「언어연구」, 제22권 제2호(2005), 87～107

A Word-based Model Approach to Synthetic Compounds

Yong-Seon Jang
(Sahm Yook University)

Jang, Yong-Seon. 2005. A Word-based Model Approach to Synthetic Compounds. *Linguistic Research* 22.2, 87-107. This paper taking a word-based model approach aims to show that, without recourse to existing lexical or syntactic rules, synthetic compounds can be accounted for in terms of word schemas describing morphological correspondence between words in the lexicon. For this purpose, this paper assumes that the lexicon consists of complete words and word schemas, not of morphemes. Consequently, it does not matter whether the internal structure for synthetic compounds is [[N V] -affix] or [N [V -affix]] in this paper while previous studies in the morpheme-based model have highly taken into account the difference between these structures. Furthermore, within the word-based model, we can effectively describe not only synthetic compounds but also [N+V] and [V+N] exocentric compounds and [N+V] compound verbs derived by back-formation. **(Sahm Yook University)**

Keywords exocentric compounds, morphological correspondence, synthetic compound, word-based model, word-schema

1. Introduction

In a synthetic compound like *bookseller*, the right-hand constituent, the head, inherits the argument structure of the verb *sell* from which it is derived by attaching the suffix *-er* and the non-head *book* satisfies the internal argument of the verbal base. Thus, depending on the different ways to explain how to inherit and satisfy the argument structure of the verb, two internal structures for synthetic compounds are proposed as follows.

(1) a. [[book sell] -er]
b. [book [sell -er]]

The structure in (1a) is postulated by Lieber (1983), Fabb (1984), Sproat (1985), Di Sciullo (1992), and Lim (2004), while (1b) is proposed by Di Sciullo and Williams (1987), Lieber (1992), Oshita (1995), and Plag (2003). The main reason for their postulating different internal structures results from the assumption that the lexicon consists of roots and morphemes and that the process of deriving synthetic compounds includes concatenation of roots such as *book* and *sell* and a morpheme like *-er*.

This paper, however, will show that it does not matter whether the internal structure for synthetic compounds is [[N V] -affix] or [N [V -affix]] since the lexicon is assumed to consist of complete words and word schemas, not of morphemes.

Without recourse to existing lexical or syntactic rules, this paper will show how not only synthetic compounds but also [N+V] and [V+N] exocentric compounds and [N+V] compound verbs derived by back-formation can effectively be accounted for in terms of word schemas describing morphological correspondence between words in the lexicon. Thus, this paper will take a word-based model approach advocated by Spencer (1988), Becker (1993), Bochner (1993), Booij (1997), and Haspelmath (2002).

2. Compound Types

English compounds can largely be classified into two types, based on whether they have a head or not. The first type is called endocentric compounds in which the right element functions as the head (Williams 1981, Spencer 1991, Katamba 1993). The other is exocentric compounds which lack a head.[1]

[1] There is another type of compounds such as singer-songwriter, scientist-explorer, poet-translator and hero-martyr in which two elements do not have any dependency holding between them. These compounds, called copulative or appositional compounds (Plag 2003:

Primary compounds belong to the first type. For example, in a primary compound like *houseboat*, the right element determines the syntactic category of the entire compound. And it also functions as the semantic head of this compound. Thus, a *houseboat* is a kind of boat. Synthetic compounds such as *bookseller* and *window-cleaning* are another class of endocentric compounds in that the right-hand constituent functions as the head.[2] Thus, for example, a *bookseller* is someone who sells books.

On the other hand, compounds such as *frostbite* and *turncoat* belong to the second type. For example, in an exocentric compound like *frostbite*, the right element does not determine the syntactic category of the entire compound, nor does it function as the semantic head. In contrast to the compound *frostbite*, *turncoat* is a noun, just like its right element *coat*. Nevertheless, it belongs to the second type since *coat* is not the semantic head of *turncoat*. A *turncoat* is a person, not a kind of coat (Plag 2003).

It appears that there is no difference between primary and synthetic compounds since they both belong to the same type of compounds and the meaning of both kinds can be predictable from the meaning of the head. However, it has been pointed out that the meanings of some primary compounds are not always derivable from their component words. Some primary compounds as a whole bear idiosyncratic meaning, undergoing lexical drift, as shown in the following examples (Chun 1995, Ahn 1998, Plag 2003).

(2) bigwig, bird dog, blackmail, egg head, lazybones, red head, red skin

---

146), will not be dealt with in this study whose main concern is compounds for which internal structure is considered to have [N+V].

[2] The head of synthetic compounds is widely assumed to consist of a verb followed by one of affixes such as agentive *-er*, nominal and adjectival *-ing*, and the passive adjectival *-en* (Roeper and Siegel 1978, Lieber 1983, Fabb 1984, Carstairs-McCarthy 1992). However, in addition to these affixes, Selkirk (1982), Sproat (1985), Di Sciullo and Williams (1987), and Grimshaw (1990) include other nominalized forms such as *slum clearance*, *consumer protection*, *car theft*, *troop deployment*, *property appraisal*, and compounds based on adjectives such as *machine-readable*, *water repellent*, *disease inhibitory*. As pointed out by Hong (2001: 822), these different views result from different definitions on synthetic compounds.

On the other hand, the meanings of synthetic compounds are assumed to be generally transparent. They tend to be derivable compositionally from the meanings of their component words (Katamba 1993: 309).

The problem is that there are also many counter-examples showing that the meanings of synthetic compounds are not always derivable transparently from their parts. In some cases, synthetic compounds as a whole bear idiosyncratic meaning, too, undergoing lexical drift, as illustrated in the following examples (Allen 1978, Bauer 1983, Spencer 1991, Oshita 1995).

(3) chain smoker, food poisoning, street seller, Sunday seller, town crier, vacuum cleaner,

As reported by Bauer (1983: 181), Oshita (1995: 189), and Plag (2003: 150), for example, a *vacuum cleaner* is not a device to clean vacuums, nor is a *town crier* a person who cries towns, nor does a *street-seller* sell streets. And a *chain smoker* does not smoke chains, rather he or she smokes a cigarette, and *food-poisoning* does not poison food, rather it is a case of illness caused by unsanitary food. Consider the contrast between the examples in (4a) and in (4b).

(4)	a. vacuum cleaner	b. pipe cleaner
	chain-smoker	cigarette-smoker
	food-poisoning	rat poisoning

The above examples show that some synthetic compounds bear idiosyncratic meaning as in (4a) while some bear compositional meaning as in (4b). In short, semantic compositionality is not the unequivocal basis on which we can distinguish primary and synthetic compounds.

Then, what is the difference between primary and synthetic compounds? Roeper and Siegel (1978), Selkirk (1982), Spencer (1991), and Katamba (1993) have attempted to explain the difference in terms of syntactic parallelism, semantic compositionality, and morphological structure. However, Lim (2004)

recently argues that the difference between two kinds of compounds is not evident in terms of syntactic parallelism and semantic compositionality. Rather, the difference can be accounted for solely by the fact that the internal structures for both compounds are different from each other. In the following section, Lim's contention will be examined briefly and then several problems will be pointed out.

3. Internal Structure for Synthetic Compounds

First of all, let us consider the ambiguous compound *tree eater* noted in Selkirk (1982: 28-29). Selkirk assumes that, on the synthetic compound interpretation, it means 'an eater of trees,' while, on the primary compound interpretation, it may mean 'a creature which habitually eats in trees.' Moreover, she assumes that whether the compound *tree eater* is assigned to sets of primary or synthetic compounds, it has the same internal structure in both interpretations as shown in (5b) since there is no $[N\ V]_V$ structure in English. However, Lim (2004) argues that the ambiguity results from the difference of the internal structures, proposing that the synthetic compound interpretation is based on the internal structure in (5a) while the primary compound interpretation is based on the structure in (5b).

(5) a. [[N V] -affix]
b. [N [V -affix]]

Contrary to Selkirk (1982), Lim argues that synthetic compounds can be produced from the non-existing structure *[N V]V in English just as the so called exocentric compounds can be produced from the same structure. In order to make this possible, he stipulates the following conditions (Lim 2004: 653).

(6) a. The concatenation of [N V] is a possible combination in English but it cannot stand alone as a free word,

b. so it needs a process of zero conversion[3] or affixation of *-er*, *-ing*, and *-ed*, to be a free word.

He supports his stipulation by the following examples.

(7) group-think, bootblack, daybreak, moonshine, spacewalk, frostbite
(8) a. *He is a goer.
b. He is a movie goer.

According to him, an exocentric compound like *frostbite* is derived from the unacceptable structure *[[frost]$_N$ [bite]$_V$]$_V$ on its own. To get grammaticality, it needs a process of zero derivation. In the same manner, a synthetic compound like *movie goer* is derived from the unacceptable structure *[[movie]$_N$ [go]$_V$]$_V$ on its own. To get grammaticality, it needs a process of attaching the suffix *-er*. He furthermore assumes that even the sentence in (8a) is ungrammatical since **goer*, a non-free form, is used alone. Even though **goer* is ungrammatical when it is used by itself, it can be used to form a grammatical compound such as *movie goer*. And, in addition to **goer*, he provides more examples such as **avoider*, **dweller*, **maker*, and **teller* to support his stipulation that a non-free word becomes a free word only when it is attached by another word, as shown in the following examples.[4]

(9) a. *maker; doll maker, game maker, movie maker, watch maker
b. *avoider; mouse avoider, proxy avoider, string avoider
c. *dweller; city dweller, cottage dweller, mountain dweller
d. *goer; concert goer, film goer, gallery goer, play goer
e. *teller; fortune teller, future teller, star teller, story teller

There are, however, several problems in Lim's study. First, the idea that the internal structure for synthetic compounds is [[N V] -affix] is also proposed by

[3] Lim uses the term zero conversion but the term zero derivation will be used hereafter in this study because conversion or zero derivation is generally used.
[4] The examples in (9) are not the same to those provided in Lim (2004: 651).

Lieber (1983) in terms of the argument structure of the verb and the argument inheritance,[5] and by Di Sciullo (1992), based on selectional requirements.[6] In addition to these lexicalists, Fabb (1984) and Sproat (1985) also propose that [[N V] -affix] is the correct constituent structure on syntactic ground.[7] However, the internal structure [[N V] -affix] has been criticized on the ground that there is no source verb like **truck drive*, **booksell*, or **beer-drink* in English (Plag 2003: 149). Furthermore, as Booij (1988) and Carstairs-McCarthy (1992: 114) point out, it seems strange that productive synthetic compounds should be based on compound verbs with the structure *[N V]$_V$ which are not productive nor grammatical in English. And there are other views on the internal structure argued by Di Sciullo and Williams (1987), Lieber (1992), Oshita (1995), and Plag (2003) who prefer the different structure [N [V -affix]]. So, to them, there is no problem of the non-existence of a source verb.

Whether synthetic compounds are accounted for by lexical rules or syntactic rules, previous studies postulating the internal structure for synthetic compounds assume that the lexicon consists of roots and morphemes and that the process of deriving synthetic compounds includes concatenation of roots and a morpheme. The theory based on this assumption can be called "the morpheme-based model" (Haspelmath 2002: 47). In this model, it matters whether the internal structure for synthetic compounds is [[N V] -affix] or [N [V -affix]].

Second, Lim (2004) provides such words as **maker*, **goer*, **avoider*, **dweller*, and **teller* to support his stipulation that a non-free word becomes a free word only when it is attached by another word. However, although they are not often used, these words can nonetheless be found in fully acceptable sentences even when they are used on their own, as illustrated in (10). The following sentences can easily be searched on the web site http://www.yahoo.com.

---

[5] Lieber assumes that the argument structure of the verb is a kind of feature which is subject to percolation to a dominating node of the same syntactic category.

[6] That is, the V selects its sister N, and the suffix selects its sister [N V]$_V$.

[7] In syntactic structure, N must be adjacent to the verb and thus be governed by it for theta-role assignment. In the same manner, in synthetic compounds with the structure [[N V] -affix], the non-head can be adjacent to the verb and thus be governed by it.

(10) a. It's not an easy thing to meet your maker.
b. Are you a seeker or an avoider? You can change avoidance behaviors; trainers play a crucial role.
c. Bored on the train? Give dweller a go.
d. Be a goer.
e. Truth never hurts the teller. (Robert Browning)

Third, he argues that exocentric compounds in (11a) are derived from the unacceptable structure *[N V]$_V$ by a process of zero derivation. In addition to exocentric compounds in (11a), there are exocentric compounds shown in (11b) whose internal structure is assumed to be [V N]$_N$. He does not deal with these compounds which have a syntactic head but, nevertheless, might need a process of zero derivation in the morpheme-based model. The reason for this is that the right element determines the syntactic category of the entire compound but it does not function as the semantic head of this compound. Consider the following examples.

(11) a. group-think, bootblack, daybreak, moonshine, spacewalk, frostbite
b. cutthroat, dreadnought, kill-joy, pickpocket, spoilsport, turncoat

In short, in order to account for synthetic compounds and [N+V] exocentric compounds, Lim (2004) proposes the same internal structure [N+V]$_V$ followed by a process of affixation or zero derivation, as illustrated in (12).

(12) a. synthetic compound : [[N+V]$_V$ -affix]
b. exocentric compound : [[N+V]$_V$ Ø]$_N$

It appears that, in (12), Lim can account for both synthetic compounds and [N+V] exocentric compounds in an effective way. However, he does not treat [V+N] exocentric compounds in (11b) and [N+V] compound verbs derived by back-formation. It is not easy to account for these compounds consistently in the morpheme-based model.

4. Reanalysis in the Word-based Model

In this section, compounds will be reclassified into four kinds such as synthetic compounds, [N+V] and [V+N] exocentric compounds, and [N+V] compound verbs derived by back-formation. And we will confirm that these compounds can be described effectively in the word-based model advocated by Spencer (1988), Becker (1993), Bochner (1993), Booij (1997), and Haspelmath (2002).[8]

4.1 [N+V] Compounds

In section 2, we classified English compounds into two types such as endocentric and exocentric compounds, based on whether they have a head or not. According to this basis, synthetic compounds belong to endocentric compounds even though there are some examples which bear idiosyncratic meaning. In contrast, [N+V] compounds like *frostbite* and [V+N] compounds like *turncoat* belong to exocentric compounds since they do not have an element which functions as a semantic head. In addition to synthetic and exocentric compounds, there are some words which apparently have the structure $[N+V]_V$, as illustrated in (13) (Selkirk 1982, Oshita 1995).

(13) air-condition, baby-sit, brainwash, breast-feed, handwrite,
housekeep, henpeck, spoon-feed, vacuum-clean

It is considered that the words in (13) are derived by back-formation. For example, the verb *vacuum-clean* which has the internal structure $[[\text{vacuum}]_N [\text{clean}]_V]_V$ is from *vacuum cleaner* and the verb *housekeep* is from *housekeeper*. We need to point out two things about these compound verbs. First, these are different from [N+V] exocentric compounds in (11a) in that they are fully

[8] The word-based model has mainly been developed to describe such phenomena as back-formation and bracketing paradox. However, in this paper, synthetic compounds and exocentric compounds will also be accounted for in this model.

grammatical on their own without the process of zero derivation even though they have the structure $[N+V]_V$. Second, these compound verbs can be thought to belong to endocentric compounds on the one hand in that the right element decides the syntactic category of the whole compound and to exocentric compounds on the other hand in that the right element does not function as a semantic head. These compound verbs have not been dealt with in the morpheme-based model. The reason for this is that back-formation has been regarded as a diachronic process. Probably, another reason is that there is no productive rule combining a noun and a verb to form [N+V] compounds in English. However, as pointed out by Bauer (1983: 65) and Becker (1993: 6), back-formation must be dealt with in a synchronic grammar since "in current English, it does still thrive." And we need to explain the fact that semantically *housekeep* and *housekeeper* have close relationship.

4.2 Overview of the Word-based Model

Here let us briefly review the basic concept of the word-based model and how to capture the relationships between relevant words in the lexicon. It has dominantly been assumed within Generative Morphology that the lexicon consists of roots and morphemes, and that complex words are derived by morphological rules which combine morphemes and roots in the same manner as syntactic rules put together words to derive sentences (Aronoff 1976, Selkirk 1982, Spencer 1991). For example, a complex word such as *improper* is derived by concatenating the derivational prefix *im-* which has the lexical entry in (14a) and the root *proper* which has the lexical entry in (14b).

(14) a. $\begin{bmatrix} /im/ \\ \_\_\ A \\ \text{'not'} \end{bmatrix}$ b. $\begin{bmatrix} /prɑpər/ \\ A \\ \text{'proper'} \end{bmatrix}$

This approach to complex words can be called the morpheme-based model in which complex words are derived by combining roots and morphemes. In

contrast, the word-based model assumes that the lexicon consists of complete words, not of morphemes, and that word-schemas[9] capture the relationships between the words in the lexicon (Plag 2003). For example, in the word-based model, the similarities among the words in (15a) are captured not by splitting up *impatient*, *impotent*, *improper*, and *impure* into parts such as the derivational prefix *im-* and roots *patient*, *potent*, *proper*, and *pure*. Rather, they can be expressed by a word-schema in (15b) or (15c).[10]

(15) a. impatient, impotent, improper, impure

b. $\begin{bmatrix} /imX/ \\ A \\ \text{‘not } x\text{’} \end{bmatrix}$ c. $\begin{bmatrix} /imX/_A \\ \\ \text{‘not } x\text{’} \end{bmatrix}$

In the word-schema formulated in (15c), the variables /X/ and (‘x’) stand for the common features between the related words in (15a). In addition to the schema in (15c) which subsumes a set of words in (15a), there exists a closely related schema in (16b) which subsumes a very similar set of words in (16a).

(16) a. patient, potent, proper, pure

b. $\begin{bmatrix} /X/_A \\ \text{‘}x\text{’} \end{bmatrix}$

Now the relationships between the words in (15a) and the words in (16a) can be captured by the rule in (17) representing the morphological correspondence.[11]

---

[9] A word-schema is similar to a lexical entry in that it contains information on pronunciation, syntactic properties, and meaning. However, it is different from a lexical entry in that it stands for complete words while a lexical entry may stand for words and individual morphemes.

[10] In (15c), the word-class information is shown as a subscript. This is an abbreviated notational variant of (15b) which will be used hereafter to save space.

[11] In order to represent morphological correspondence between word schemas, Becker (1993) uses the unidirectional arrow(→) and Haspelmath (2002) the bidirectional arrow(↔). In this paper, however, the notation (≈) will be used since using the arrow notation might mean the direction of word formation. The notation (≈) means that for some words matching the

(17) $\left[\begin{array}{l} /X/_A \\ `x' \end{array}\right] \approx \left[\begin{array}{l} /imX/_A \\ `\text{not } x' \end{array}\right]$

4.3 An Analysis of Exocentric Compounds

Let us first account for [N+V] exocentric compounds in (18) in the word-based model.

(18) bootblack, daybreak, moonshine, spacewalk, frostbite

Exocentric compounds in (18) are assumed to have the internal structure *[N V]$_V$ in the morpheme-based model. As examined in section 3, in order to derive acceptable compounds from the ungrammatical structure *[N V]$_V$ which does not occur on its own, Lim (2004: 653) stipulates the process of zero derivation. In fact, the use of zero in deriving complex words is controversial, as pointed out by Katamba (1993: 55) and Lieber (1992: 84). However, in the word-based model, we do not need to stipulate such an ungrammatical structure followed by zero derivation. As an example, (19) shows the English rule that yields an exocentric compound like *frostbite*.[12]

(19) $\left[\begin{array}{l} /frost/_N \\ `\textit{frost}' \end{array}\right]$ & $\left[\begin{array}{l} /bite/_V \\ `\textit{bite}' \end{array}\right] \approx \left[\begin{array}{l} /frostbite/_N \\ `\textit{tissue damage by intense cold}' \end{array}\right]$

The property of exocentric compounds is apparently represented in the rule in (19). That is, the syntactic category of the compound is not the same as that of the head. Moreover, the compound word-schema on the right contains the additional meaning that cannot be predicted exclusively from the meanings of its constituent elements. The additional meaning is rather associated with the pattern as a whole.

---

schema on the left, there is a corresponding word matching the schema on the right.

[12] In compounds, the notation (&) represents that the left-hand side of the correspondence must consist of two word-schemas.

Let us then deal with [V+N] exocentric compounds in (20) in the word-based model.

(20) cutthroat, dreadnought, kill-joy, pickpocket, spoilsport, turncoat

For example, an exocentric compound like *turncoat* is assumed to have the internal structure $[[\text{turn}]_V\ [\text{coat}]_N]_N$. Even though the syntactic category of the compound is the same as that of the head, the process of zero derivation is still needed in order to get the additional meaning the compound has in the morpheme-based model. In the word-based model, however, we do not need to stipulate the process of zero derivation. The following rule shows how we yield exocentric compounds in (21).[13]

(21) $\left[\begin{array}{l} /X/_V \\ \text{`do}_x\text{'} \end{array}\right]$ & $\left[\begin{array}{l} /Y/_N \\ \text{`}y\text{'} \end{array}\right]$ $\approx$ $\left[\begin{array}{l} /XY/_N \\ \text{`person who does}_x\ y\text{'} \end{array}\right]$

In (21), /X/ stands for a verb like *turn* and /Y/ a noun like *coat*. As shown above, the compound word-schema on the right contains the additional semantic element 'person'. This additional meaning is not predicted from the meanings of its constituent elements but is associated with the pattern as a whole.

The similar examples to those in (20) can be found in Italian exocentric compounds as follows (Haspelmath 2002: 88).

(22) portabagagli 'trunk' (portare 'carry' + bagagli 'luggage')
lavapiatti 'dishwasher' (lavare 'wash' + piatti 'dishes')
asciugacapelli 'hair dryer' (asciugara 'dry' + capelli 'hairs')

In an Italian exocentric compound like *portabagagli*, there is no morpheme that can be assigned the additional meaning 'instrument.' The additional meaning is associated with the compound as a whole. Without recourse to the

[13] In (21), 'do_x' represents a variable action meaning.

process of zero derivation, Haspelmath formulates the following rule in (23) for Italian exocentric compounds.[14]

(23) $\left[\begin{array}{l} /\mathrm{Xre}/_{\mathrm{V.INF}} \\ \text{`do}_x\text{'} \end{array}\right]$ & $\left[\begin{array}{l} /\mathrm{Y}/_{\mathrm{N}} \\ \text{`}y\text{s'} \end{array}\right] \approx \left[\begin{array}{l} /\mathrm{XY}/_{\mathrm{N.SG}} \\ \text{`instrument for doing}_x\ y\text{s'} \end{array}\right]$

4.4 An Analysis of Synthetic Compounds

Now let us discuss synthetic compounds within the word-based model. It has been assumed that a synthetic compound like *bookseller* is derived by putting together roots such as *book* and *sell* followed by a morpheme *-er* in the morpheme-based model. Thus, it has been important to postulate the internal structure for synthetic compounds as examined in previous sections. However, in the word-based model, there is no process of attaching a morpheme like *-er* to the root *sell* to form a synthetic compound. The lexicon is assumed to consist of complete words such as *book* and *seller* and word-schemas which capture the relationships between the words. Thus, we do not need to postulate the internal structure for synthetic compounds. Let us then consider how we yield a synthetic compound like *bookseller*. The compound *bookseller* matches two word-schemas at the same time. First, it matches the word-schema of agent nouns given on the right in (24).

(24) $\left[\begin{array}{l} /\mathrm{X}/_{\mathrm{V}} \\ \mathrm{SUBJ} - \mathrm{OBJ} \\ \quad | \qquad | \\ \mathrm{agent_i} \quad \mathrm{patient_j} \\ \text{`A}_i\text{ does}_x\text{ on B}_j\text{'} \end{array}\right] \approx \left[\begin{array}{l} /\mathrm{X}er/_{\mathrm{N}} \\ \mathrm{OBJ} \\ \quad | \\ \mathrm{patient_j} \\ \text{`person who does}_x\text{ on B}_j\text{'} \end{array}\right]$

In (24), /X/ stands for a verb like *sell* and /X*er*/ an agent noun like *seller*. The word-schema of a verb on the left contains information on the valence of

[14] In (23), '*y*s' represents the plural of '*y*'(e.g. luggage, dishes, hairs).

the verb besides information on pronunciation, the word-class, and the meaning. The corresponding agent noun to the verb on the right inherits the valence of the verb. As pointed out by Fabb (1984) and Rappaport Hovav and Levin (1992), when the agent noun inherits the valence of the verb, it absorbs the external argument of the verb since it corresponds to an entity understood as the external argument of the verb.

Second, the compound *bookseller* matches the nominal compound schema on the right in (25).

(25) $\left[\begin{array}{l} /X/_N \\ \\ \\ \\ \text{`}x\text{'} \end{array}\right] \& \left[\begin{array}{l} /Y/_N \\ \text{OBJ} \\ \quad | \\ \text{patient}_i \\ \text{`person who does}_y \\ \text{on } A_i\text{''} \end{array}\right] \approx \left[\begin{array}{l} /XY/_N \\ \\ \\ \\ \text{`person who does}_y \\ \text{on } x_i\text{'} \end{array}\right]$

In (25), /X/ stands for a noun like *book* and /Y/ an agent noun like *seller* which corresponds to the word schema on the right in (24). As shown in (25), the valence of the agent noun requires a patient, which is satisfied by the non-head. Thus, the synthetic compound *bookseller* is predicted to be grammatical. Since the word-schema of an agent noun does not involve the external argument in its valence, however, in synthetic compounds, it is impossible to occur with a noun which functions as the subject of the head, as illustrated in the following examples (Selkirk 1982: 34).

(26) *girl swimming, *weather changing, *kid eating

A transitive verb like *sell* takes an obligatory internal argument. Thus, the word-schema is formulated as in (24). On the other hand, a verb like *drive* takes an obligatory internal argument when the verb is used as transitive but it does not take an obligatory internal argument when it is used as intransitive, which is indicated by the word-schema formulated in (27).

(27) $\left[\begin{array}{l} /X/_V \\ \text{SUBJ} \\ \quad | \\ \text{agent}_i \\ \text{`A}_i \text{ acts}_x\text{'} \end{array}\right] \approx \left[\begin{array}{l} /Xer/_N \\ \\ \\ \\ \text{`person who acts}_x\text{'} \end{array}\right]$

If the verb *drive* necessarily takes an internal argument, the direct internal argument of the verb will be satisfied by the non-head of a synthetic compound as in *truck driver*, while, if the verb *drive* does not take an internal argument, the left element of the compound will not function as the direct internal argument of the verb. In this case, the compound *truck driver* has the primary compound interpretation, as pointed out by Spencer (1991). So it might unusually mean 'someone who drives around on trucks' just as compounds in (28) have the primary compound interpretation (Oshita 1995, Haspelmath 2002, Plag 2003).

(28) city driver, desert driver, moon driver, Sunday driver

Let us now discuss the ambiguous compound *tree eater* noted in Selkirk (1982: 28-29). The synthetic compound *tree eater* which means 'an eater of trees' is yielded by the rule in (25). On the other hand, in the primary compound *tree eater* which may mean 'a creature which habitually eats in trees,' the agentive noun *eater* is yielded by the rule in (27) and the compound is described by the following rule.

(29) $\left[\begin{array}{l} /X/_N \\ \text{`}x\text{'} \end{array}\right] \text{ \& } \left[\begin{array}{l} /Y/_N \\ \text{`}y\text{'} \end{array}\right] \approx \left[\begin{array}{l} /XY/_N \\ \text{`a } y \text{ that has something to do with } x\text{'} \end{array}\right]$

In (29), /X/ stands for a noun like *tree* and /Y/ an agent noun like *eater* which corresponds to the word schema on the right in (27). Haspelmath (2002: 49) formulates the rule in (29) to yield the primary compound *baby-sitter*. Based on his proposal, I employ (29) to yield *tree eater*, *truck driver*, and

other compounds in (28) which have the primary compound interpretation.

Let us finally consider synthetic compounds with the head derived from a verb like *give*. As is well known, the verb *give* has two internal arguments. When we form synthetic compounds with the verb *give*, Theme argument appears as the non-head, not Goal argument, illustrated in the following examples.[15]

(30) a. He gives his child money.
b. *child giver
c. money giver

The compound *money giver* whose head is derived from a verb like *give* which has two internal arguments matches two word-schemas at the same time. First, it matches the word-schema of an agent noun given on the right in (31).

(31)
$$\left[\begin{array}{l} /X/_{V} \\ \begin{array}{ccccc} \text{SUBJ} & - & \text{OBJ} & - & \text{IOBJ}_{to} \\ | & & | & & | \\ \text{agent}_i & & \text{theme}_j & & \text{goal}_k \end{array} \\ \text{`A}_i \text{ acts}_x \text{ B}_j \text{ to C}_k\text{'} \end{array}\right] \approx \left[\begin{array}{l} /Xer/_{N} \\ \begin{array}{ccc} \text{OBJ} & - & \text{IOBJ}_{to} \\ | & & | \\ \text{theme}_j & & \text{goal}_k \end{array} \\ \text{`person who acts}_x \text{ B}_j \text{ to C}_k\text{'} \end{array}\right]$$

As shown in (31), an agent noun like *giver* on the right is assumed to have the argument structure [theme goal]. When we form synthetic compounds with the agent noun, Theme is satisfied inside the compound and Goal may be satisfied optionally outside the compound, not vice versa. Thus we have an acceptable synthetic compound like *money giver to children*, not like **child giver money*. Second. the compound *money giver* matches the nominal compound schema on the right in (32).

---

[15] Lim (2004: 644) also reports similar instances such as *advice giver*, *bribe giver*, *charity giver*, *job-giver*, *support giver*.

(32) $\left[\begin{array}{l} /X/_N \\ \\ \\ \\ \\ \text{`}x\text{'} \end{array}\right] \& \left[\begin{array}{l} /Y/_N \\ \\ \begin{array}{ll} OBJ & -\ IOBJ_{to} \\ | & \quad | \end{array} \\ \begin{array}{ll} theme_i & goal_j \end{array} \\ \text{`person who acts}_y \\ \text{A}_i \text{ to B}_j\text{'} \end{array}\right] \approx \left[\begin{array}{l} /XY/_N \\ \\ IOBJ_{to} \\ \quad | \\ goal_j \\ \text{`person who acts}_y \\ x \text{ to B}_i\text{'} \end{array}\right]$

In (32), /X/ stands for a noun like *money* and /Y/ an agent noun like *giver* which corresponds to the word schema on the right in (31). As shown in (32), the valence of an agent noun like *giver* requires Theme argument, which is satisfied by the non-head inside the compound and Goal argument may be satisfied optionally outside the compound.

Lim (2004) argues that *church giver* is a counterexample demonstrating that Goal argument can appear as the non-head as well. However, in *church giver*, what is given is not church but an object, for instance, money which functions as the implicit Theme argument. Thus, it seems that the rule in (32) is appropriate to describe synthetic compounds with the verb *give*.

4.5 An Analysis of [N+V] Compound Verbs

Let us deal with a compound verb like *housekeep* derived from *housekeeper* by back-formation. Since back-formation has been regarded as a diachronic process, it is disregarded in a synchronic grammar based on the morpheme-based model. In the word-based model, however, based on the word schema formulated in (29), we can describe the morphological correspondence between *housekeep* and *housekeeper* by the following rule.

(33) $\left[\begin{array}{l} /X\textit{er}/_N \\ \text{`one who act}_x\text{'} \end{array}\right] \approx \left[\begin{array}{l} /X/_V \\ \text{`act}_x\text{'} \end{array}\right]$

In (33), the word schema on the left stands for a noun like *housekeeper*

and the word schema on the right represents a compound verb like *housekeep*. As shown in (33), the compound verb *housekeep* and the compound noun *housekeeper* have semantically close relationship.

5. Conclusion

Previous studies on synthetic compounds have been based on the assumption that the lexicon consists of roots and morphemes. And synthetic compounds are formed through a process of affixation and compounding governed by lexical or syntactic rules. So, in the morpheme-based model, it matters whether the internal structure for synthetic compounds is [[N V] -affix] or [N [V -affix]], depending on different manners of ensuring the inheritance and satisfaction of argument structure of the verbal base.

However, this paper has taken a word-based model approach, based on the assumption that the lexicon consists of complete words and word schemas, not of morphemes. Since there is no combining process of morphemes, we don't need to postulate the ungrammatical structure *[N V]$_V$ for a base of synthetic compounds and [N+V] exocentric compounds. In an alternative way, I have described both kinds of compounds in terms of word schemas describing morphological correspondences between words. As an additional benefit of this approach, I have shown that not only synthetic compounds but also [N+V] and [V+N] exocentric compounds and [N+V] compound verbs derived by back-formation can effectively be described in the word-based model.

References

Ahn, Sang-Cheol. 1998. *Hyeongtaeron*(Morphology). Seoul: Mineumsa. (in Korean)

Allen, Margaret. 1978. *Morphological Investigations*. Ph.D. dissertation. University of Connecticut.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, Mass.: MITPress.

Bauer, Laurie. 1983. *English Word-formation*. Cambridge: Cambridge University Press.

Becker, Thomas. 1993. Back-formation, cross-formation, and 'bracketing paradoxes' in paradigmatic morphology. In Booij, Geert and van Marle, Jaap (eds.), *Yearbook of Morphology* 1993. 1-25. Dordrecht: Kluwer.

Bochner, Harry. 1993. *Simplicity in Generative Morphology*. Berlin: de Gruyter.

Booij, Geert. 1988. The relation between inheritance and argument linking: deverbal nouns in Dutch. In Everaert, M., Evers, A., Huybregts, R. and Trommelen, M. (eds.), *Morphology and Modularity: in Honour of Henk Schultink*. 1988. 29-44. Dordrecht: Foris.

Booij, Geert. 1997. Autonomous morphology and paradigmatic relations. In Booij, Geert and van Marle, Jaap (eds.), *Yearbook of Morphology* 1996. 35-53. Dordrecht: Kluwer.

Carstairs-McCarthy. 1992. *Current Morphology*. London: Routledge.

Chun, Sang-Bum. 1995. *Hyeongtaeron*(Morphology). Seoul: Hanshin Publishing Co. (in Korean)

Di Sciullo, Anna-Maria. 1992. Deverbal compounds and the external argument. In I. M. Roca, (ed.), *Thematic Structure: Its Role in Grammar*. 65-78. New York: Foris.

Di Sciullo, Anna-Maria. 1997. On Word-structure and Conditions. In Anna-Maria Di Sciullo, (ed.), *Projections and Interface Conditions*: *Essays on Modularity*. Oxford: Oxford University Press.

Di Sciullo, Anna-Maria, and Edwin Williams. 1987. *On the Definition of Word*. Cambridge, Mass.: MIT Press.

Fabb, Nigel. 1984. *Syntactic Affixation*. Ph.D. dissertation, MIT.

Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, Mass.: MIT Press.

Haspelmath, Martin. 2002. *Understanding Morphology*. New York: Oxford University Press Inc.

Hong, Ki-sun. 2001. Yeongoui dongsaseong boghapeo: nonhanggwa bugaeo. (Verbal Compounds in English: Arguments and Adjuncts). *Korean Journal of Linguistics*, 26-4: 821-843.(in Korean)

Hovav, Malka Rappaport and Beth Levin. 1992. -Er nominals: implications for the theory of argument structure. In E. Wehrli and T. Stowell (eds.), *Syntax and the Lexicon*. Syntax and Semantics 26. 127-153. New York: Academic Press.

Katamba, Francis. 1993. *Morphology*. London: The Macmillan Press Ltd.

Lieber, Rochelle. 1983. Argument linking and compounds in English. *Linguistic Inquiry* 14, 251-285.

Lieber, Rochelle. 1992. *Deconstructing Morphology.* Chicago: University of Chicago Press.

Lim, Jongseong. 2004. Remarks on Synthetic Compounding. *Korean Journal of Linguistics*, 29-4: 635-660.

Oshita, Hiroyuki. 1995. Compounds: A view from suffixation and a-structure alteration.

In Geert Booij and Jaap van Marle, (eds.), *Yearbook of Morphology* 1994. 179-205. Dordrecht: Kluwer.

Plag, Ingo. 2003. *Word-Formation in English*. Cambridge: Cambridge University Press.

Roeper, Thomas and Muffy E. A. Siegel. 1978. A lexical transformation for verbal compounds. *Linguistic Inquiry* 9, 199-260.

Selkirk, Elisabeth. 1982. *The Syntax of Words*. Cambridge, Mass.: MIT Press.

Spencer, Andrew. 1988. Bracketing paradoxes and the English lexicon. *Language* 64: 663-82.

Spencer, Andrew. 1991. *Morphological Theory*. Oxford: Basil Blackwell.

Sproat, Richard. 1985. *On Deriving the Lexicon*. Ph.D. dissertation, MIT.

Williams, Edwin. 1981. On the notions 'lexically related' and 'head of a word.' *Linguistic Inquiry* 12, 245-274.

Division of English Studies
Sahm Yook University
26-21 Kongeung-dong, Nowon-Gu
Seoul 139-742, Korea
E-mail: jangys@syu.ac.kr

접수일자: 2005.10.19
게재일자: 2005.12.16