning. In order to modify their strategy, learners probably need more time to adopt a new strategy acquired through a corpus-based approach. According to Tarone and Yule (1989, p. 153), learners pass through a ‘plateau stage’ where they are struggling to accept new content and methodologies. For this reason, their improvement is not linear, but is stagnant at a certain stage (the plateau stage) for some time. More specifically, Johnson (2001) discusses learners’ automatisation which refers to fluent skills in language use needing less conscious attention. According to Johnson, learners’ fluent language use — in other words automatisation — is based on cognitive transfer from declarative knowledge to procedural knowledge. This transfer might have different patterns according to individual
differences. For example, some learners require a longer period in the plateau stage to transfer declarative knowledge to procedural knowledge. However, all learners need this kind of transfer to achieve automatisation. Even though the experimental group achieved a certain level of accuracy in their use of the target grammar, they were likely to need sufficient time and input to acquire a higher level of consciousness of procedural knowledge (i.e., understanding) for automatisation. The relationship between a corpus-based approach and learners’ cognitive development is neither simple nor straightforward. There might be many other factors such as instruction and individual difference, which Schmidt (1990) also presents, and which we need to consider seriously. One of the most distinctive contributions from a cognitive perspective is that a corpus-based approach, focusing on a process-oriented perspective, is much more powerful in that learners can notice how language is used. According to Skehan (1998), language learning is considered as process rather than product, so learners can benefit from a process-oriented approach, in other words a corpus-based approach to raise their consciousness of language use.

Consequently, the two groups show a clear difference in terms of consciousness-raising. The Corpus group raised their consciousness both of knowledge of how determiners are used and knowledge of what determiners are. However, the Traditional group raised their consciousness of knowledge of what determiners are rather than how determiners are used. Even though the consciousness of the Corpus group is at the level of noticing, it is quite different from that of the Traditional group.

3. Limitations of the Study

Some limitations can be described in terms of four factors. The first limitation is a problem of concerning the validity and reliability of the learners’ own descriptions, such as their personal journals. Learners might have described what the researcher wanted to know rather than what they really learnt. Furthermore, Korean learners might have a lot of difficulty in describing their target-language learning. The second limitation concerns learners’ psychological factors with regard to new materials (i.e., corpus data) and process-oriented methodology. In psychology, it has been variously discussed under the names as the Hawthorne effect, referring to the effect from the fact that people are taking part in an experiment (Chiesa & Hobbs, 2008), and the Halo effect, referring to the effect caused by the novelty of the treatment in an experiment (Draper, 2008). For example, the better results of the Corpus group might have been related to learners’ attention being drawn to the novelty of a process-oriented methodology or materials. One of the solutions to the influence of learners’ psychological factors is to investigate their improvement over a longer period (Harrison, 2002). This also relates directly to the third limitation; this is
that the three-month research period which might not have been enough to investigate learners’ consistent patterns of improvement (Cohen, Manion, & Morrison, 2000). In particular, highly-motivated learners may have better results over a short period (Harrison, 2002). However, in long-term and classroom-based experiments, learners’ dropout rates might increase (Cohen, Manion, & Morrison, 2000). The fourth and final limitation is the sample size of the case study which could have had an effect on the results of the two groups. For this reason, their distinctive patterns may not represent the general patterns of all EFL learners. However, sample sizes in this kind of longitudinal educational research are largely dependent on course enrolment figures, over which the researchers have no control. For this reason, the dropout rate in a longer experiment might mean that they still end up with similarly small sizes at the end of the day.

VI. CONCLUSION

This case study of Korean learners’ acquisition of determiners shows that a corpus-based approach, focusing on cognitive process, contributes to developing the interlanguage system. One of the most distinctive findings of this study is that a corpus-based approach became much more powerful when it focused on raising learner consciousness at the level of noticing. In particular, the learners showed significant improvement in increasing procedural knowledge rather than declarative knowledge. For this reason, the most important contribution of the corpus-based approach is that learners can develop their cognitive skills and ability through noticing how language is actually used on the basis of their own analysis of corpus data. The researcher believes that this finding can be a sensible way to apply corpus data to language pedagogy from a balanced perspective of both theoretical descriptions and pedagogical considerations. Even though this methodology is focused on a limited area of language learning (grammar learning) and has some limitations, it may have many possibilities and much potential to be applied to other areas (e.g., reading, listening, testing, and so on).

With regard to future research, the researcher would like to emphasise the use of process-oriented methodology in a computer setting. As a result of advanced computer technology, the analysis of corpus data is now the job not only of researchers and teachers, but also of learners too. For this reason, corpus-based approaches in a computer setting are likely to have more pedagogical potential and possibilities than other methodologies in traditional settings, such as paper-based corpus-data formats. Since a computer-based corpus-data format helps learners to construct their own learning strategies, they may have more chance to develop their cognitive skills for language learning. From this viewpoint, it is necessary to investigate whether or not a process-based corpus-data computer format is
more efficient than a process-based paper format. One of the key points here is how learners can make use of concordancers in their language learning if a process-based corpus-data computer format is better than a process-based paper format in terms of cognitive development. In particular, learners’ use of concordancers should focus on how they contribute to raising their consciousness of both types of knowledge (procedural and declarative). Therefore, as a future research question, a computer-based format for corpus approaches needs to be investigated with the focus being on learners’ consciousness-raising and how it connects to their subsequent learning.

REFERENCES

Baars, B. (1986). What is a theory of consciousness a theory of? The search for critical constraints on theory. Imagination, Cognition and Personality, 6, 3-23.
London: Routledge.
(Eds.), *The communicative approach to language teaching* (pp. 5-26). Oxford: Oxford University Press.


acquisition. Cambridge: Cambridge University Press.


York: Newbury House.

Raines, A. (1985). What unskilled ESL writers do as the write: A classroom study of
composing. TESOL Quarterly, 19, 229-258.

Cambridge University Press.

piece to the puzzle. Studies in Second Language Acquisition, 21, 511-556.

129-158.


Schmidt, R. (1994). Deconstructing consciousness in search in useful definitions for

attention and awareness in learning. In R. Schmidt (Ed.), Attention & awareness in
foreign language learning (pp. 12-55). Honolulu: Hawai’i University Press.

instruction (pp. 3-32). Cambridge: Cambridge University Press.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second

University Press.

Shekary, M., & Tahririan, M. H. (2006). Negotiation of meaning and noticing in text-based

London: Collins.

Fligelstone, T. McEnery & G. Knowles (Eds.), Teaching and language corpora (pp.

Linguistics, 17(1), 38-62.


Applicable levels: higher education
Key words: Consciousness-raising, noticing, self-discovery learning, process-oriented perspective.

Shin Chul Hong
Department of English Language and Literature,
Dong-A University
840 Hadan 2-dong, Saha-gu,
Busan 604-714, Korea
Tel: 051-921-4895/ H.P.: 010-4127-4895
Email: shinchul@yahoo.com