WORD FORMATION IN PHASE THEORY

MAKIKO MUKAI
(University of Kochi)

Abstract

Recursive compounds, such as *peanut butter sandwich* pose a problem within Phase Theory (Chomsky 2001, 2008, Marantz 2007) with regard to Economy Condition. Harley (2009) uses Baker's (1988) incorporation, arguing compounding shows a case of morphology-as-syntax, as it is productive. But if roots are involved in word formation, there is no feature to trigger the incorporation. So her theory cannot be right. I propose a solution to this problem and show how this structure is applicable to recursive compounds.

1. Introduction

Recursion is a fundamental property of human language that potentially differentiates language both from other human cognitive domains and known communication systems in animals (Hauser, Chomsky, Fitch 2002, Corballis 2011). The main aim of this paper is to propose structure for recursive compounds (Roeper, Snyder, Hiramatsu 2002, Bisetto 2010, Tokizaki 2011) within Phase Theory (Chomsky 2001, 2008, Marantz 2007). Therefore, it is not to criticize other frameworks of word-formation, but to understand some aspect of recursion in Phase Theory.

Before going on to talk about recursion in compounding, let us first define what recursion in language is. Summarising a number of linguists, such as Chomsky (1965), Ralli (2013), (Bisetto 2010), Corballis (2011), and many others, I defined recursion as follows in Mukai (2013): it is a phenomenon of embedding structures within structures in cyclic fashion to create sentences, as complex and long as we like. Here, complex means embedding of phrases within phrases of the same kind. In principle it is possible to construct a limitless embedding structures in human language, at least, within the limitless of one's memory and processing capacity. According to Miyagawa & Nóbrega (2015) merge is the recursive operation of the language faculty. In this paper I follow this claim and use Phase Theory to propose structure for recursive compounds.

In this paper I define recursive compounds as compounds with embedding structures within structures. Bisetto (2010) and Tokizaki (2011) define recursive compounds with head addition, not without head addition. Let us show the difference between them.

(1) Waste disposal plan
(2) American English teacher

In (1), the compound *waste disposal* is expanded on the left side by means of the merger of a new constituent, *plan*. In contrast, in (2) the compound *English teacher* is not expanded by merging another head, but it is expanded by merging a modifier *American.* Compounds such as (2) are called iterative compounds by Bisetto (2010). However in this paper I will not follow their definition but argue both (1) and (2) types are recursive compounds in English, Mainland Scandinavian and Japanese and some other languages in the world.

The type (1) is left-branching compound while (2) is right-branching compound. The difference in the branching corresponds to a meaning difference. For example, the example

waste disposal plan means 'plan for waste disposal', not '*disposal plan for waste'*. In contrast, *American English teacher* means ‘English teacher from America’.

This paper will be organised as follows. First, characteristics of recursive compounds will be explained. This section is followed by another section where I explain and criticize Harley’s (2009) theory of compound word formation in Distributed Morphology. Then based on the criticisms I will propose a new theory in Phase Theory for word-formation. The conclusion of this paper will be the new analysis and implications for future research.

2. Characteristics of Recursive Compounds

2.1. Recursive compounds in different languages

As discussed in the previous section recursion can be seen at word level just as at phrase formation. Also, recursive compounds, both right-branching and left-branching is often accepted in many languages of the world, including Japanese, English, Mainland Scandinavian, Chinese, Korean, Dutch, German, Hungarian, Finnish, Thai and Vietnamese.

The first two examples are right-branching and the last two ones are left-branching recursive compounds, except in Korean, Vietnamese and Thai where only one type of branching can be found.

(3)	[kagaku [benkyo beya]] science study room	Japanese

The word *room* in Japanese is translated as *heya* and when this word is the second constituent of a compound, the initial sound /h/ becomes voiced. So the word *benkyo beya* clearly shows it is a compound word.

(4)	[kochi [kenritsu daigaku]] Kochi prefectural university 'University of Kochi'	Japanese
(5)	[booeki gaisha] shachoo] trade company president	Japanese
(6)	[[nise danuki] shiru] mock badger soup	Japanese
(7)	[mail [delivery service]]	English
(8)	[student [feedback system]]	English
(9)	[[waste disposal] plan]	English
(10)	[[peanut butter] sandwich]	English
(11)	[barn-[bok-klub]] child-book-club	Mainland Scandinavian
(12)	[aften-[komputer-klasse]] evening-computer-class	Mainland Scandinavian
(13)	[[jule-mand] costume] Christmas-man costume ‘Santa Clause costume’	Mainland Scandinavian
(14)	[[bo-stad-s] kvarter] live-place-Link area ‘residential area’	Mainland Scandinavian
(15)	[informatie-[terugkoppeling-system]] information-feed back-system	Dutch (Google Netherlands)

(16)	[beroep-s-[auto+andelaar] professional-Link-car+delar ‘professional car dealer’	Dutch
(17)	[[grond-wet-s]-artikel] ground-law-Link-article ‘constitution’	Dutch
(18)	[[scheep-s-bouw] maatschappij] ship-building-Link+company	Dutch
(19)	[Plastik-[garten-zwerg]] plastic-garden-dwarf 'plastic garden dwarf'	German
(20)	[Nord-[bahn-hof]] North-train-court ‘North station’	German
(21)	[[fahrrad-kurrier] jacket] bicycle-courier jacket ‘bicycle courier jacket’	German
(22)	[[Blut-[druck]-apparat] blood pressure apparatus	German
(23)	[Barbie [selkä-[reppu]] Barbie rucksack	Finnish (Google Finland)
(24)	[seinä[kirja-hylly]] wall book case	Finnish (Google Finland)
(25)	[[huone-kalu]-tehdas] room+thing+factory 'furniture factory'	Finnish
(26)	[[ti-tarve]-myynti] home+need+scale 'household sale'	Finnish
(27)	[fal-[könyv-szekrény] wall-book-case	Hungarian (Google Hungarian)
(28)	[fedett-[autó-[parkoló]] roofed-car-park ‘carport’	Hungarian (Hungarian Google dictionary)
(29)	[[vér-nyomás]-mérő] blood pressure apparatus	Hungarian (Kiefer 2009: 527)
(30)	[[pót kerek] csapagy] spare wheel bearing	Hungarian (Tokizaki 2013: 290)
(31)	[suān-[zhòng-dú]] acid-hit-position ‘acidosis’	Chinese (Packard 2004:178)
(32)	[tiě-[fàn-[wǎn]] iron-rice-bowl ‘secure job’	Chinese (Packard 2004:178)
(33)	[[fèiwù chǔ zhì] jìhuà] waste disposal plan	Chinese (Tokizaki 2011:9)

(34) [[zhongguo gongchan] dang] Chinese
China communist party
'Communist Party of China'

(35) [upyeon-[baedal-seobiseu]] Korean (right-branching)
mail-delivery service (Google Koera)

(36) [on-chən]-yok] Korean (left-branching)
warm-spring-bathe
'bathing in a hotspring'

(38) [[dɯ̀ŋ cʌn] tɤ̀i] Vietnamese
line foot sky
'horizon'

(37) [[bá:n sɛ] bɔ̀] (Dang 2013: 21)
wheel vehicle cow/ox
'cow/ox cart wheel'

(38) [[raan khai] khanom] Thai
shop sell sweets
'candy store' Post (personal communication)

As the examples show recursive left-branching and right-branching compounds are productive in these languages. The degree of productivity is different in different languages, but these languages seem to have recursive compounds.

It should be noted that Tokizaki (2013) comments that Thai does not have productive recursive compounds and three-member-compounds are found in menu of dishes. However, Post (personal communication) argues that there are three-member-compounds in Thai, too. Certainly there needs to be more research on this language. However, the definition for 'productivity' of word-formation seems to conflict among different researchers. I leave this to future research.

Also, even though they do not allow as productive recursive compounds as the languages above, Italian and French, allow exocentric VNN compounds. They are right-branching compounds (iterative in Bisetto (2010)) and the VN compound becomes the complement of a verb. Even though they are not as productive as NNN recursive compounds in the above languages, recursion and the presence of self-embedding in such constructions, show that these compounds are derived by a Merge (Miyagawa & Nóbrega 2015). Examples of this type of compounds from these languages are as follows.

(39) [porta-[stuzzica-denti]] Italian
carry 3.pres-pick-teeth
'toothpick-holder' Bisetto (2010)

(40) [porta-[ombrello bici]] Italian (Google Italia)
carry 3. pres-umbrella bike
'umbrella holder for bike'

(41) [porte-[parapluie vélo]] French (Google France)
carry 3. pres-umbrella bike
'umbrella holder for bike'

2.2. Characteristics of recursive compounds in Japanese, English and Mainland Scandinavian

The last sub-section showed some examples of recursive compounds in different languages of the world. The aim of this paper is to propose structures for recursive

compounds which are appropriate for any language of the world. As time and space are limited to show characteristics of recursive compounds in all of the languages discussed in 2.1, I focus on Japanese English and Mainland Scandinavian.

Recursive compounds are phrase-like in the following sense. First, they are pronounced like two independent words and there is a slight pause in between the second and third constituents. Also, they are very productive and internal structure is visible to syntax. However, they do show word-like characteristics in that functional category is excluded and it is impossible to delete part of them.

Let us show the above arguments are true. First the examples from Japanese (3)-(6) are shown.

(42) [kagaku|[benkyo beya]]
(43) [Kochi|[kenritsu daigaku]]
(44) [[nise danuki]| shiru]
(45) [[booeki gaisha]| shachoo]

'|' stands for a short pause. In both right-branching and left-branching compounds there is a short pause. However, it is impossible to have a functional category.

(46) [kagaku [benkyo (*no) beya]]
science study (*GEN) room
(47) [kochi [kenritsu (*no) daigaku]]
(48) [[nise (*no) danuki] shiru]
mock (*GEN) badger soup
(49) [[booeki (*no) gaisha] shachoo]
trading (*GEN) company president

The example (49) is possible if there is a short pause between the words *kenritsu* and *daigaku*. However when the two words are pronounced together like a compound word a functional category cannot intervene.

Also, it is impossible to delete a part of the compounds.

(50) *Gakko de [kagaku [benkyo ~~beya]] o~~ tsuka-tta-ga, daigaku de-wa
School dat [science [study ~~room]] ACC~~ use-PAST but, university dat-const
[kagaku [benkyo beya]] o tsuka-wana-ka-tta.
[science [study room]] ACC use-NEG-past
'I used science study room at school but I did not at university.'

(51) *Ibaragi de [[nis danuki] s~~hiru]o~~ tabe-ta ga, Kochi de [[nise danuki]
Ibaragi dat [[mock badger] ~~soup] acc~~ eat- PAST but, Kochi dat [[mock badger]
shiru o tabe-naka-tta.
eat-neg-PAST
'I ate mock badger soup in Ibaragi, but didn't eat in Kochi.'

(52) *Kobe de wa [[booeki gaisha] ~~shachoo]~~ ~~ni~~ a-tta ga, Osaka de wa
Kobe dat top [[trading company ~~pres. DAT~~ meet-PAST but, Osaka DAT TOP
[[booeki gaisha]shachoo] ni awa-na-ka-tta.
trading company pres. DAT meet-NEG-PAST
'I met a trading company president in Kochi but I did not in Osaka.'

The functional category, genitive case marker cannot intrude the examples or it is impossible to delete its part in recursive compounds in English and Mainland Scandinavian, like in Japanese. First let me show that there is a short pause between the constituents.

(53) [mail|[delivery service]]
(54) [student| [feedback system]]
(55) [[peanut butter]| sandwich]
(56) [[waste disposal]|plan]
(57) [barn|[bog-klub]]
child book club
(58) [aften|[komputer-klasse]]
Evening computer class
(59) [jule-mand]| kostume]
Christmas-man costume
'Santa Claus man'
(60) [[bo-stad]-s|-kvarter]
Live-place-link-area

Right-branching has a short pause after the first constituent, which is like a modifier of the 'compound word', while left-branching has one after the second constituent. This phenomenon is also seen in Japanese, like the examples (44)-(47) show. How about intervention of functional element?

(61) [mail [delivery (*of) service]]
(62) [student [feedback (*of) system]]
(63) [barn [bog-(*s)-klub]]
child book-(*link)-club
(64) [aften [komputer-(*s)-klasse]]
evening computer-(*link)-class
(65) [peanut (*of) butter] sandwich]
(66) [[jule-(*s)-mand] costume]
(67) [[bo-(*s)-stad]-s-kvarter]

The above examples show that both right-branching and left-branching recursive compounds cannot have a functional element in-between.

Also, it should be noted here that there is a linking morpheme in Scandinavian languages (also found in other non-English Germanic languages). According to Josefsson (1997), it is a morpheme without independent meaning. The linking morpheme is used to get this structure and interpretation. Without the linking morpheme, Swedish is strictly right-branching compounds. For example, without a linking morpheme, the compound in (69) is interpreted as "place area for resident", which cannot be logical in the real world. Bisetto (2010) and Mukai (2013) discuss linking morphemes found in compounds in other languages of the world. However there does not seem to have any consensus on their function.

In Section 2 it has been shown that there are examples of endocentric right-branching and left-branching recursive compounds in different languages and exocentric right-branching compounds in Romance languages. In addition focusing on recursive compounds in Japanese, English and Mainland Scandinavian I have shown both phrase-like and lexical characteristics of these compounds.

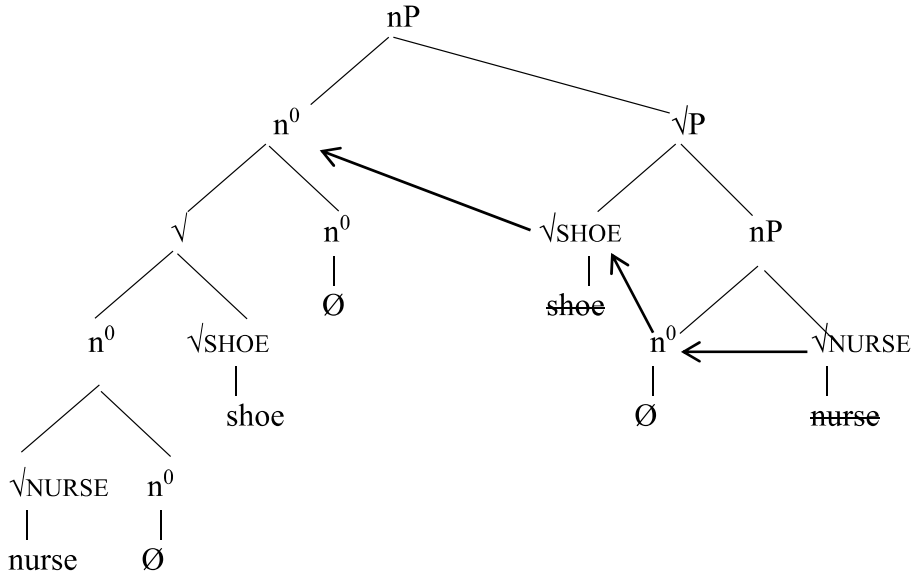
The next section discusses and criticizes Harley's theory of word-formation.

3. Harley's word-formation

In Distributed Morphology it is argued that all generalised transformations are assumed to be of the same kind, showing that they take place in a single module (Marantz 2007). Thus, there is no single place called a 'lexicon' distinct from syntax where some generalised transformations take place. Another point is that only formative aspects of morphemes, but not phonological aspects, are contained in the list of morphemes. With above in mind Harley (2009) argues compounding shows a case of morphology-as-syntax as it is productive and uses Baker's (1988) incorporation for derivation of words.

The following is a structure for primary compounds *nurse shoe*, proposed by Harley.

(68)

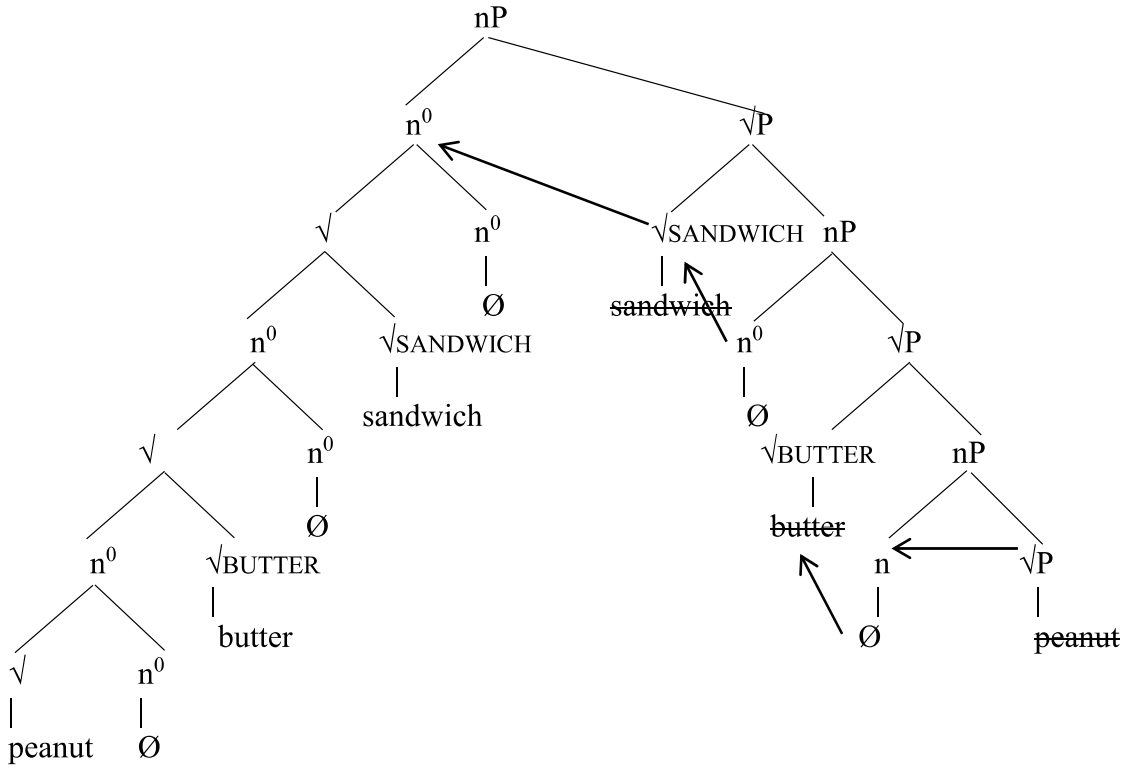


Harley (2009: 140)

Harley uses Baker's (1988) incorporation to capture the Lexical Integrity, namely indivisibility of the elements in the compounds, the impossibility of the movement out of them, and the impossibility of discourse antecedent from within a compound. The structure shows that the roots without categorical feature (Marantz 2007) are merged with their category-determining feature bundle, or otherwise, they cannot appear in the narrow syntax. Also, Harley assumes the modificational relationship between the non-head (*nurse*) and the head (*shoe*) in the compound.

Harley does not propose a structure for recursive compounds, discussed in Section 2. However, I assume that the structure is as follows.

(69)



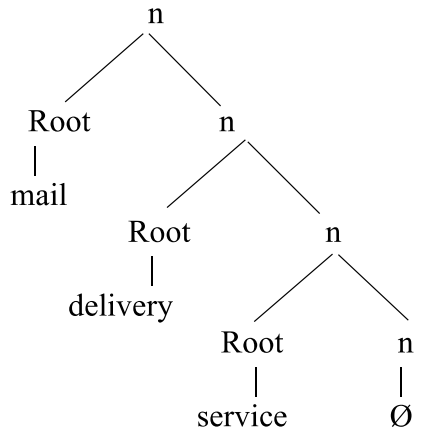
The extra constituent means more embedded structure. As discussed above, Harley assumes Baker's incorporation of roots. However, if roots are involved in word formation, there is no feature to trigger the incorporation. Harley herself admits that roots do not have any categorical feature. Also, in the Distributed Morphology roots are not assumed to have any categorical feature. So her theory cannot be right. Head-movement should be Last Resort, so there should be another theory for word-formation without any movement. In the next section I will propose a new theory based on this criticism in Phase Theory.

4. Analysis of recursive compounds in Phase Theory

Bauke (2009) assumes lexical and syntactic compound word-formation should be formed in different places. However, following Marantz (2007) I assume that both lexical and syntactic compound word-formation takes place in the narrow syntax. In addition, as word formation is sound and meaning as phrase formation is, I assume there is a phase at word-level, as well as phrase level (Marantz 2007). Finally, recursion in word-formation shows that word-formation is derived by the operation Merge and it must contain an internal hierarchical structure (Miyagawa & Nóbrega 2015).

With the above arguments in mind a proposed structure for recursive compounds is as follows.

(70)



The above structure is one for right-branching recursive compounds. This structure is constructed as follows: First, a root without word class feature (Zhang 2007) is merged with a syntactic head, thus turning the root into an n (Marantz 2007). This is labelling, in terms of Chomsky (2008). Then, another root is merged to form compound word and this 'compound' is transferred to the interpretational component and spelled out as phase (Chomsky 2001, 2008).

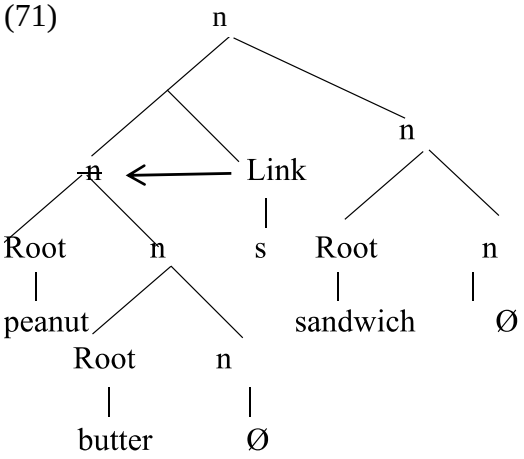
The reason why the two-member compound is assumed to be a phase is that by assuming it is a phase we can capture the Lexical Integrity of compounding (Section 2). That is, as once the compound is a phase, it is impossible to delete part of it or no functional element can intervene between them.

Another root is merged to form the right-branching compound and the head of the whole compound is the 'right-most' element.

The three roots are not merged immediately because the interpretation of the whole compound allows for alternative and compositional interpretation. Also, roots cannot be directly combined with each other as they are void of grammatical feature (Zhang 2007, Mukai 2006, 2008). As a result the n is the head of the whole compound word, as there is no categorizing feature to be interpreted semantically and phonologically.

How about left-branching recursive compounds? I propose the following structure for this type of compounds.

(71)



Like the structure for right-branching recursive compounds, the structure for left-branching recursive compounds does not need any movement, unlike in Harley's theory. The

structure above is derived in the following ways: First, a root without word class feature is merged with a syntactic head, thus turning the root into an n (Marantz 2007). Another root is merged to form a two-member compound word. Here, the compound is transferred to the interpretational component and spelled out. The two-member compound with its syntactic head is interpreted semantically and phonologically. The interpretation is different depending on the roots.

Like in the structure (72) the two roots are not merged immediately in (73), as the interpretation of the whole compound allows for alternative and compositional interpretation. Also, the reason why the two-member compound is assumed to be a phase is that by assuming it is a phase we can capture the Lexical Integrity of compounding (see Section 2). That is, as once the compound is a phase, it is impossible to delete part of it and to be referred back as a discourse antecedent.

The categorizing head is merged, as in recursive compounds there is a slight pause between the second and third constituent of the recursive compound (Tokizaki 2011, 2013). Here, Tokizaki argues that there is a head.

However, as it is impossible to have two heads in the syntax, a linking morpheme, which is phonetically there in Mainland Scandinavian and several other languages (see Section 2), but not in other languages, like Japanese or English, is merged to check the categorizing head. The resulting structure is merged with another root which is merged with a categorizing head, which is derived in parallel. As a result, the categorizing head n is the head of the whole compound.

Unlike in Harley's theory there is no unnecessary movement in this theory.

In this section I proposed structures for right-branching recursive compounds and left-branching recursive compounds in Phase Theory. Unlike Harley's theory there is no movement but only the operation merge. Assuming a phase in the structure can capture the Lexical Integrity of compounds, as discussed in Section 2. The linking element is there for Economy reason to check the feature on the head.

The present theory is applicable for recursive compounds in the languages discussed in Section 2.

What about productivity of recursive compounds? As mentioned in Section 2.1 it is argued that Romance languages do not allow as productive recursive compounds as in the other languages, especially Germanic languages. Tokizaki (2013) argues that this might be due to phonology. I leave this problem for future research.

5. Conclusion

In this paper I proposed structures for recursive compounds in Phase Theory. As discussed in Section 2 recursive compounds show both phrase-like and word-like characteristics. By assuming there is a phase in the structure Lexical Integrity of recursive compounds can be captured without any unnecessary movement, unlike in Harley's theory.

However, there remains some problems to be solved in this theory. First, what kind of feature is there in the head? How is structure for right-branching recursive compounds with a linking morpheme inside in German or Greek languages? More research is required to answer these questions, along with many other questions.

References

Baker, M. C. (1998). *Incorporation*. Chicago & London: Chicago University Press.

Bauke, L. (2009). *Nominal Root Compounds.* Ph.D dissertation, UMass Amherst.
Bisetto, A. (2010). Recursiveness and Italian Compounds. *SKASE Journal of Theoretical Linguistics* 7, 14-35.
Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.
Chomsky, N. (2001). Beyond Explanatory Adequacy. *MIT Occasional Papers in Linguistics* 20, 1-28.
Chomsky, N. (2008). On Phases. In Freidin, R., Otero, P., & Zubizarreta, M. L. (eds.), *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, 133-166, Cambridge, MA: MIT Press.
Corballis, M, C. (2012). *The Recursive Mind: The Origins of Human Language, Thought, and Civilization.* Princeton: Princeton University Press.
Dang, P. H. (2013). *Does prosodic word recursion cause phonetic initial strengthening?* MA Thesis, California State University, Fresno.
Harley, H. (2009). Compounding in Distributed Morphology. In Lieber, R., & Stekauer, P. (eds.), *The Oxford Handbook of Compounding*, 129-144. Oxford: Oxford University Press.
Hauser, M. D. Chomsky, N., & Fitch, T. (2002). The Faculty of Language: What is it, Who has it, and How did it Evolve? *Science* 98, 1569-1579.
Josefsson, G. (1997). On the Principles of Word Formation in Swedish. Lund: Lund University Press.
Kiefer, F. (2009). Uralic, Finno-Ugric: Hungarian. In Lieber, R., & Stekauer, P. (eds.), *The Oxford Handbook of Compounding*, 527-541. Oxford: Oxford University Press.
Marantz, A. (2007). Phases and words. In Choe, S.H. (ed.), *Phases in the Theory of Grammar* (pp. 191-222). Soeul: Dong-In Publishing Co.
Miyagawa, S., & Nóbrega, N.A. (2015). The precedence of syntax in the rapid emergence of human language in evolution as defined by the integration hypothesis. Frontiers. Psychology, 6: 271.
Mukai, M. (2006). *A Comparative Study of Compound Words in English, Japanese and Mainland Scandinavian*, Ph.D dissertation, University of Newcastle.
Mukai, M. (2008). Recursive Compounds. *Word Structure* 1 (2): 178-198.
Mukai, M. (2013). Recursive Compounds and Linking Morpheme. *International Journal of English Linguistics* 3 (4), 36-49.
Packard, J. (2004). *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge: Cambridge University Press.
Ralli, A. (2013). Compounds and its Locus of Realization: Evidence from Greek and Turkish. *Word Structure* 6 (2), 181-200.
Roeper, T., Snyder W., & Hiramatsu, K. (2002). Learnability in a Minimalist Framework: Root Compounds, Merger, and the Syntax-Morphology Interface, 25-37. In Lasser, I. (ed.). *The Process of Language Acquisition.* Frankfurt: Peter Lang Verlag.
Tokizaki, H. (2011). The Nature of Linear Information in the Morphosyntax-PF Interface. *English Linguistics* 28 (2), 227- 257.
Zhang, N. (2007). Root Merger in Chinese Compounds. *Studia Linguistics* 61 (2), 170- 184.

Makiko Mukai
Department of Cultural Studies
2-22 Eikokuji
Kochi 780-8515
Japan

mukai@cc.u-kochi.ac.jp