

Two Types of Compounds in Distributed Morphology*

Tatsuhiko Okubo

1. Introduction

Internal structures of compounds have been one of the arguable topics and been treated across several approaches. One of the approaches is a syntactic approach to internal compositions of words (cf. Roeper and Siegel (1978), Lieber (1992), among others). In this approach, all of compounds as well as derivatives are formed only in syntax. This approach has progressed along with the development of minimalism (Chomsky (2000, 2001, 2008)). Currently, the development crystallized into Distributed Morphology (Halle and Marantz (1993, 1994), Marantz (1997), Embick and Marantz (2008), among others). Many researchers in the realm have focused on the mechanism of inflections and the structures of derivatives since the rise of Distributed Morphology. However, there are few works about internal structures of compounds except Siddiqi (2006), Zhang (2007), and Harley (2009) in the framework. So, internal structures of compounds are worth investigating.

In this paper, I aim to clarify the structures of compounds in Distributed Morphology by employing a Root merger analysis proposed by Zhang (2007) and two distinct domains for ‘word’ formation devised by Marantz (2001).^{1, 2} In so doing, it is found that in addition to Zhang’s structure, an additional structure of compounds is necessary.

The organization of this paper is as follows. Section 2 overviews the definition of compounds proposed by Harley (2009) and a Root merger analysis for compounds proposed by Zhang (2007). In section 3, adapting the definition of compounds and adopting the Root merger analysis, I will propose, based on the two places for word-formation formulated by Marantz (2001), that there are two types of compounds. One has Zhang’s (2007) Root-merged structure. The other is my proposal. Section 4 offers supporting evidence for both structures. Section 5 shows that compounding proposed by Siddiqi (2006) becomes

* I am grateful for helpful comments to Yukio Hirose, Nobuhiro Kaga, Masaharu Shimada, Naoaki Wada, Masaru Kanetani, and Akiko Nagano. My thanks also go to Wenwen Ding, Shotaro Namiki, Ryohei Naya, and Masanao Asano. Needless to say, any remaining errors and shortcomings are my own.

¹ Here, I intend to show by putting a single quotation mark that in Distributed Morphology, words are epiphenomenal and have no theoretically privileged status. Hereafter, I will not use a single quotation mark for intending this just for an expository purpose.

² In this paper, I will mainly focus on English compounds. English compounds obey the right-hand head rule proposed by Williams (1981). Accordingly, in English, the first constituents of compounds, which are left to the heads, are non-heads.

unnecessary if my analysis is on the right track. Section 6 concludes.

2. Compounds and Compounding in Distributed Morphology

2.1. *The Definition of Compounds: Harley (2009)*

Harley (2009) touches upon the definition of compounds. She claims, under the framework of Distributed Morphology, that a compound is a word-sized unit that includes two or more Roots:³

- (1) Compound: a word-sized unit containing two or more Roots
(Harley (2009:130))

According to this definition of compounds, we can correctly distinguish derivatives like *curiosity* made of the Root $\sqrt{\text{CURIOS}}$ and the suffix *-ity* from compounds like *blackbird* composed of the Roots $\sqrt{\text{BLACK}}$ and $\sqrt{\text{BIRD}}$ because derivatives contain only one Root but compounds contain two or more Roots.

2.2. *The Root-Merger Compounds: Zhang (2007)*

Independently of Harley (2009), Zhang (2007) proposes a structure of compounds, as shown in (2):

- (2) $\sqrt{\text{ROOT}} \sqrt{\text{ROOT}}$

The structure in (2) is formed by combining two Roots. Zhang (2007) calls this combination Root merger.

According to Zhang, Root merger explains the abnormality of Chinese compounds: exocentric compounds, the freedom of projectivity, the disappearance of subcategorization, the issue of Case and theta role assignment, the effect of Lexical Integrity in movement, and the effect of Lexical Integrity in pronominalization. Among them, let us show how Root merger captures exocentric compounds in Chinese.⁴ It is well-known that unlike phrases, compounds can be exocentric. Exocentric compounds are very productive in Chinese. Witness the following data:

³ For avoiding confusion of a root, a morphological unit, with $\sqrt{\text{ROOT}}$, a syntactic object used in Distributed Morphology, I henceforth use a label 'Root' to refer to the latter notion.

⁴ For the other abnormal aspects of Chinese compounds, see Zhang (2007).

- (3) a. [da_A-xiao_A]_N (lit.) big-small ‘size,’ [hao_A-dai_A]_{Adv} (lit.) good-bad ‘anyhow,’ [kai_V-guan_V]_N (lit.) open-close ‘switch,’ [bao_V-shou_V]_A (lit.) keep-defend ‘conservative,’ [wun_N-sen]_V (lit.) thing-color ‘to look for,’ [mao_N-dun_N]_A (lit.) spear-shield ‘contradictory’
- b. [kai_V-xin_N]_A (lit.) open-heart ‘happy,’ [pin_A-zui_N]_V (lit.) poor-mouth ‘to talk garrulously,’ [xiao_A-shuo_V]_N (lit.) small-say ‘novel’
(Zhang (2007:172-173))

The compounds in (3a) are composed of two constituents having the same categories, as the subscripts show. If the whole categories of the compounds match with those of their constituents, then the compounds have endocentric structures. However, these compounds are exocentric because the categories of the compounds are not inherited from their constituents. Likewise, compounds in (3b) made of constituents that have different categories from each other are categorized by categories other than those of the constituents. Namely, compounds in (3b) are exocentric. A Root merger analysis of compounds can correctly capture the peculiarity of exocentric compounds. Zhang, based on Embick and Noyer (2007), argues that Roots are category-less and that the constituents of all compounds in (3) are Roots but not categorized words:

- (4) a. kai-xin ‘happy’
b.
-

The tree diagram in (4b) represents the structure of *kai-xin* in (4a). The Roots in (4b) are category-free as mentioned above, so that the Root complex $\sqrt{KAI}-\sqrt{XIN}$ must be categorized in order for the complex to be interpreted. Accordingly, the adjectivalizer *a* attaches to the complex structure in this context. What is important here is that the resultant structure is endocentric but not exocentric; the whole category of *kai-xin* is assigned from the compound’s constituent *a*. In this way, the abnormal behaviors of exocentric compounds in Chinese is explainable by Root merger.

3. Two Types of Compounds

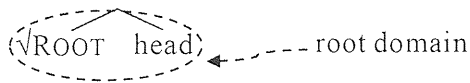
In this paper, adopting the Root merger analysis proposed by Zhang (2007) and adapting the definition of compounds suggested by Harley (2009), I argue

under the framework of Distributed Morphology (Marantz (2001) and Embick and Marantz (2008), among others) that there are two types of compounds. As a first approximation, we have to make clearer what “a word-sized unit” in (1) means by introducing the mechanism of creating a word in Distributed Morphology.

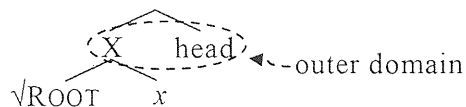
3.1. Two Types of Words and Revising Harley’s Definition of Compounds

Marantz (2001) offers a specific proposal with respect to structures of words. He argues that there are two places for word-formation, namely two types of words. One is a word derived in a root domain, and the other is derived in an outer domain:

- (5) a. A word derived in a root domain

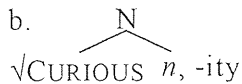


- b. A word derived in an outer domain

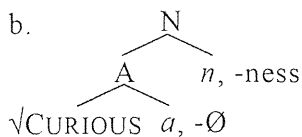


(Marantz (2001) with slight modifications)

- (6) a. curiosity



- (7) a. curiousness



(Embick and Marantz (2008:23))

A root domain illustrated in (5a) is a place where a word is composed by directly attaching a morpheme represented as *head* to a Root.⁵ An outer domain, on the other hand, is a domain where a word is created by combining a *head* to a structure that is already rendered its syntactic category. For better understanding the difference between the two domains, take the derivations of *curiosity* and

⁵ In Distributed Morphology, there are two types of morphemes: abstract morphemes and Roots. Abstract morphemes include categorizers. For the exact definitions of the two morphemes and the difference between them, see section 4.2.

curiousness as examples.⁶ The derived noun *curiosity* is formed in a root domain in such a way that the nominalizer assigned the phonological content *-ity* attaches to the Root $\sqrt{\text{CURIOS}}$ directly, as shown in (6b). By contrast, in (7b), *curiousness* occurs in an outer domain as a result of combining the nominalizer realizing *-ness* to the adjective *curious*. I suggest here that “a word-sized unit” in Harley’s definition of compounds corresponds to the two structures in (5). Under this view, the definition of compounds in (1) is revised as the following:

- (8) The revised definition of compounds
 Compounds are word-sized units containing two or more Roots. The units are derived in root or outer domains.

The fuzziness of “word-sized unit” in Harley’s definition of compounds become clear as shown in (8).

3.2. Proposal: Two Types of Compounds

If there are two types of words as Marantz (2001) argues, it is natural to say that compounds that are words have two types of structures: namely, one derived in a root domain and the other formed in an outer domain. Moreover, Roots are not categorizers, so that they cannot be inserted into the *x* position in (5b). This means that there is only one position for Roots in (5a) and (5b): that is, the *head* positions. Assuming that this reasoning is on the right track, I propose two types of compounds having structures like (9a) and (9b):

- (9) a. A compound derived in a root domain

$$\begin{array}{c} \wedge \\ \sqrt{\text{ROOT}} \quad \sqrt{\text{ROOT}} \end{array}$$
- b. A compound derived in an outer domain

$$\begin{array}{c} \wedge \\ \text{X} \quad \sqrt{\text{ROOT}} \\ \wedge \quad \backslash \\ \sqrt{\text{ROOT}} \quad x \end{array}$$

In (9a), a compound is formed in a root domain where the constituents are two

⁶ Embick and Marantz (2008:11) generalizes that words that have idiosyncratic meanings are formed in root domains, whereas in outer domains semantically compositional words are created. According to this generalization, on the one hand, *curiosity* is a word derived in a root domain because of its semantic non-compositionality, but the semantically compositional word *curiousness* is, on the other hand, formed in an outer domain.

Roots. The structure in (9a) is equal to that derived by Root merger proposed by Zhang (2007). In (9b), a compound is created in an outer domain in which a categorized non-head is merged with a Root.

The next section is devoted to discussing the evidence for the structures in (9a) and (9b).

4. Supporting Evidence

4.1. Neoclassical Compounds

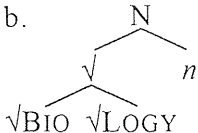
In this paper, the Root merger analysis proposed by Zhang (2007) is employed as shown in (9a). Even if that is the case, Zhang applies Root merger only to Chinese compounds but not to other languages. Accordingly, we need evidence to show that Root merger can capture the behavior of compounds in other languages.

Now, let us turn our attention back to the structure in (9a). The structure is composed only of two Roots. Considering that Roots cannot be pronounced and interpreted without categorizers (Embick and Marantz (2008:6)), it turns out that neither of the two Roots can appear alone. The presence of compounds that are made of bound stems bears out the validity of the reasoning:

- (10) a. bio-logy, psycho-logy, socio-logy
 b. geo-graphy, tomo-graphy

(Booij (2012:88))

The compounds listed in (10a-b) are called neoclassical compounds. This kind of compounds are created by combining two bound stems borrowed from Greek and Latin. For example, *bio-logy* in (10a) is composed of the two bound stems *bio-* and *-logy* that cannot appear alone. It can be safely said from the fact that neoclassical compounds listed in (10a-b) are derived in root domains. (11) is the sample structure of *bio-logy*:

- (11) a. bio-logy
 b. 

The two Roots in $\sqrt{\text{BIO}}$ and $\sqrt{\text{LOGY}}$ are combined in a root domain as clearly shown in (11b). Neither of the two Roots is not hence categorized and cannot occur alone.

the categorizer in (9b) gets assigned some phonological content. This reasoning is supported by the following data:

- (14) a. grammaticality judgment
 b. grading session
 c. participation grade
 d. marketing suggestion
 e. cooler unit
 f. shifter knob
 g. copier service
 h. unhappiness factor

(Siddiqi (2006:86))

The underlined suffixes in (14) are realizations of nominalizers. For example, the structure of (14b) is shown in (15):

- (15) a. grading session
 b.
-
- ```

graph TD
 N1[N] --- V1[√]
 N1 --- N2[n, -∅]
 V1 --- N3[N]
 V1 --- V2[√SESSION]
 N3 --- V3[√GRADE]
 N3 --- N4[n, -ing]

```

As clearly shown in (15b), *-ing* realizes a nominalizer attached directly to the Root  $\sqrt{\text{GRADE}}$ . On the other hand, a null phonological content represented as  $\emptyset$  is assigned to the other nominalizer.

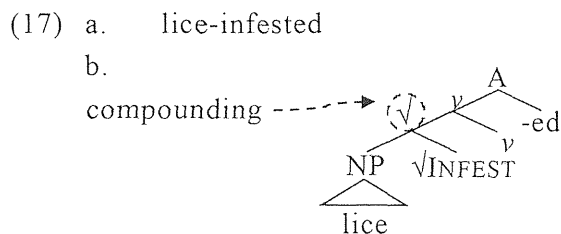
### 5. Consequence: The Rejection of Siddiqi's (2006) Compounding

As far as I am concerned, there has been no literature that deals with compounding except Siddiqi (2006). Siddiqi (2006) defines compounding as a process of combining a Root with an already created phrase:

- (16) Compounding is an application of morphological merger to a pair of nodes  $\alpha$  and  $\beta$ , where  $\alpha$  is a phrase ( $X^n$ ,  $n > 0$ ) and  $\beta$  is a Root, dominated by the phrase (or  $\sqrt{\quad}$ ). (Siddiqi (2006:89) with slight modifications)

A sample structure obtained in this way is given below:





(Siddiqi (2006:91) with slight modifications)

A compound *lice-infested* is derived in the following steps. First, an NP *lice* is formed. Next, a Root  $\sqrt{\text{INFEST}}$  is combined with the NP. At this point, compounding occurs because the NP corresponding to  $\alpha$  in (17) and the Root  $\sqrt{\text{INFEST}}$  corresponding to  $\beta$  in (17) constitute a root phrase  $\sqrt{\phantom{x}}$  where a compounding takes place. The resultant structure does not have a lexical category. Accordingly, it needs to be assigned a lexical category so as to be interpreted. In order to do this, the verbalizer  $v$  is attached to  $\sqrt{\phantom{x}}$  and renders  $\sqrt{\phantom{x}}$  a verb. Finally, an adjectival suffix *-ed* is merged with the resultant structure, changing the verb into an adjective.

Although Siddiqi's compounding is fascinating, it faces a problem if there are two types of compounds as the present paper argues. Namely, Siddiqi's compounding can capture compounds derived in outer domains, but cannot capture those derived in root domains. This is because with respect to compounds derived in root domains, non-heads corresponding to  $\alpha$  in (16) are not phrases or  $X^n$  ( $n > 0$ ) but just Roots or  $X^0$ . Accordingly, the configuration in (9a) is not derivable from Siddiqi's compounding, which is a valid reason for the rejection of his compounding.

## 6. Conclusion

This paper has argued in the framework of Distributed Morphology that there is an additional structure of compounds in addition to a root-merged structure proposed independently by Zhang (2007). The presence of the two types of compounds was clarified by taking into consideration two places for word-formation, a root domain and an outer domain, which are devised by Marantz (2001). I have also showed that Siddiqi's (2006) compounding is unnecessary.

## REFERENCES

- Booij, Geert (2012) *The Grammar of Words: An Introduction to Linguistic Morphology* (3rd Edition), Oxford University Press, Oxford.

- Chomsky, Noam (2000) "Minimalist Inquiries: The Framework," *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michael and Juan Uriagereka, 89–155, MIT Press, Cambridge, MA.
- Chomsky, Noam (2001) "Derivation by Phase," *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1–52, MIT Press, Cambridge, MA.
- Chomsky, Noam (2008) "On Phases," *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, ed. by Robert Freidin Charlos Otero and Maria Luisa Zubizarreta, 133–166, MIT Press, Cambridge.
- Embick, David and Rolf Noyer (2007) "Distributed Morphology and the Syntax-Morphology Interface," *The Oxford Handbook of Linguistic Interfaces*, ed. by Gillian Ramchand and Charles Reiss, 289–324, Oxford University Press, Oxford.
- Embick, David and Alec Marantz (2008) "Architecture and Blocking," *Linguistic Inquiry* 39-1, 1–53.
- Halle, Morris and Alec Marantz (1993) "Distributed Morphology and the Pieces of Inflection," *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. by Ken Hale and Samule Keyser, 111–176, MIT Press, Cambridge, MA.
- Halle, Morris and Alec Marantz (1994) "Some Key Features of Distributed Morphology," *MITWPL* 21, 275–288.
- Harley, Heidi (2009) "Compounding in Distributed Morphology," *The Oxford Handbook of Compounding*, ed. by Rochelle Lieber and Pavol Štekauer, 129–144, Oxford University Press, Oxford.
- Lieber, Rochelle (1992) *Deconstructing Morphology: Word Formation in Syntactic Theory*, the University of Chicago Press, Chicago.
- Marantz, Alec (1997) "No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon," *Proceedings of the 21st Penn Linguistics Colloquium*, ed. by Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark and Alexander Williams, 201–225, University of Pennsylvania, Philadelphia.
- Marantz, Alec (2001) "Words and Things," Handout, MIT, available at <http://dingo.sbs.arizona.edu/~hharley/courses/Oxford/Marantz2.pdf>.
- Niemi, Jussi (2009) "Compounds in Finnish," *Lingue e Linguaggio* VIII-2, 237–256.
- Roeper, Thomas and Muffy Siegel (1978) "A Lexical Transformation for Verbal Compounds," *Linguistic Inquiry* 9-2, 199–260.
- Siddiqi, Daniel (2006) *Minimize Exponence: Economy Effects on a Model of the Morphosyntactic Component of the Grammar*, Doctoral dissertation, the University of Arizona.
- Štichauer, Pavel (2009) "Compounds in Czech," *Lingue e Linguaggio* VIII-2, 293–314.
- Williams, Edwin (1981) "On the Notions 'Lexically Related' and 'Head of a Word,'" *Linguistic Inquiry* 12-2, 245–274.

Zhang, Niina (2007) "Root Merger in Chinese Compounds," *Studia Linguistica* 61, 170–184.

Doctoral Program in Literature and Linguistics

University of Tsukuba

e-mail: okubo.tatsuhiko@gmail.com

