Validating the Lexical Network of *Take*: A Corpus-based Study*

Hyung-Sun Kim & Baeg-seung Kim**

(Chonnam National University)

Kim, Hyung-Sun & Kim, Baeg-seung. 2011. Validating the Lexical Network of *Take*: A Corpus-based Study. *Linguistic Research* 28(1), 117-142. The lexical network accounts for the relationship among a number of senses of a polysemous lexeme in the cognitive semantics, where each sense is connected to another by links. This study is an attempt to validate the lexical network of *take* proposed by Norvig and Lakoff (1987) via newly compiled corpora from an English-only local journal, and the outcome of a sentence-elicitation experiment. Under the assumption that the intuitive model of the lexical network theory would hold more explanatory adequacy in the native use of language than in the non-native one, the data were arranged into two comparable sets accordingly. While the model does not meet linguistic validation by means of measuring the sense frequency, the result of the saliency experiment was in concurrence with the model for the most central sense, take-1 (*grab*). There was no sign that the "lexical network theory" (LNT) applies differently between the native and non-native use of *take*. The result of this study which does not concur with the previous studies on the LNT model of *take* can be attributed to the different decoding of sense distinction in part; thus further clarification is needed to define the nature of centrality. **(Chonnam National University)**

Key Words polyseme, sense, lexical network theory (LNT), frequency, saliency

1. Introduction

The prospect of adopting a data-driven methodology such as the corpus approach has been discussed in a diverse range of linguistics (Aston, 2000; Granger & Tribble, 1998; Johns, 1991). Many studies in this vein have focused on its pedagogical benefit of obtaining 1) enhanced description of a language suggested by the native speakers' intuition; 2) a better language learning methodology discovered from various patterns and uses in native corpora; and 3) information from learner corpora

in the areas of difficulty that learners experience (Johns, 1991; Kim, 2008; Na & Kim, 2010; Oh, 2007). On the other hand, there have been studies in the corpus framework that center on pure linguistic descriptions such as the semantic path from modality to aspect (Aijmer, 2004), lexico-grammar (Nevalainen, 2004), subject clitics in English (Brinton, 2004), phraseology (Altenberg, 2001), and the differentiation between *actually* and *in fact* (Oh, 2000).

The tradition of 'contrastive linguistics,' especially in the form of a comparative analysis of native speaker/expert vs. language learner/novice production of language may be found in both realms. There have been extensive studies with relatively more pedagogical implications covering various topics of the use of adverbials (Altenberg & Tapper, 1998), phrasal verbs (Lam & Hung, 1998), and vocabulary (Altenberg & Granger, 2001). Conversely, with focus on the explanation of linguistic phenomena, Kim and Heine's (2011) comparative study reported a suggestive piece of evidence regarding grammaticalization that the intermediate forms of the existential *there* were mostly found in the native use of *there* and seldom in its non-native use. It is the latter trend of research that the current study takes in its attempt to verify the lexical network model of the polysemous lexeme, *take*.

The lexical network is concerned with the relationship among a number of senses of a polysemous lexeme within cognitive semantics. The lexical network theory (LNT) claims that the semantic part of a lexical item is best explicated as a network of lexical senses, in which the use of a word based on its subcategorization frame may be an instance of a previously established sense of the word (Norvig, 1989). A polyseme of multiple senses "is structured in terms of a network such that each sense is a minimal variant of some other sense," (Norvig & Lakoff, 1987, p.197) and a link functions on one sense to produce its minimal variant. As traditional studies in this framework turn to linguists' introspections rather than being grounded on empirical evidence, much of the recent research has focused on the validity of the theoretical models employing the corpus approach (Croft, 1998; Gries, 2006; Nordlund, 2008), and has tried to build empirical models (Gilquin, 2008a) on the basis of 'ready-made' corpora. There is no conclusive evidence yet, however, about how the intuitive model LNT may be properly verified, and what kind of, and/or how many sources should be examined to reach adequate "converging evidence" (Rice, 1996, p. 162).

Given its restrictions, this study attempts to validate Norvig and Lakoff's (1987)

intuitive model of a polyseme *take* under the LNT via the newly compiled sets of native and non-native speakers' corpora from a native English speaker, and non-native speaker produced magazine published in Korea. As the model was established on the grounds of the linguists' intuition (Gilquin, 2008a), we expected that there would be a better chance for the native-speakers' (NS) contributions to draw on the model than for the non-native speakers' (NNS) writings. A sentence-elicitation experiment was also performed to investigate the applicability of the model in terms of cognitive saliency. Overall, the following research questions have guided this study:

1. Does the LNT model of *take* stand up to the NS and NNS corpus validation examining its linguistic reality and/or to the experimental validation examining its cognitive reality?
2. Is there any difference in the sense pattern of *take* between the NS and NNS corpora and/or between the NS and NNS experimental writings?
3. Is the intuitive LNT model better applied to the native uses than to the non-native uses of *take*?

2. Background of the Study

2.1 *Take* in Lexical Network Theory

Since Brugman (1984) applied the semantic relatedness criterion to explicate the internal semantic structure of polysemes, many studies have developed the theory of lexical networks (Amagawa, 1996; Brugman & Lakoff, 1988; Lakoff, 1987; Norvig, 1986), in which a polyseme is considered as a network which is constructed of minimally differing senses connected with links. Norvig and Lakoff (1987) established six link types proposing two additional types (see e & f) in addition to the existing four types in the theory of link types (pp. 197-198).

a. Image-schema transformation links: These are links given by natural relationships among image-schemas.
b. Metaphoric links: These are links that are established by metaphoric mappings that exist independently of the given lexical items.

c. Metonymic links: These are links that are established by metonymic mappings that exist independently of the given lexical items.
d. Frame addition links: Here the minimal difference is the addition of a frame.
e. Semantic role differentiation links: There are cases where one sense identifies two semantic roles, and a minimally related sense distinguishes those roles.
f. Profile shift: These are cases where what is back-grounded in one sense is fore-grounded in a minimally related sense.

On the basis of the link types and the semantic roles listed in Table 1, Norvig and Lakoff (1987) listed seven distinctive senses of *take* as illustrated by the examples in Table 2. Both tables are adapted from the research.

Table 1. Semantic roles used in the lexical network of *Take*

	Semantic role	Description
	Agent (A)	active actor or causer of an action
	Source (S)	initially has the patient
Participant	Recipient (R)	receives the patient
	Patient (P)	object acted upon or affected by the agent
	Instrument (I)	used to transport the patient
Setting	Origin (O)	location where patient started out
	Destination (D)	location where patient ends up

The central sense of *take* (take-1) is hypothesized to be the sense in *The baby took the toy from its mother* in Norvig and Lakoff (1987), from which other senses are ramified through minimal variation links. In take-1, the agent uses his/her hand as the instrument of movement, and the patient is a relatively small physical object. As the agent is the recipient, it does not need to move, and actually receives the patient as the typical consequence of the taking.

Take-2 (*take patient to recipient*) is a minimal variant of take-1 (*grab*). The minimal difference that produced take-2 from take-1 is that the agent is not the recipient any longer. So, the agent in take-2 needs to move to the destination of the patient where the recipient is. Take-3 (*take patient to destination*) is different from

Table 2. Senses of *Take* in the lexical network

Sense	Representation	Linked to	Minimal difference	Example
Take-1	Grab		Default	The baby took the toy from its mother.
Take-2	Take P to R	T-1	A ≠ R	The messenger took the book to Mary.
Take-3	Take P to D	T-2	Settings are profiled.	I took a suitcase to Bali.
Take-4	Take action at P	T-2	The metaphor that applying force is transferring an object.	I took a punch at him.
Take-5	Take action from A	T-4	Result is profiled.	I took a punch from him.
Take-6	Take to the movies	T-3	Add "going-to-D" Schema.	John took Mary to the movies.
Take-7	Take a glance at	T-1	The metaphor that perceiving is receiving	John took a whiff of this roast beef.

take-2 in that the setting is profiled (fore-grounded), and consequently the source and recipient are back-grounded. Take-4 (*take action at patient*) is a metaphorical extension of take-2 which is a transfer of a patient to a recipient by an agent. In the metaphor, a quick and forceful action is understood as an object delivered by the agent to the recipient. Take-5 (*take action from agent*) minimally differs from take-4 in that the result is profiled, and the recipient is the subject of the sentence since the recipient is profiled while the agent is back-grounded. Take-6 (*take to the movies*) is a minimal variant of take-3, where the metonymy in which the occurrence of "going-to-D" stands for doing the relevant activity. Contrary to the sense of *take* in *Max took Sadie to the doctor* (take-2), the sense in *Max took Sadie to the theater* (take-6) entails the agent's participation in the activity concerned.

Lastly, take-7 (*take a glance at*) is a minimal variant of take-1 in that it is a metaphorical use of take-1, which is a special case of the very general metaphor that perceiving is receiving. The patient is not limited to a physical object but extended to a sensory impression through the metaphor. Figure 1 is a schematic representation

of what has been described so far, the network of *take* with the link-types. The Norvig and Lakoff's (1987) diagram has been adapted to be reflective of its layered structure. The degree of shading signals the extent of centrality of the senses; as the number of links increases, the centrality of the sense decreases.
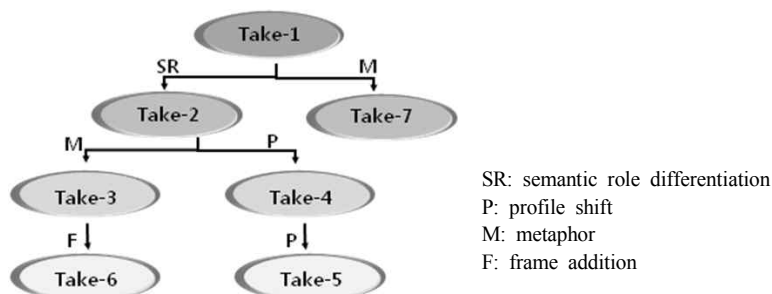
SR: semantic role differentiation
P: profile shift
M: metaphor
F: frame addition

Figure 1. Lexical network of *Take*

2.2 Prototypicality: Frequency and/or Saliency

Within the LNT model, the different senses of a polyseme are organized by means of positioning the prototypical sense in the center of the radial network. The remainder move farther from the center with increasing degree of peripherality (Brugman and Lakoff, 1988). The concept of prototypicality has been used differently by linguists: 1) Viberg (2002) described the prototype of a word as one 'central' sense from which the other senses may be derived in a principled way; 2) Williams (1992) associated prototypicality with 'saliency' by stating that the prototypical or central sense carries the greatest amount of cognitive saliency; and 3) to Stubbs (2005), the prototype is the most 'frequent' item in its actual use, which is a typical stance in corpus linguistics.

Gilquin's (2008a) bearing on prototypicality is inclined to the first among those three uses of the concept; she identified the centrality, which was the determinant of the default state of sense in Norvig and Lakoff (1987), with the prototypicality. Yet, Gilquin (2008b) rightly attended to the fact that saliency is in the mind and frequency in language, although those two concepts of saliency and frequency are often considered to work in tandem with each other since mind and language are usually associated. The research design of analyzing corpora, running parallel with a

sentence-elicitation experiment, which was originally implemented by Gilquin (2008a; b), was motivated by this observation: the former was chosen to measure frequency of a sense and the latter to see saliency, and the results were intended to be compared to verify their association.

2.3 Previous Studies on *Take* in Lexical Network Theory

Previous studies have attempted either to enhance or to verify the explanatory adequacy of the *take* model under the lexical network theory which was the framework of Norvig and Lakoff (1987). For example, Amagawa (1996) suggested schematic networks starting from the observation that the senses of the verb *take* have either the schema of *take x from y* or the schema of *take x to y*. She argued that the senses of the same schema form a network in one of two domains: spatial or psychological. In doing so, she pointed out three problems in Norvig and Lakoff (1987): 1) it is not clear how take-5 is a minimal variant of take-4 in that there is a semantic-role shift occurring between the two senses in addition to the profile shift; 2) it remains unclear how take-7, linked to take-1 in reliance on the general metaphor "perceiving is receiving" could deal with an expression such as **take a listen to a radio* which is not well-formed English; and 3) some syntactically equivalent expressions to take-7 such as *take a stroll* and *take a walk* were excluded despite the fact that they make perfect sense.

More careful examination of Norvig and Lakoff (1987), however, indicates that the answers to those problems raised by Amagawa (1996) had been already provided: 1) "When we say that sense A is a minimal variant of sense B, we mean that, even though A and B may differ in many ways, all their differences boil down to a single difference from which the others can be predicted" (p. 197); 2) "This metaphor interacts with a general process in English by which a perfective action⋯is nominalized⋯Why doesn't **take a stare* occur [whereas *take a look* does]? Because the nominalization process only applies to perfective actions, and *stare* is inherently imperfective," (p. 204) and so is *listen*; and 3) their study centered not on the syntax of *take* but on the "semantic relatedness criterion" (p. 195) of the verb, and despite the syntactic similarity between *take a look* and *take a walk*, they do not share the same sense of *take*.

In an attempt to verify the analysis of *take* in Norvig and Lakoff (1987), Gilquin

(2008a) utilized the data obtained from the Frown corpus for the written samples and the Switchboard corpus for the spoken data. The research also carried out sentence production experiments of forty native speakers of English to measure the degree of perceptual saliency of the senses of *take*. The corpus methodology and experiment tested the *take* model against empirical data, for its linguistic and cognitive reality respectively. Based on the result of the quantitative analysis, she concluded that the LNT model of *take* in Norvig and Lakoff (1987) is not realistic in both linguistic and cognitive perspectives in that the model did not correspond with natural language production. A point to note is that while she followed the overall sense distinction in Norvig and Lakoff (1987), which is the improved version of Norvig (1986), she did not fully adhere to the distinction. Specifically for take-4 (*take a punch at someone*), which involves "a quick, forceful action" (Norvig & Lakoff, 1987, p. 201), she argued that there are very few nouns which could take the place of "punch" and decided to extend the category to "any type of action" for this somewhat arbitrary reason (this take-4 sense issue will be further discussed in Section 4). Despite the impeccable research design, her modification of the existing sense distinction gave rise to the necessity of follow-up research which would prove or disprove the conclusion of her research.

3. Method

Motivated by the assumption that the lexical network model of *take*, which was claimed to be intuitive, would hold more explanatory adequacy in native uses of *take* than in its non-native uses, this study has chosen the comparative research framework and collected data accordingly. Within the framework, the data analysis is aimed at obtaining the central or more central sense among the seven senses of the lexeme in each set of data. The highest frequency as evidenced by corpus data has been equated with prototypicality (Stubbs, 2005) which, according to Gilquin (2008a), may be taken to correspond to the centrality (or default) introduced by Norvig and Lakoff (1987). On the other hand, the experiment asking participants for the item that comes first to mind has been considered to be a standard test to uncover the prototype, as an equivalent of saliency of a category (Radden, 1992; Ungerer & Schmid, 1996). With this rationale as the basis, the corpus methodology

was chosen for linguistic validation, i.e. to check if the model holds natural language occurrences in terms of the token frequency. The experiment for cognitive validation was to see if the model adequately reflected the cognitive saliency as evinced in the sentences which draw the most salient sense of *take* from the participants' mind.

3.1 Data: Corpus Construction and Elicitation Experiment

When studies of this type resort to 'ready-made' corpora, it is difficult to secure the proper level of compatibility of two sets of data. Therefore, researchers often end up controlling the genres and/or registers of the corpora, in order to guarantee a minimal level of compatibility of the data. We desired to see if the intuitive model would be better reflected in native language use than in its non-native counterpart, while maintaining the comparability of the two sets of corpus; the comparison of learner language in general against that of professional writers, journalist, and so on hardly seemed compatible (McEnery & Kifle, 2002). The first set of the contrasting data was collected from a series of issues of an exceptional magazine that both native speakers (NS) and compatibly advanced level non-native speakers (NNS) contribute to, which would assure adequate comparability. The monthly journal is published in English by an international center, supported by city administrative authorities. The NS writers are mostly English teachers at schools or institutions, and the NNS writers are professors, journalists, EFL teachers, and university students. The former are from the United States of America, England, Australia, and Canada, etc., and the latter are mostly Koreans, but some are from other countries. For the purpose of the study, only the contributions of Koreans were collected for the NNS data to minimize further variables. The articles compiled to construct the corpora were downloaded from the weblog based on phpBB (PHP Bulletin Board) designed for the purpose of the journal publication by the international center, which provided consent for using the data for research purposes. When clicking an issue, every article of the journal appears, and is allocated to a section of the web page, and the section shows the writing from the first contribution to the final published version with interim revised versions, if any. It also shows that the correspondence between the contributor and the editor in charge of the column. The originally contributed versions, in order to avoid any NS editors' lexical intervention, were collected from each section of an issue one by one. As the purpose of the journal is not only

sharing information about living in Korea but also discussing humanitarian issues worldwide, the topics were literature, politics, economy, art, human rights, and travel, etc. A thorough process of compiling the appropriate articles from the issues of four years from November 2006 to October 2010 constructed corpora that reached a size of 266,025 words (269 articles) for the NS; and 117,527 words (144 articles) for the NNS.

The second part of the research data was the result of sentence production experiments to elicit the most salient sense of *take* in the participants' mind. Adapted from Gilquin's (2008a) experimental method, the working theory was that participants would use the most central sense among many in their minds when asked to produce the very first sentence they could think of containing the verb. The 27 native speakers who volunteered to participate in the research were English teachers from a university and from two language institutes. The 55 non-native speaker (Korean) participants were juniors and seniors in the department of English Language and Literature and the department of English Education of a university. The informed consent was obtained from all parties. For the experiment, five sheets of paper were provided to participants, with a word on the top of each paper (one with the verb *take*, and the others with four random distracters: *make*, *there*, *some*, and *get*) with the instruction to "Write down the very first sentence you can think of containing the word presented."

3.2 Data Analysis

In analyzing the NS and NNS corpus constructed, NLPtools (Lee, 2007) was utilized to get the frequency and to extract the concordance lines of the lemma *take* and its grammatical forms (*taken*, *takes*, *taking*, and *took*) from each of the two corpora. Then the concordance outputs were carefully examined to sort out the cases of the seven senses as presented in Table 2 (in Subsection 2.1) from all the occurrences of the target word. There were cases that could obviously be categorized into one of the seven senses, but many others were found to require close studies and discussions between researchers based on the sense description of Norvig and Lakoff (1987) and the relevant papers reviewed in Subsection 2.2. For example, there were some cases of take-2 (*take patient to recipient*) and take-3 (*take patient to destination*) that were confused with take-6 (*take to the movies*) for their subtle

differences. The metonymic limitation of "going-to-D schema" (Norvig & Lakoff, 1987, p. 203) was exclusively applied for take-6; the difficulty was that in take-6 the agent should take part in the activity that normally takes place at the destination, i.e. seeing a movie. Therefore, on the basis of either the explicit mention or contextual implication, the cases in which the agent does not intend to be involved in the activity were allocated to take-2 or take-3. Note the following examples:

(1) I arrived in Seoul on a steamy July night in 2007, and my host family ─a couple from the school where I was to work─offered to [take] me out for a late dinner. (take-6)

(2) It is also recommended that all interested in adopting should [take] the pet to the vet to get a checkup. (take-2)

(3) So whenever my aunt [took] me to the beauty salon, all the ajummas came to me and said 'Look at that cute little girl!' (take-3)

(4) Once you are under the rainbow sign, you will see some stairs going into the building that will [take] you to the second-floor coffee shop. (take-3)

The next step was arranging the two sets of data (NS and NNS) to be comparable with each other and to the previous studies: calculating the ratio of each sense type against the total occurrence, and making clusters of senses which demonstrate similar frequencies. Some sample sentences were extracted for the qualitative discussion as well.

The similar procedure was taken with the result of the elicitation experiment: each sentence produced by the NS and NNS participants containing *take* was examined and classified according to the sense types of *take*; and the seven senses were grouped into clusters according to the number of occurrences, so that the results may be comparable between the NS and NNS use and to those of the previous studies.

4. Results and Discussion

In this section, the result of empirical (linguistic and cognitive) validation of the

Norvig and Lakoff's (1987) network is provided, and discussed in comparison to the model of Gilquin (2008a). Before doing so, clarification of the sense criterion adopted in this study, which is differentiated from the one adopted in Gilquin (2008a), is required.

4.1 Some Sense Distinction

As pointed out in Subsection 2.2, Gilquin (2008a) has proposed some modifications of the existing sense distinction, in which a critical change was made for the sense criterion of take-4. According to the original model, take-4 (*take action at patient*) as in *I took a punch at him* is linked to take-2 (*take patient to recipient*), which involves a patient that is transformed to a "quick, forceful action" through "the metaphor that applying force is transferring an object." Those two senses are linked by a minimal difference which is the metaphor.

The discussion arose because 1) the verb usage *take a punch* is very similar to that of light verbs, in which the sense of a verb becomes a "delexical or light-verb sense" (p. 28), and 2) very few nouns may occur in the place of *punch* (Gilquin, 2008a). Therefore, she suggested that take-4 should include all the light verb usage of *take*, which would involve any type of actions or "take a V" (Amagawa, 1996, p. 23), including cases such as *take a walk*, or *take a picture*. The amendment, however, creates a more serious problem. In applying the formula *take action at patient*, the above-mentioned cases predicate that the action does not affect a patient different from the agent, in other words, the counterpart is missing. It is hardly acceptable considering that the link between take-2 and take-4 is a single minimal variation, i.e. the metaphoric link, from which the absence of a patient cannot be predicted. In addition, as for the issue of the scarcity of take-4 cases that she raised, no concern is necessary since, as declared by herself, "the seven senses proposed by Norvig and Lakoff (1987) do not cover all the uses of *take* as exemplified in the data" (Gilquin, 2008a, p. 29), but account for a subset of the *take* usages after all.

Thus, this paper stuck to the sense criterion of take-4 as originally stated by Norvig and Lakoff (1987), and only the cases where the agent affects the patient were taken to belong to the take-4 category, as exemplified in (5)-(9). As the light-verb sense is overall the most common use of polysemous verbs, it was also the case of *take*; the number of the light-verb uses of *take* was 62 in the NS corpus,

and 40 in the NNS corpus, occupying 13.96% and 20.94% of all occurrences of *take* respectively, and overall 16.06%. Table 3 lists the light-verb cases identified in the corpora examined in this study, which are excluded from the take-4 category. Since Gilquin (2008a) counted the light-verb use as part of take-4, it is a matter of course that take-4 was the most central sense in both the written and spoken corpus in her research.

Table 3. Light-verb uses of *Take* in the NS and NNS corpus

	Light-verb uses of *take*	Occurrences
NS	take example, take photos, take a walk, take note of, take a pick, take care of, take notes, take advantage of, take vows, take a drink, take test, take a trip, take notice of, take a step back, take steps (=walk), take control of, take a nap, take a break, take a stroll, take a trip, take a stroll, take trips, take approach, take a turn, take prediction, take pictures, take a poll, take a toll, take the plunge, take a shower, take pause	62 (13.96%)
NNS	take a rest, take an approach, take care of, take a walk, take a rest, take exam, take picture of, take advantage of, take turn, take trips, take refuge in, take turn, take charge of, take sunbath, take snap shots, take step (walk), take notice of	40 (20.94%)

(5) ... and then goes on to [take] a swipe at the teachers union. (take-4, NS Corpus)

(6) Thailand must [take] appropriate measures to ensure that Hmong children are protected against all forms of discrimination or punishment on the basis of their ethnic and national status. (take-4, NS Corpus)

(7) Fate, again, decided to [take] one last shot at me and delay my visa. (take-4, NS Corpus)

(8) They need our consideration, and we must [take] action for them now. (take-4, NNS Corpus)

(9) It was on the verge of extinction unless some drastic measure was [taken] to save it. (take-4, NNS Corpus)

4.2 Linguistic Validation: Sense Frequency

The size of the corpora used in this study reached 383,552 words (266,025 words for the NS corpus and 117,527 words for the NNS corpus) as detailed in Table 4. Total 635 concordance lines of *take* were extracted from the NS and NNS corpus, 444 and 191 lines respectively, among which the seven senses of Norvig and Lakoff (1987) were identified. As it appears in Table 5, the lexical network model covers only 16.67% (NS) and 19.32% (NNS) of the total occurrence of *take* (overall, 17.17%) in the current study while 34.45% was covered in the Gilguin's (2008a) data. However, considering the fact that she counted the light-verb use of *take* into the take-4 category, it can be said that the coverage of the model that has reported in these two different studies is strikingly similar.[1]

Table 4. Corpus size and the frequency of *Take*

Corpus	Word counts	Tokens of *take*	*Take* within N&L* network
NS	266,025	444	74 (16.67%)
NNS	117,527	191	35 (19.32%)

* Norvig and Lakoff (1987)

Table 5. Proportions of the seven senses in NS and NNS corpus

Sense	NS		NNS		Chi-square value
Take-1	11	2.5%	6	3.1%	0.22
Take-2	2	0.5%	0	0.0%	0.86
Take-3	43	9.7%	9	4.7%	4.03
Take-4	10	2.3%	8	4.2%	1.77
Take-5	0	0.0%	6	3.1%	13.95***
Take-6	3	0.7%	2	1.0%	0.23
Take-7	5	1.1%	4	2.1%	0.88
Total	74	16.7%	35	18.3%	

*** $p < .001$

---

[1] Gilquin (2008a) reported that the seven senses in Norvig and Lakoff (1987) covered 34.45 % in her corpus data, which included the light-verb uses of *take*. In this study, the seven senses reached overall 17.17%, excluding the light-verb sense of *take* (16.06%). It makes 33.23% as per the Gilquin's (2008a) modified sense distinction.
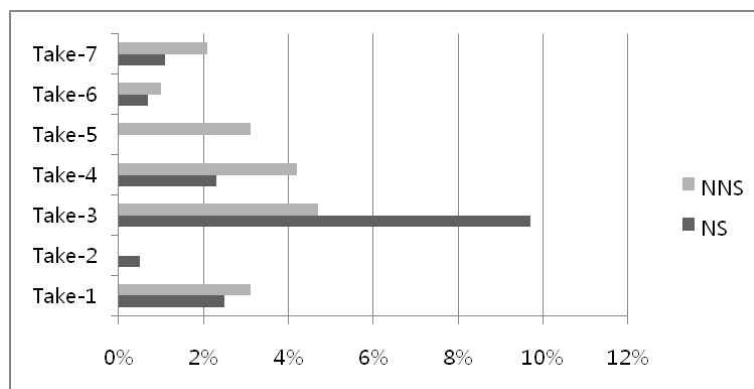
Figure 2. Proportions of the seven senses in NS and NNS corpus

Table 5 and Figure 2 show the proportion of each sense of *take* in the NS and NNS corpus. The most frequent, thus central sense in both corpora was take-3 (*take patient to destination*), which amounted to 9.7% and 4.7% in the NS and NNS corpus respectively. Take-3 is considered to be peripheral as much as two "links" to the central sense take-1 (*grab*) in the LNT model. Yet the corpus data indicate that the most central sense in terms of the usage frequency is not "grabbing" but "carrying." The examples of take-3 in both corpora are shown in (10)-(11).

(10)   After another ten drinks, you can design another mug to [take] home. (take-3, NS Corpus)

(11)   Buses no. 64 and 38 will [take] you to the venue. (take-3, NNS Corpus)

Another sense which showed relatively high frequency was take-4 (*take action at patient*), appearing 2.3 % in the NS corpus and 4.2% in the NNS corpus as shown in (5)-(9). The occurrence of take-1 which is the default and central sense in the LNT model occupied just 2.5% in the NS corpus, and 3.1% in the NNS corpus. Take-2 (*take patient to recipient*), which is considered to be the second most central sense in the LNT model, was rarely shown in both corpora, indicating 0.5% only in the NS corpus. These senses are exemplified in (12)-(13).

(12)   ... but the teacher constantly warned them never to [take] anything out of his basket, because it's poison to children. (take-1, NS Corpus)

(13)   Also have documents translated before [taking] them to a lawyer to save you money and time. (take-2, NS Corpus)

Interestingly, take-5 (*take action from agent*) never occurred in the NS corpus, but the NNS corpus showed 6 cases (3.1%), which will be discussed in the following subsection in detail. The two senses left, namely take-6 (*take to the movies*) and take-7 (*take a glance at*) appeared to be very minimal. Examples are presented as follows:

(14)   If the patient's cornea's tissue is thick enough to remove, it is possible to [take] the surgery again. (take-5, NNS Corpus)

(15)   Later in the interview the mother said that she doesn't like visiting zoos in general but that she wanted to [take] her kids to see the animals. (take-6, NS Corpus)

(16)   You can [take] a peek here: http://free-ali.blogspot.com/ (take-7, NS Corpus)

(17)   You can [take] a look at the short movie about his life that includes English subtitles. (take-7, NNS Corpus)

Overall, the linguistic evidence in the form of the sense frequency in natural use of language did not correspond to the lexical network proposed by Norvig and Lakoff (1987). Figure 3 represents the clusters of senses which may be grouped together according to their frequency level. The degree of shading signals the extent of centrality of the senses which belong to the cluster; no shading indicates no occurrence of the sense.
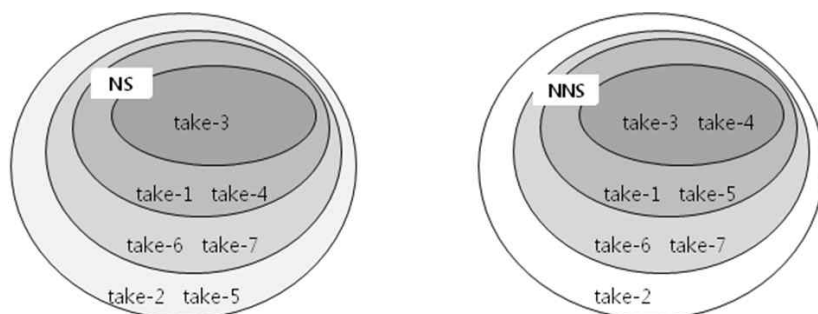

Figure 3. *Take*-network in the NS and NNS corpus

While the networks produced by the NS and NNS corpus were not distinct from each other, the use of take-5, as it appears in the two corpora, turned out to be significantly different (X^2 = 13.95, *df* = 1, *p* < .001), as Table 6 indicates. In the following subsection, the difference between the native and non-native use of *take* in the corpus data is discussed.

4.3 Overuse of Prefabricated Chunks in the NNS Corpus

Further investigation into the concordance lines of *take* showed that the verb in the NNS corpus often combined with same words. For example, six occurrences of take-1 were identified in the NNS corpus, and five cases out of the six involved *out* or *out of*, as shown in (18)-(22).

(18) Always [take] the plug out if you're not using the appliance.
(19) But we did not have enough money to pay the fine so we showed our pockets and [took] out all the money we had.
(20) Rodney knew that it would be helpful to cure her if he [took] her out of the shelter first.
(21) The Creator made the first man of clay and baked him in the oven but [took] him out too soon.
(22) We [took] out all the money we had and started to negotiate a deal, despite the communication barrier.

The same tendency was also noted in the case of take-4 where *take action(s)/measure(s)* were found six times in eight occurrences.

(23) They need our consideration, and we must [take] action for them now. (=8)
(24) The government started [taking] actions to move the dog meat eateries from the main thoroughfares to inconspicuous back alleys in Seoul.
(25) It [took] immediate action to bring up the North's invasion with the UN Security Council.
(26) Then US President Harry Truman also [took] decisive action in the Security Council through the US delegation to save Korea.

(27) University authorities should have [taken] through campus media some preparatory measures regarding cultural differences that their home students may encounter soon.

(28) It was on the verge of extinction unless some drastic measure was [taken] to save it. (=9)

Another noteworthy point is that the sense take-5 was not found in the NS corpus at all but six times in the NNS corpus. The difference, however, is undermined by a close look into the cases. The five occurrences of them were all *take surgery*,[2] and the rest was *take treatment*, both of them conveying the same meaning.

(29) If the patient's cornea's tissue is thick enough to remove, it is possible to [take] the surgery again. (=14)

(30) If they didn't [take] the surgery, the biggest reasons are the patients are afraid, there's less desire to get an operation, and the inconsistency of the eye condition after operation.

(31) If LASIK is safe and good, why do many ophthalmologists not want to [take] the surgery?

(32) [Taking] surgery can be used in case some treatment such as artificial tears is not working.

(33) And most ophthalmologists don't recommend [taking] the surgery if patient does not have a big problem to wear glasses and contact lens.

(34) Hence many people who wear viable alternatives such as corrective eyeglasses or contact lenses for a while are considering [taking] the surgical treatment for vision problems.

As there was no case of take-5 in the NS corpus, it was decided to identify any 'verb/verbal equivalent + *surgery/treatment* (meaning medical treatment)' constructions in order to find any repertoire of verbs that native speakers would use

---

2 While the phrase *take surgery* may easily be found, via www.google.com research for example, a native speaking professor of an English Language and Literature Department commented that one can 'take treatment' but does not 'take surgery'; one 'takes a pill,' but 'has' or 'endures' surgery, thus the corpus is poor in this case.

in collocation with the words. The word *surgery* was not found, but the word *treatment* was used together with various verbs as in *[receive] continuous treatment, [seeking] treatment, [access to] treatment, was [given] the modern treatment, [prompt] medical treatment* —the verb *take* was not chosen at all by the NS in any equivalent context to the NNS's take-5 examples. On the other hand, in the NNS corpus, the other verbs than *take* that are used in collocation with *surgery* were found to be *get* (2 occurrences), *have* (2 occurrences), and *undergo* (1 occurrence). Considering that the three verbs are listed for *surgery* in the Oxford Collocations Dictionary while *take* is not, the take-5 cases in the examples (29)-(34) should rather be taken as the repetition of a less desirable verb choice by the NNS for the 'verb + *surgery*' construction.

Indeed, the overuse of the formulaic sequences has been identified by the various studies that examined interlanguage in comparison to the native use of language (Kim & Heine, 2011; Palacios-Martinez & Martinez-Insua, 2006; Selinker & Lamendella, 1979; Vigil & Oller, 1976). The lexical approach (Lewis, 1993) proclaims that language is comprised of grammaticalized lexis, and collocation is a principle which organizes language. Within the approach, the prefabricated "chunks" (p. vi) are considered to be a core of language learning. However, they often cause the so-called fossilized expressions, i.e. those that are not grammatically incorrect but not native-like, in learner language when the learner's inventory of prefabricated chunks is limited. The overuse of *take surgery* in the NNS corpus may be explained in the same vein.

4.4 Cognitive Validation: Sense Saliency

The cognitive validation of the LNT model of *take* was carried out by means of the sentence production experiment on the NS and NNS speakers. Among the whole sentences produced by the participants (n=82), 24 sentences contained one of the seven senses in the LNT model, which amounts to 29.27% of the sentences created (cf. 17.17% of the NS and NNS corpus was covered by the model). Table 6 and Figure 4 show the proportion of each sense as identified in the first sentence that came into mind with *take* written by the participants.

Table 6. Proportions of the seven senses in the experimental data

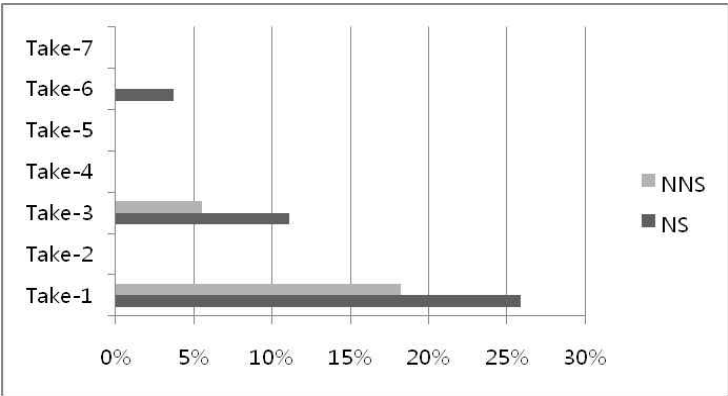| Sense	NS		NNS		Chi-square value
Take-1	7	25.9%	10	18.2%	0.52
Take-2	0	0.0%	0	0.0%	
Take-3	3	11.1%	3	5.5%	0.79
Take-4	0	0.0%	0	0.0%	
Take-5	0	0.0%	0	0.0%	
Take-6	1	3.7%	0	0.0%	2.04
Take-7	0	0.0%	0	0.0%	
Total	11	40.7%	13	23.6%	

Figure 4. Proportions of the seven senses in the experimental data

The results illustrated in Table 6 and Figure 4 indicate that the most salient, so central sense is take-1 (*grab*) as affirmed by the proportion in both the NS and NNS output, being 25.9% and 18.2% respectively, which support the LNT model of Norvig and Lakoff (1987). The second-most central sense in this cognitive perspective, however, was take-3 (*take patient to destination*) instead of take-2, showing 11.1% and 5.5% for the NS and NNS output respectively. Yet, the result may be taken to support, partially at least, the intuitive lexical network model in that the core of the *take*-network in the theoretical framework is the centrality of take-1 (*grab*) as the default. Most uses of *take* in the experiments were converged onto the two senses and others were minimal or nil in both the NS and NNS experiments. As

the Chi-square values (X^2) attested, there was no significant difference between the two groups. The example sentences are (35)-(42).

(35) Take a seat. (take-1, NS experiment)

(36) I want to take it. (take-1, NS experiment)

(37) I am looking at books right now, and hope I find one I like enough to take home. (take-3, NS experiment)

(38) Take me to the movies. (take-6, NS experiment)

(39) Don't take all of them. (take-1, NNS experiment)

(40) You can take it. (take-1, NNS experiment)

(41) Take me to somewhere. (take-3, NNS experiment)

(42) He takes me to the school by car. (take-3, NNS experiment)

Figure 5 shows the clusters of senses which grouped together according to their saliency level based on the experimental outputs. The degree of shading signals the extent of centrality of the senses which belong to the cluster; no shading means no occurrence of the sense.
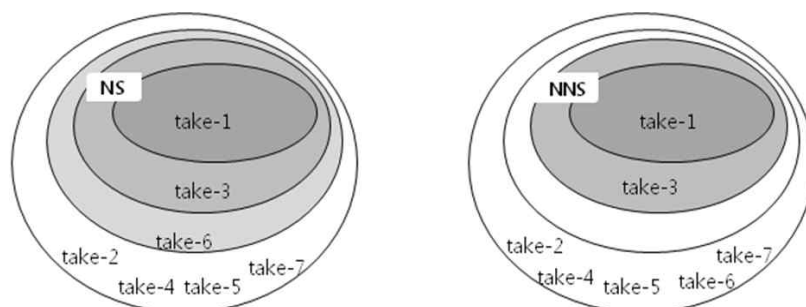

Figure 5. *Take*-network in the NS and NNS experiment

5. Summary and Concluding Remarks

This paper has examined the verb *take* in the framework of the lexical network theory (LNT) to validate the *take*-network proposed by Norvig and Lakoff (1987) against the corpora, and against the output of sentence-elicitation experiments.

Overall, the study was performed in two dimensions: the NS versus NNS use of *take*, and corpus-based versus experimental validation. The results of this study are summarized as Table 7.

Table 7. Summary of the lexical network structure of *Take*

Centrality	N&L* (1987)	Corpus data		Experimental data	
		NS	NNS	NS	NNS
Level 1	take-1	take-3	take-3, take-4	take-1	take-1
Level 2	take-2 take-7	take-1, take-4	take-1, take-5	take-3	take-3
Level 3	take-3, take-4	take-6, take-7,	take-6, take-7	take-6	
Level 4	take-5 take-6	take-2, take-5	take-2	take-2, take-4, take-5, take-7	take-2, take-4, take-5, take-6, take-7

* Norvig and Lakoff (1987)

The findings of this study: first, while the high frequency of a sense as evidenced in the corpora does not correspond to the Norvig and Lakoff's (1987) sense centrality, which has been equated with prototypicality (Gilquin, 2008a, b; Viberg, 2002), the result of the experiment is in concurrence with their model for the most central sense, take-1. Taken superficially, this finding seems to entirely refute Gilquin (2008a) that argued the Norvig and Lakoff's LNT model does not stand up to empirical validations, neither through the corpora nor by the cognitive experiment. However, as pointed out in 4.2, her modification of take-4 to include the light-verb sense of *take* led to a sizable difference in the total ratio of the seven senses between her corpus analysis and this study's; both would have produced not significantly different results otherwise. It was the saliency test that brought about a notable distinction between the two studies: while her experiment had led to the view that the LNT model of *take* does not stand for cognitive reality, the one conducted in this study concurred with the model for the most central sense (default).

Second, our assumption that the intuitive LNT model would be better applied to the native uses of *take* than to the non-native uses, which called for the need of a

contrastive study, turned out to be false. Only take-5 (*take action from agent*) showed a statistically significant difference between the NS and NNS corpus through the Chi-square validation, which, however, should be attributed to some non-native speakers' overuse of a 'fossilized prefab,' as evidenced in 4.3, than be taken as their distinctiveness in using take-5. The result of the elicitation experiment also did not demonstrate any significant difference between the NS and NNS production of sentences. Thus, the current study may be taken as a piece of evidence that the intuitive LNT model of *take* bears the empirical validation by means of a sentence-elicitation experiment, and its application to the NS language and the NNS language does not make a significant difference.

Limitations of this study, mostly related to the nature of research design, need to be noted. As we pursue the compatibility of the NS and NNS use of language, locating the corpora that would fit this specification was a considerable challenge; we were able to secure a reasonable size of compatible NS vs. NNS corpora from the magazine issues of four years, yet the balance of data in terms of registers, genres, language users' gender and ages, etc. was not secured. It was fortunate that the results of the corpus analysis of this study and the previous one based on the existing corpus did not go in the reverse direction. Yet it is possible that research on the basis of other sets of corpora would not produce a result that affirms this research. Second is the homogeneity/representability issue of the participants particularly of the NNSs in the cognitive experiment. They were university students who enrolled in two English-medium courses which are optional, so assumed to hold a relatively advanced level of English competence; this may result in a severe restriction in the possibilities of generalizing the experiment outcomes.

Then a further research suggestion would be how to interpret and/or validate the 'centrality' of a sense. Gilquin (2008a) suggested that the center of the lexical network corresponds to the 'prototype' that has often been considered as the equivalent of either high frequency in linguistic data or the cognitive saliency, as reviewed in 2.2. However, the saliency and frequency may or may not go together just as mind and language may or may not be associated, as the disagreeing results of previous studies and the current one show. The exact nature of the centrality of a sense in the LNT model in relation to cognitive saliency and/or linguistic frequency, and also to the relevance of the latter two concepts, needs to be clarified by future research.

References

Aijmer, K. 2004. The semantic path from modality to aspect: 'be able to' in a cross-linguistic perspective. In H. Lindquist, & C. Mair (Eds.), *Corpus approaches to grammaticalization in English* (pp. 57-79). Amsterdam: John Benjamins.

Altenberg, B. 2001. On the phraseology of spoken English. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and application* (pp. 101-122). Oxford: Clarendon.

Altenberg, B., & Granger, S. 2001. The grammatical and lexical patterning of make in native and non-native student writing. *Applied Linguistics, 22*(2), 173-194.

Altenberg, B., & Tapper, M. 1998. The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80-93). London: Longman.

Amagawa, T. 1996. Another look at *taking* in lexical network theory. *Tsukuba English Studies, 15*. 23-42.

Aston, G. 2000. Corpora and language teaching. In L. Burnard, & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 7-17). Berlin: Peter Lang.

Brinton, L. 2004. Subject clitics in English: A case of degrammaticalization? In H. Lindquist, & C. Mair (Eds.), *Corpus approaches to grammaticalization in English* (pp. 227-256). Amsterdam: John Benjamins.

Brugman, C. 1984. The very idea: A case study in polysemy and cross-lexical generalizations. In D. Testen et al. (Eds.), *Papers from the parasession on lexical semantics* (pp. 21-38). Chicago: Chicago Linguistic Society.

Brugman, C., & Lakoff, G. 1988. Cognitive topology and lexical networks. In G.W. Cottrell, S. L. Small, & M.K.Tannenhause (Eds.), *Lexical ambiguity resolution*: *Perspectives from psycholinguistics, neuropsychology and artificial intelligence* (pp. 477-508). San Mateo: Morgan Kaufman.

Croft, W. 1998. Linguistic evidence and mental representations. *Cognitive Linguistics 9*(2), 151-173.

Gilquin, G. 2008a. Taking a new look at lexical networks. *Lexis, 1*, 23-39.

Gilquin, G. 2008b. What you think ain't what you get: Highly polysemous verbs in mind and language. In J. R. Lapaire (Ed.), *Grammaire et Cognition*. Bordeaux: Presses Universitaires de Bordeaux. (Retrieved January 3, 2011, from http://sites-test. uclouvain.be/cecl/archives/ Gilquin_2008_What_You_Think_Aint_What_You_Get.pdf)

Granger, S., & Tribble, C. 1998. Learner corpus data in the foreign language classroom: Form-focused instruction and data-driven learning. In S. Granger (Ed.), *Learner English on computer* (pp. 199-209). London: Longman.

Gries, S. 2006. Corpus-based methods and cognitive semantics: The many senses of *to run*. In S. Gries, & A. Stefanwitsch (Eds), *Corpora in cognitive linguistics: Corpus-based*

Approaches to syntax and lexis (pp. 57-99). Berlin: Mouton de Gruyter.

Johns, T. 1991. Should you be persuaded – two samples of data-driven learning materials. *English Language Research Journal, 4,* 1-16.

Kim, H.-S. 2008. A study on the incidental development of English competence in CMC, focusing on syntactic and lexical mastery. *Korean Journal of Applied Linguistics, 24*(3), 105-128.

Kim, H.-S., & Heine, B. 2011. Native and non-native use of existential *there* in CMC. *Korean Journal of Linguistics, 36*(1), 69-91.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago.

Lam, P., & Hung, J. 1998. The use of multi-word verbs in advanced Chinese ESL learners. In S. Granger, & J. Hung (Eds.), *Proceedings of the first international symposium on computer learner corpora, second language acquisition and foreign language teaching* (pp. 80-82). Hong Kong: The Chinese University of Hong Kong.

Lee, Y.-H. 2007. *NLPTools*. Seoul: Cambridge.

Lewis, M. 1993. *The lexical approach*. Boston: Heinle ELT.

McEnery, T., & Kifle, N. 2002. Epistemic modality in argumentative essays of second-language writers. In T. Flowerdue (Ed.), *Academic discourse* (pp. 182-195). London: Longman.

Na, Y.-H., & Kim, H.-S. 2010. Tracing L2 learners' accuracy development in CMC. *English Language Teaching, 22*(1), 195-221.

Nevalainen, T. 2004. Three perspectives on grammaticalization: Lexico-grammar, corpora and historical sociolinguistics. In H. Lindquist, & C. Mair (Eds.), *Corpus approaches to grammaticalization in English* (pp. 1-31). Amsterdam: John Benjamins.

Nordlund, M. 2008. *From physical to mental acquisition: A corpus-based study of verbs*. Unpublished doctoral dissertation. Lulea University of Technology, Lulea, Sweden.

Norvig, P. 1986. On defining polysemous words. In *Proceedings of the Third Annual Conference on Theoretical Issues in Conceptual Information Processing*. Philadelphia, USA. Retrieved on January 12, 2011, from http://norvig.com/ticip.pdf.

Norvig, P. 1989. Building a lexicon with lexical network theory. *Proceedings of the First International Language Acquisition Workshop*. Detroit, USA. Retrieved on January 12, 2011, from http://www.norvig.com/lnt.pdf.

Norvig, P., & Lakoff G.. 1987. Taking: A study in lexical network theory. In J. Aske, N. Beery, L. Michaelis, & H. Filip (Eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 195-206). Berkeley: BLS.

Oh, S.-Y. 2000. *Actually* and *in fact* in American English: A data-based analysis. *English Language and Linguistics 4*. 243-268.

Oh, S.-Y. 2007. A corpus-based study of epistemic modality in Korean college students' writings in English. *English Teaching, 62*(2), 147-175.

Palacios-Martinez, I., & Martinez-Insua, A. 2006. Connecting linguistic description and language teaching: native and learner use of existential *there*. *International Journal of Applied Linguistics, 16*(2), 213-231.

Radden, G. 1992. The cognitive approach to natural language. In P. Martin (Ed.), *Thirty years of linguistic evolution: Studies in honour of René Dirven on the occasion of his sixtieth birthday* (pp. 513-541). Amsterdam: John Benjamins.

Rice, S. 1996. Prepositional prototypes. In P. Martin, & D. René (Eds.), *The construal of space in language and thought* (pp. 135-165). Berlin: Mouton de Gruyter.

Selinker, L., & Lamendella, J. (1979). The role of extrinsic feedback in interlanguage fossilization: a discussion of "Rule fossilization: A tentative model." *Language Learning, 29*, 363-375.

Stubbs, M. 2005. Conrad in the computer: examples of quantitative stylistic methods. *Language and Literature, 14*(1). 5-24.

Ungerer, F., & Schmid, H. 1996. *An introduction to cognitive linguistics*. London: Longman.

Viberg, A. 2002. The polysemy of Swedish *ge* "give" from a crosslinguistic perspective. In A. Braasch, & C. Povlsen (Eds.), *Proceedings of the Tenth Euralex International Congress* (pp. 669-682). Copenhagen: Copenhagen University.

Vigil, N., & Oller, J. 1976. Rule fossilization: A tentative model. *Language Learning, 26*, 281-295.

Williams, N. 1992. Processing polysemous word in context: evidence for interrelated meanings. *Journal of Psycholinguistic Research 21*(3). 193-218.

Hyung-Sun Kim & Baeg-seung Kim
Department of English Language and Literature
Chonnam National University
300 Yongbong-dong, Buk-gu,
Gwangju 500-757, Korea
E-mail: hskim4@dreamwiz.com, baegseung@hanmail.net