The Boundary between Word and Phrase: The Case of A-N Expressions within Compounds

Reiko Shimamura
Tsuda College

1. Introduction

An expression consisting of an Adjective and a Noun, henceforth referred to as an A-N expression, can function either as a N(ominal) P(hrase) in a DP or as a compound word. Consider, for example, the contrast between the NP (1a) and the compound (1b) in English, and also between the NP (2a) and the compound (2b) in Japanese.

(1) a. [$_{NP}$ black board]
b. [$_{N}$ blackboard]
(2) a. [$_{NP}$ ama-*i* sake]
sweet-INFL alcoholic drink
b. [$_{N}$ ama zake]
sweet alcoholic drink

In English, the Adjective *black* in (1a) and the one in (1b) are indistinct in form. On the other hand, in Japanese, the Adjective *ama* 'sweet' in (2a), in contrast to the one in (2b), has the inflectional suffix *-i* attached.

It is usually the case that the difference in structure between a phrasal expression and its corresponding lexical expression is closely correlated with their semantic and phonological differences. A-N compounds in English and Japanese, as opposed to their corresponding phrases, tend to be semantically lexicalized. What is called a 'blackboard' can occasionally be green rather than black. In the case of Japanese as well, (2b) is more than a composite of *ama* 'sweet' and *sake* 'alcoholic drink,' having its own additional elements of meaning; that is, it means 'sweet alcoholic drink made from fermented rice.' Furthermore, A-N expressions behaving as phrases and as words are phonologically distinguished from each other. Examples (1a) and (1b) are different in stress contour. In the case of (2b) in Japanese, the initial consonant of the second element has to be the voiced /z/ due to the so-called *rendaku* phenomenon.[1]

Another important difference between phrases and words is that in contrast to a phrase, no element can modify a part of a word or can be inserted between the components of a word, as illustrated in the English and Japanese A-N compounds in (3) and (4) respectively.

(3) a. *[$_{N}$ very blackboard]
b. *[$_{N}$ black hard board]
(4) a. *[$_{N}$ totemo ama- zake]
very sweet alcoholic drink
b. *[$_{N}$ ama oisi-*i* sake]

sweet delicious-INFL alcoholic drink

The phenomenon mentioned above is termed 'lexical integrity' or 'syntactic atomicity,' the principle of which is generally defined as follows (Di Sciullo and Williams 1987, Bresnan and Mchombo 1995, among others):

(5) Lexical Integrity Principle:
Rules of syntax do not have access to the parts of words directly.
(Di Sciullo and Williams 1987: 45)

Here let me pay attention to compounds containing an A-N expression. As shown below, A-N compounds can occur within a larger compound in both English and Japanese.

(6) [$_N$ [$_N$ blackboad] eraser]
(7) [$_N$ [$_N$ ama- zake] manzyuu]
sweet alcoholic drink bean-jam bun

Example (6) means 'eraser for a blackboard,' and (7) 'bean-jam bun covered with a thin wrapping of dough which contains *ama-zake*.' On the other hand, in the case of A-N expressions which seem to behave as NPs rather than as compounds, English and Japanese exhibit a sharp contrast. In English it seems possible to incorporate this type of A-N expression within a compound, while in Japanese this is impermissible. Consider:

(8) [$_N$ [black board] eraser]
(9) *[$_N$ [ama-*i* -sake] manzyuu]

The compound (8) means 'eraser of a black board,' and the A-N expression *black board* in (8), which looks like an NP, retains its semantic transparency in this compound. In contrast, the compound (9) in Japanese, where the A-N expression *ama-i sake*, which is obviously an NP, is unacceptable.

The distinction between the lexical status of the A-N expression *blackboard* in compound (6) and the phrase-like status of *black board* in compound (8) is reflected in the difference in stress contour between these two compounds, as indicated in Chomsky and Halle 1968. In the case of (6), the Compound Rule applies first to the inner A-N expression *blackboard* to assign the main stress to the lefthand A, and next to the compound as a whole. As a result, its stress contour is 132. On the other hand, the compound (8) undergoes the Compound Rule, which assigns the primary stress to the first N which forms the A-N expression *black board*, after the Nuclear Stress Rule assigns the main stress to the righthand N of this A-N expression. This results in the stress contour 312 for this compound. A possible assumption, therefore, would be that A-N expressions inside compounds in English can behave as genuine syntactic phrases (i.e., NPs), as in the case of (8). In fact, Chomsky and Halle 1968 propose that the A-N expression *black board* in (8), to

which the Nuclear Stress Rule applies, is identified as phrasal, i.e., as an NP (in its traditional sense). [2] In this paper, however, I will neither accept this assumption nor support the conclusion which I reached in Shimamura 2001 that lexical and phrase-like A-N expressions inside compounds should indistinguishably be regarded as lexical. In the present paper I will give an alternative solution to the question of the phrase/word distinction of seemingly phrasal A-N expressions within compounds by referring to Chinese and Dutch as well as English and Japanese.

2. A-N expressions within compounds in four languages

In this section, after investigating the behavior of A-N expressions inside compounds in English, Japanese, Chinese, and Dutch, I will suggest that there are two types of languages with respect to whether phrase-like A-N expressions inside compounds such as the one in (8) are permissible or not.

2.1. English and Japanese

Below I will give additional compounds with an A-N expression in English and Japanese to make clearer the difference in the internal structure of such compounds between the two languages, and also examine whether A-N expressions inside compounds as in (8) can be qualified as genuine syntactic NPs by reviewing Shimamura 2001 and also by referring to some suggestions made by other researchers.

In Section 1, I indicated that in Japanese, A-N compounds, but not their corresponding NPs, can be incorporated into larger compounds, as shown by the contrast between (7) and (9). Similar examples from Shimamura 2001 are given below.[3]

(10) a. [yasu- mono] gai
cheap article buying 'the buying of cheap articles'
b. *[yasu-*i* mono] gai
cheap-INFL article buying

(11) a. [huru-hon] ya
old book store 'store where old books are sold'
b. *[huru-*i* hon] ya
old-INFL book store

A-N compounds, as opposed to the corresponding NPs, are not transparent in meaning. The A-N compound *yasu-mono* 'cheap article' within the larger compound (10) usually refers to articles which are not only cheap but also bad. Similarly, the A-N compound *huru-hon* 'old book' inside the compound (11) refers to an old secondhand book.

In the case of English, the following A-N expressions inside the compounds, as with *blackboard* inside compound (6), are regarded as genuine compounds, as shown in Shimamura 2001: 169 and Bennett 2002. [4]

(12) a. [old-boys] network, [white-water] rafting, [blackboard] jungle, [big bang] theory
(Shimamura 2001)

b. [high-school] dance, [smallpox] vaccination, [shortbread] recipe (Bennett 2002)

As explained in Section 1, not only A-N compounds but also phrase-like A-N expressions like *black board* in (8) can occur as the non-head element inside a larger compound. Further examples of this type of compound are shown in (13), where the compounds in (a), (b), and (c) are synthetic compounds, root compounds, and genitive compounds respectively. (They are given in Shimamura 2001. For the sources of the compounds below, see Shimamura 2001.)

(13) a. [purple people] eater, [small car] driver, [gray elephant] hunter, [used computer] seller, [low-wage] earners, [equal opportunity] employer
b. [silent-film] star, [fresh fish] shop, [rare-manuscript] room, [dual-action] sprays, the [easy-action] hook, two [wide-brim] straw hats, a [full-colour] catalogue, his [early-life] experiences, a [cheap-price] series, a [single-storey] factory, [balanced budget] amendment, an [early-retirement] clause, [hazardous-duty] pay, [high power] radar systems, the banks' [low-interest] savings account, a [high-profit] trade, a [little-girl] voice, [good-class] properties, [low-cost] service, a [late-model] car, a [late-night] meeting, [poor quality] goods, the [new style] activists, [low income] group, [old book] shops, [rare book] trade, [single parent] family, [equal rights] amendment
c. [gifted children's] school, [old people's] house, small green [little girl's] bicycle

The following compounds listed in Bennett 2002 are also of this type, where the compounds in (a) are root compounds and the ones in (b) synthetic compounds.

(14) a. [new-road] inquiry, [old-clothes] shop, [old-growth] forest, [direct-action] politics, [small-car] design, [early-morning] feeling, [small-town] life, [live-music] venue, [cold-weather] clothing, [fresh-fruit] shop, [mixed-ability] group, [best-film] list, [new-books] shelf, [high-tension] wire, [modern style] building
b. [cold-milk] drinkers, [warm-beer] lovers, [modern-history] teacher, [new-syllabus] proposal

Moreover, Liberman and Sproat 1992: 161, citing the compounds in (15), mention that the stress pattern of inner A-N expressions is assigned by the Nuclear Stress Rule and therefore that the main stress is placed on the righthand N of the A-N expressions within the compounds. The stress contour of each compound in (15) is 312, as is the case in the compound (8).

(15) [collective bargaining] agreement, [balanced budget] amendment, [civil rights] bill, [used car] business, [due process] clause, [floating underflow] trap, [fresh fish] shop

The compounds indicated in (13)-(15) suggest that various phrasal-looking A-N expressions

can be rather freely formed inside root and genitive compounds as well as synthetic compounds, contrary to Roeper and Siegel 1978, Roeper 1988, and others, who claim that A-N expressions like (16) and phrasal expressions of other forms are impermissible within synthetic compounds. [5]

(16) a. *[good dark coffee] maker cf. coffee-maker (Roeper and Siegel 1978: 213)
b. *[odd day] lover (Roeper 1988: 205)

As for the stress pattern of the A-N expressions in (13)-(15), it seems safe to say that almost all the A-N expressions in (13) and (14), as well as the ones in (15) cited by Liberman and Prince 1992, are given a phrasal stress pattern and pronounced with a phrasal accent according to the Nuclear Stress Rule, though in a few cases there may be disagreement in the stress pattern.

Bennett 2002 claims that phrasal-looking A-N expressions inside compounds like those in (14) are regarded as what is called 'weakly lexical constructions' in Sadler and Arnold 1994, which are considered to be N^0 structures (not NP structures) but nonetheless are syntactic in nature. This claim by Bennett is mainly based on the fact that such A-N expressions are transparent in meaning and therefore not lexicalized.

As explained above, Liberman and Sproat 1992 mention that the A-N expressions in (15) are assigned righthand stress. Their basic assumption is that stress placement is derived from some syntactic or lexical property of a construction. They propose, on the basis of extensive computer analysis, that the Compound Rule applies to N^0 structures and that the Nuclear Stress Rule applies to N'-level syntactic structures, by showing that more than 90% of A-N expressions have righthand stress, whereas about 75% of N-N expressions display lefthand stress. It is also claimed in Sproat 1993 that the phrasal N' structure is possible in synthetic compounds by making reference to *[purple people] eater* and *[small car] driver* given in (13a), though the stress contour of A-N expressions is not referred to in Sproat 1993.

In sum, what is shared by Liberman and Sproat 1992, Sproat 1993, and Bennett 2002 with respect to the analysis of phrasal-looking A-N expressions is that such A-N expressions, even when they occur inside compounds as in (13)-(15), are regarded as genuine syntactic structures. If so, it follows that the compounds in (13)-(15) are all in violation of the Lexical Integrity Principle in (5). However, I will point out two facts which argue against the syntactic status of phrase-like A-N expressions inside compounds.

I have stated above that phrasal-looking A-N expressions are rather freely generated inside compounds in English. However, it should be noted that this pattern is not completely productive; in other words, not all phrases of the form [A-N] are permissible inside compounds. For example, as pointed out in Bauer 1998: 73, the expression in (17a) is usually interpreted as in (17b), i.e., as a broad bed for a river, and not as in (17c), i.e., not as a bed for a broad river.

(17) a. broad river-bed

b. broad [$_N$ river-bed]
c. * [$_N$ [broad river] bed]

If phrasal-looking A-N expressions inside compounds were true syntactic phrases or constructions, (17c) would be perfectly acceptable as a compound. Moreover, according to Sproat 1993: 251 and Lieber 1992: 209, the compound (18) is lower in acceptability than the compound *[small car] driver* given in (13a).

(18) ?[green car] driver (? attached by the present author)

I think that Sproat 1993 correctly considers that because compound nouns are used to name things (cf. Downing 1977 and Bauer 1983), the compound *[small car] driver*, as opposed to the compound (18), is perfectly acceptable because *small car* generally denotes a type of car. According to him, the expression *green car* is regarded as a description rather than a type of car and does not serve appropriately as an element incorporated into a compound. Similar facts indicating that not all phrasal A-N expressions can occur inside compounds are the dubious compounds in (19), in contrast to (20). (Examples (19) and (20) appear in Bresnan and Mchombo 1995: 193.)

(19) ??[recent history] teacher, ??[amplified guitar] player
(20) [American history] teacher, [acoustic guitar] player

The second piece of evidence that A-N expressions inside compounds do not qualify as genuine syntactic phrases or constructions is that compounds like the ones listed below are impossible.

(21) a. *[very small car] driver, *[surprisingly cheap price] service, *[rather late night] meeting
b. *[small green car] driver
*[small new books] shelf, *[new small books] shelf
(cf. [small books] shelf, [new books] shelf)
*[small quiet town] life, *[quiet small town] life
(cf. [small town] life, [quiet town] life)

The unacceptability of the compounds in (21) suggests that seemingly phrasal A-N expressions inside compounds must have a fixed A-N form without modification or insertion in accordance with the Lexical Integrity Principle in (5).[6] If such A-N expressions within compounds were true syntactic phrases or constructions, all the compounds in (21) should be completely acceptable.

To sum up, as explained in Shimamura 2001, we should consider that phrase-like A-N expressions inside compounds like the ones in (13)-(15) are neither fully phrasal nor fully lexical but have an intermediate status. They should regarded as phrasal, because i) they are

not semantically lexicalized, and ii) they are assigned righthand stress by the Nuclear Stress Rule. But at the same time they have something in common with words, because i) the formation of them is not completely productive, and ii) they occur in the fixed A-N form in compliance with the Lexical Integrity Principle in (5).

In Shimamura 2001 I proposed that not only A-N compounds inside larger compounds like the ones in (6) and (12a) but also phrase-like A-N expressions inside compounds like the ones in (13) should be regarded as examples of what is termed 'weakly lexical construction' in Sadler and Arnold 1994, though I pointed out that the former type of A-N expression within a compound is neither fully phrasal nor fully lexical. However, this proposal lacks in adequacy, because it would make it impossible to distinguish between the above-mentioned two types of A-N expressions within a compound. In Section 3, I will discuss how we should account theoretically for the intermediate status of phrase-like A-N expressions inside compounds.

2.2. Chinese

I now turn to Chinese. Suggesting that in Chinese, there are two kinds of nominals of the forms given in (22), Duanmu 1998 argues that phrases and words are respectively of the forms (22a) and (22b), where the Adjective is considered to be a kind of modifier.

(22) a. [M *de* N] (M = modifier, *de* = particle)
b. [M N]

Duanmu offers many morpho-syntactic, semantic, and phonological facts which support the above-mentioned argument, one of which I will explain below.

The distinction in the two forms given in (22) between phrases and words can be seen by examining whether adverbial modifiers can occur before expressions of the forms in (22). The contrast between (22a) and (22b) is exemplified by (23) and (24), which are shown in Duanmu 1998: 150.

(23) a. xin *de* shu
new DE book 'a new book'
b. hen xin *de* shu
very new DE book 'a very new book'
c. geng xin *de* shu
more new DE book 'a newer book'

(24) a. xin shu
new book 'a new book'
b. *hen xin shu
very new book 'a very new book'
c. *geng xin shu
more new book 'a newer book'

Xin de shu in (23a) is of the form (22a) and can be modified by adverbials, whereas *xin shu*

in (24a) is of the form (22b) and cannot be modified by adverbials. The unacceptability of adverbial modification in the case of (24) can be attributed to the violation of the Lexical Integrity Principle in (5). Furthermore, Duanmu 1998 indicates that in the case of the form (22b), it can be expanded into [[X M] N]. (25) illustrates this.

(25) [N N] → [[A N] N]
bu shou-tao [lan bu] shou-tao
'cloth glove' 'blue-cloth glove'

Example (25) shows that the A-N sequence *lan bu* is a compound, so it can occur inside a larger compound to give a compound of the form [[A-N] N].

Assuming that the above-mentioned analysis of Chinese by Duanmu 1998 is fundamentally correct, we can consider that in the case of Chinese, like Japanese, A-N expressions behaving as phrases and as words are clearly distinct in form and that Chinese, as with Japanese, has compounds which contain an A-N compound but not an A-N phrase.

2.3. Dutch

Compounds in Dutch are treated in detail in Booij 2002. I will investigate the properties of A-N expressions inside compounds in Dutch by referring to what is claimed by Booij 2002.[7]

First, I would like to give a brief explanation of agreement phenomena observed in the NP in Dutch. Prenominal adjectives take the inflectional ending *-e* except when the NP has the feature [+Indefinite, +Singular, +Neuter] (Donaldson 1997, Booi 2002), as seen in (26).

(26) a. het mooi*e* boek [+Definite, +Singular, +Neuter]
the nice book
'the nice book'
b. een mooi boek [+Indefinite, +Singular, +Neuter]
a nice book
'a nice book'

Booij 2002: 146 claims that A-N expressions occurring in the non-head position of a compound as in (27) are phrasal but not lexical, and that their NP status is confirmed by the fact that the inflectional suffix *-e*, which is only possible within NPs, is attached to the Adjective of an A-N expression inside a compound, as shown in (27).[8]

(27) a. [$_{NP}$ blot*e*- vrouwen] blad
nude-INFL women magazine 'nude women magazine'
b. [$_{NP}$ het*e*- lucht] ballon
hot-INFL air balloon 'hot air balloon'

Booij further mentions that phrasal A-N expressions, even when they are used as a non-head constituent of a compound, are not semantically lexicalized.

The additional fact referred to by Booij 2002 which proves the phrasal status of the A-N expressions inside the compounds in (27) is concerned with their stress pattern; that is, such A-N expressions have the stress pattern of a phrase, i.e., righthand stress, not the lefthand stress of a compound.

Compounds like (27) lead Booij 2002 to suppose that the morphological component of a grammar cannot be placed before the syntactic component, since the syntactic rule of A-N agreement must be allowed to apply within such compounds. Consequently, claiming that the grammar has to consist of the unordered modules, Booij proposes that the morphological module defines the set of well-formed words and that it should be stated in this module that the non-head part of N-N compounds can be an NP, to which the rules of the syntactic module apply.

Interestingly enough, however, Booij 2002: 147 points out that only the Adjective and the Noun are allowed to occur in the NP within a compound, by presenting the fact that compounds in (28) are impossible.

(28) a. *een [heel blot*e* vrouwen] blad
'a very nude women magazine'
b. *[vier donker*e* kleuren] druk
'four dark colour*s* print'
c. *[de oud*e* mannen] huis
'a the old men's home'

Booij proposes that it is the morphological module that states the above-mentioned restriction that the non-head NP in a compound has to consist only of the Adjective and the Noun. Booij suggests that this restriction at least partially has something to do with semantics. Example (28c) is considered to be unacceptable, because A-N expressions within compounds in general cannot be referential. However, it seems that what Booij 2002 does is only to mention the fact that in Dutch, there is such a restriction in the non-head NP within a compound. I consider that Booij's proposal does not account for why such a restriction has to be generally imposed on NPs within compounds.

My claim about A-N expressions inside compounds like the ones in (27) in Dutch is that they are intermediate between words and phrases, as is the case with English A-N expressions within compounds like the ones in (13)-(15). Such expressions in Dutch are regarded as phrasal, because i) prenominal adjectives inside them are inflected, ii) they are semantically compositional, and iii) they bear a phrasal stress. At the same time, they are regarded as lexical, because they follow the Lexical Integrity Principle in (5). The unacceptability of (28), I suppose, can be ascribed to violation of the Lexical Integrity Principle. As for the productivity of phrase-like A-N expressions inside compounds in Dutch, no data are available to me which show to what extent they are productively formed, though Booij 2001: 146 states that they are a productive pattern. So, in the case of Dutch, unlike in English, I cannot give the matter of productivity as a piece of evidence to prove the phrasal/lexical status of A-N expressions within compounds.

In Booi 2002 it is not mentioned whether A-N compounds can also occur inside larger compounds.[9] However, Jack Hoeksema (p.c.), after indicating the exocentric A-N compounds given below, states that they can occur in non-head positions in compounds, though he does not provide concrete examples.

(29) a. de roodhuid 'the red-skin = American Indian'
(cf. de rod*e* huid 'the red skin')
b. dat leeghoofd 'that empty-head = stupid person'
(cf. dat leg*e* hoofd 'that empty head')

2.4. A-N expressions inside compounds in English, Japanese, Chinese, and Dutch

In this section I have been concerned with compounds which contain A-N expressions in non-head positions, by focusing on four languages, that is, English, Japanese, Chinese, and Dutch. It has been made clear that A-N expressions which can occur inside compounds are of the following two types:

(30) a. fully lexical
b. neither fully phrasal nor fully lexical

It has also been shown that in English and Dutch, A-N expressions of both types given in (30) can occur inside compounds, while in Japanese and Chinese, only A-N expressions of the type (30a) can occur inside compounds.

3. A-N expressions inside compounds and a morphological category called Word Plus

We are now in a position to account for the intermediate status of A-N expressions inside compounds like the ones in (13)-(15) in English and the ones in (27) in Dutch. In Section 3.2, I will propose that such expressions belong to a novel morphological category called 'Word Plus' (W^+), for which Kageyama 2001 argues convincingly. It is generally assumed that words constitute the category of X^0 and that morphology concerns the internal structures of X^0 or below, whereas syntax generates the structures of X' and XP. Kageyama, however, suggests that the category W^+ should be introduced, which is larger than Word or X^0 but nonetheless is a morphological category distinct from the syntactic categories X' and XP.

3.1. Morphological category W^+ proposed in Kageyama 2001

Kageyama 2001 deals with many prefixed words and compounds in Japanese which bear a phrasal accent with a slight pause or minor phrase (indicated by "/" below) after the prefixes or between the non-head and head members of compounds, as in (31), and investigates the problem of whether such words can be truly identified as words or not.

(31) a. zen/sooridaizin 'the ex-Prime Minister,' doo/daigaku 'the above-mentioned university,' gen/kaityoo 'the current president'
b. ryooriten/ keieisya 'restaurant manager,' sinki/kakutyoo-keikaku 'new expansion

plan', missyon-kei/siritsu-daigaku 'private universities founded by missionaries'

It is contended by Kageyama that prefixed words and compounds like the ones in (31), in spite of their phrase-like properties in terms of phonology, maintain their lexical integrity or syntactic atomicity and therefore that such expressions are qualified to be classified as words. Kageyama gives the following two observations which are regarded as manifestations of the Lexical Integrity Principle in (5).[10] First, they obey the 'No Phrase Constraint' (Roeper and Siegel 1978), and therefore, neither the head nor the non-head of such expressions can incorporate NPs involving genitives, adjectives, or other modifiers, as illustrated by the examples in (32) and (33) (p. 250).

(32) a. kaku / [$_{N}$ tihoo- tosi] 'each provincial city'
each province city
b. *kaku / [$_{NP}$ tihoo-*no* tosi] 'each provincial city'
each province-GEN city
(33) a. zenkoku / [gassyoo-konkuuru] 'All-Japan chorus contest'
All -Japan chorus contest
b. *zenkoku / [$_{NP}$ gassyoo-*no* konkuuru] 'All-Japan chorus contest'
All-Japan chorus -GEN contest

The second manifestation of the Lexical Integrity Principle indicated by Kageyama 2001: 251 is the impossibility of modifying a part of a phrase-like word with an external adjective or relative clause, as indicated by (34).

(34) a. Tookyoo-ni iru [booeki-gaisya / syatyoo]
Tokyo-in is(animate) trading-company president
'the ex-president of a trading company, who is in Tokyo'
b. *Tookyoo-ni aru [booeki-gaisya / syatyoo]
Tokyo-in is(inanimate) trading-company president
'the ex-president of a trading company located in Tokyo'

Kageyama's claim is that phrase-like prefixed words and compounds in Japanese are identified as words which belong to the above-mentioned morphological category W^{+}, because, as has been indicated, such expressions obey the Lexical Integrity Principle in (5). The most interesting Japanese W^{+} words treated by Kageyama are words prefixed by *doo-* 'above-mentioned,' as in *doo-daigaku* 'the above-mentioned university' given in (31). According to Kageyama, because the prefix *doo-* has the function of referring to a previously mentioned entity (for details, see Kageyama 2001: 256-257), the syntactic principle which specifies the reference of *doo*-words is considered to access the internal structure of such words, which is illustrated by the following (p. 258):

(35) *New South Wales-syuu*$_{i}$ o osotta yamakazi wa yooyaku tinka no kizasi o misete-kita.

[*Doo* / *syuu*$_{i}$ / syooboosyo-honbu] wa kinkyuu-zitai-sengen o kaizyo-sita.
'The forest fire which raged in New South Wales, Australia, is finally going to be contained. The fire department headquarters of the above-mentioned state has lifted the emergency measures.'

In the case of (35), the Noun Phrase *New South Wales-syuu* 'New South Wales' behaves as an antecedent of the word *syuu* 'state' inside the larger compound *doo-syuu syooboosyo-honbu* 'the fire department headquarters of the above-mentioned state.'

As candidates for W^+ words in English, Kageyama 2001: 269-272 refers to compounds with a phrasal accent like (36) and also to compounds which realize plurals or comparative and superlative inflections on their first members as in (37) and (38).

(36) world-fámous, crystal-cléar, student rebéllion
(37) lookers-on, passers-by, hangers-on
(38) well-known, better-known/ best-known, *more well-known/ *most well-known
kind-hearted, kinder-hearted

To sum up, Kageyama proposes a new morphological category W^+ in order to account for the fact that both in Japanese and English, there are some derivatives and compounds which maintain lexical integrity but nonetheless have phrasal properties at the same time. According to Kageyama, phonological and syntactic properties both prove the phrasal character of W^+ words. Such words display a phrasal accent, while internal structures of them can be made available to some particular syntactic rules or principles.

3.2. Category assignment of two types of A-N expressions inside compounds

For A-N expressions which can occur inside compounds, I have explained that there are two types, as indicated in (30a) and (30b). Here I would like to claim that compounds with A-N expressions of the type shown in (30a), i.e., compounds with a non-head A-N expression which is a compound itself, have the structure (39) and that compounds with A-N expressions of the type (30b), i.e., compounds with an A-N expression which has an intermediate status, are assigned the structure (40).

(39) $[_{N0}[_{N0}\ A^0\ N^0]\ \ N^0]$
(40) $[_{N0}[_{N+}\ A^0\ N^0]\ \ N^0]$

In the structure (39) A and N constitute N^0, whereas in (40) they constitute N^+. I propose that in the case of Japanese and Chinese, only the structure (39) is available for compounds with A-N expressions, whereas in the case of English and Dutch, not only the structure (39) but also the structure (40) are available for such compounds.

The category N^+ is a morphological category and not a syntactic one. Therefore, A-N expressions belonging to this category, such as the ones in (13)-(15) in English and the ones in (27) in Dutch, follow the Lexical Integrity Principle in (5), and at least in the case of English, it is evident that such expressions are not completely productive, as exemplified by

(17)-(19). Nonetheless, in the case of Dutch, due to the N^+ status of such A-N expressions, the syntactic rule of agreement can access their internal structures; by the application of this syntactic rule to A-N expressions assigned N^+ which occur inside compounds, prenominal adjectives in such A-N expressions take the ending *-e* except when N^+ bears specific features mentioned in Section 2.3. In the case of English, phrase-like A-N expressions inside compounds do not undergo a syntactic rule of agreement or any other syntactic rules. However, we should consider that such expressions, like the corresponding A-N expressions in Dutch, belong to the category N^+, because they are neither fully syntactic nor fully lexical but has properties of both.

Moreover, the difference in semantic compositionality and in stress contour between phrase-like A-N expressions and lexical ones can be reduced to the difference between the categories N^+ and N^0, to which these two expressions respectively are assigned.

It would be interesting to know whether there are any languages which tolerate compounds containing an A-N expression which violates the Lexical Integrity Principle, namely, compounds containing an A-N expression which belongs to a syntactic phrase (N' or NP).

4. Summary and residual problems

I have argued for the existence of A-N expressions inside compounds which are neither fully phrasal nor fully lexical but are given intermediate status, and have claimed that such A-N expressions are assigned the morphological category W^+ (in this case, N^+) as proposed by Kageyama 2001.

I have also suggested that there are two types of languages with respect to the structures which A-N expressions inside compounds have. The first type is languages in which the category assigned to A-N expressions inside compounds is limited to the N^0, as in Japanese and Chinese, and the second type is languages in which the category assigned to such A-N expressions can be extended into a larger category N^+, as in English and Dutch. However, it is not clear to me at present what other syntactic or morphological phenomenon or phenomena this typology correlates with.

Another matter to consider is the role of pragmatics in word formation. I consider that Sproat 1993 is on the right track when he suggests that pragmatics is involved in judging whether a certain compound with an A-N expression sounds natural or not. As stated above, the compound *[green car] driver* in (18) is less acceptable than the compound *[small car] driver* in (13a). However, according to Sproat 1993, the former would sound rather natural if it is used in a certain context (see p. 251).

Sproat (p.c.) applies the above-mentioned reasoning to the compounds given in (21) as well, which are judged as unacceptable in this paper and in Shimamura 2001. He agrees that the compounds *[very small car] driver* and *[small green car] driver*, for example, are less available than the compounds *[small car] driver* in (13a) and *[green car] drive* in (18), because 'very small car' or 'small green car' is not likely to be an interesting category, but he adds that these might be regarded as interesting categories in an appropriate situation. He told me that there is a Web site with the title shown in (41):

(41) [Very Large Array] Home Page

(Hoeksema (p.c.) also informed me of the following Web site, which is similar to (41): *[very large] telescope survey*.) Sproat, therefore, claims that there is nothing structurally wrong with *[very small car] driver* or *[small green car] driver* in (21), and states that in appropriate contexts one could expect to find similar constructions. At present, I cannot offer any well-grounded argument against this proposal of Sproat's. We will have to consider seriously the involvement of pragmatics in word formation.

Let me turn back to the proposal of this paper that because of their W^+ status, phrase-like A-N expressions inside compounds must obey the Lexical Integrity Principle in (5). Following this proposal, how should we account for the fact that even though in rare cases, there do occur compounds like (41), although A-N expressions inside such compounds violate the Lexical Integrity Principle in (5)? I suppose that there are some cases in which the category N^+ is extended into the syntactic category N' or NP, because of its Janus-faced nature. In contrast, in the case of languages like Japanese and Chinese in which the category assigned A-N expressions inside compounds is limited to N^0, compounds with a genuinely phrasal A-N expression (e.g., (4)) would be rejected in any context.

Finally, I would like to refer to a theoretical implication of the postulation of W^+ as a morphological category. If it is truly adequate to suppose the existence of this category, it will lead us to consider that some syntactic rules can access the parts of words. As examples of such syntactic rules, Kageyama 2001 points out a rule in Japanese which defines the anaphoric relation between an NP and a derived word with the prefix *doo*-, which has a referential property, and rules in English which specify comparative and superlative inflections or plurals. I have mentioned that the rule of A-N agreement in Dutch is one such syntactic rule. It should be made clear what kinds of syntactic rules can refer to parts of a W^+ word in violation of the Lexical Integrity Principle. If such syntactic rules can be delimited in future, it will mean that it is necessary to weaken the Lexical Integrity Principle in its usual sense given in (5), as is done by Haspelmath 2002: 162, for example. In this paper I cannot afford to propose explicitly a weakened version of the Principle. But it is certain that the existence of W^+ implies that the difference between morphology and syntax is not so sharp as has been assumed by the Lexicalist Hypothesis (e.g., Di Sciullo and Williams 1987, Bresnan and Mchombo 1995).

Notes

***** Earlier versions of this paper were presented at the Lexicon Study Circle on February 11, 2003 and at the Tokyo Circle of English Linguistics on April 19, 2003. I would like to express my gratitude to the members of the two Circles for their invaluable comments. I am also indebted to Professor Richard Sproat and Jack Hoeksema for their helpful comments on my paper Shimamura 2001. I am also grateful to Scott Reynolds for suggesting stylistic improvements.

[1] *Rendaku* is a phenomenon in which the initial consonant of the second element in a compound

undergoes voicing. This phenomenon is observed in Yamato compounds, that is, compounds which are of Japanese origin . (The vocabulary of Japanese is divided into three morpheme classes, one of which is Yamato.) This *rendaku* phenomenon obeys Lyman's Law. For details, see Otsu 1980, Itô and Mester 1986, and Vance 1987.

[2] According to Chomsky and Halle 1968, the sequence *black-board eraser* can be analyzed as [$_{NP}$black [$_{N}$board eraser]], where *board eraser* constitutes a compound modified by *black*. This paper is not concerned with such a construction, since A-N expressions in such constructions do not form a unit.

[3] In this paper syntactic A-N expressions in Japanese are assumed to have an NP structure in which a prenominal A modifiers a head N. But an alternative assumption would be that syntactic A-N expressions are assigned a structure with a relative clause. Various discussions concerning this problem are found in Nishiyama 1999, Yamakido 2001, and Namai 2002. I leave this problem open for future research.

[4] A-N expressions such as the following which consist of a relational adjective and a noun inside a compound also are considered to be compounds: *[medical building] site*, *[solar house] planning*, *[adjectival phrase] insertion*, *[atomic bomb] test*, *[criminal law] revision*. As for such compounds, see Levi 1978, Beard 1995, Liberman and Sproat 1992, and Bennett 2002. However, because relational adjectives seem to differ in some ways from ordinary adjectives, I do not treat such A-N compounds in this paper, though it seems that what is claimed in this paper applies to them as well.

[5] Bennett 2002 states that the occurrence of phrase-like A-N expressions within compounds is more restricted in synthetic compounds than in root compounds, pointing out that, as illustrated by the following, A-N expressions are often awkward: ?*[young-lion] tamer*, ?*[new-car] driver*, ?*[long-letter] writer*. It is not clear to me whether there is such a difference between the two types of compounds.

[6] The Lexical Integrity Principle predicts that coordination also is impossible in phrase-like A-N expressions inside compounds. However, coordination is a moot point. The compound *[[new and late model] car]* was judged as unacceptable by some native speakers of English, and I regarded this compound as unacceptable in Shimamura 2001. But later I discovered that there are native speakers who accept this compound. In fact the following compounds containing A-N expressions with coordinated adjectives are found in Okada 1999: 343: *[higher- and lower-status] members*, *[soft- and hard-surface] cleaning*.

[7] I owe it to Professor Jack Hoeksema that I can pay attention to Dutch A-N expressions within compounds.

[8] Booij 2002: 146 further mentions that the [Quantifier-Noun] sequence as well forms an NP and that it can occur inside compounds.

[9] Booij 2002: 146-147 states that Dutch has compounds of the form [$_{N}$ [$_{N}$ A-N]N] (e.g., *[breed band] antenne* 'broadband aerial'), where the [A-N] sequence is a compound, not a syntactic phrase, but according to him, such [A-N] sequences are only productive as the lefthand constituent of a compound.

[10] In Kageyama 2001, identity deletion in coordinate sentences is regarded as the third observation which indicates that words like (31) obey the Lexical Integrity Principle. My judgment of sentences relevant to this phenomenon is different from Kageyama's. So, I do not introduce it in this paper.

References

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

Bauer, Laurie 1998. "When a sequence of two nouns a compound in English?" *English Language and*

Linguistics 2, 65-86.

Beard, Robert. 1995. *Lexeme-Morpheme Base Morphology*. Albany, New York: State University of New York Press.

Bennett, Paul. 2002. "English adjective-noun compounds and related constructions." *GEMA: Online Journal of Language Studies* 2.

Booij, Geert. 2002. *The Morphology of Dutch*. Oxford: Oxford University Press.

Bresnan, Joan and Sam A. Mchombo. 1995. "The lexical integrity principle: Evidence from Bantu." *Natural Language and Linguistic Theory* 13, 181-254.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. Cambridge, Mass.: MIT Press.

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

Donaldson, Bruce. 1997. *Dutch: A Comprehensive Grammar*. London: Routledge.

Downing, Pamela. 1977. "On the creation and use of English compound nouns." *Language* 53, 810-842.

Duanmu, San. 1998. "Wordhood in Chinese." In *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, ed. Jerome L. Packard ,135-196. Berlin: Mouton de Gruyter.

Haspelmath, Martin. 2002. *Understanding Morphology*. London: Arnold.

Itô, Junko and R. Armin Mester. 1995. "Japanese phonology." In *The Handbook of Phonological Theory*, ed. John A. Goldsmith, 817-838. Cambridge, Mass.: Blackwell.

Kageyama, Taro. 2001. "Word plus: The intersection of words and phrases." In *Issues in Japanese Phonology and Morphology*, ed. Jeroen van de Weijer and Tetsuo Nishihara, 245-276. Berlin: Mouton de Gruyter.

Levi, Judith N. 1978. *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.

Lieber, Rochelle. 1992. *Deconstructing Morphology: Word Formation in Syntactic Theory*. Chicago: University of Chicago Press.

Liberman, Mark and Richard Sproat. 1992. "The stress and structure of modified noun phrases in English." In *Lexical Matters*, ed. Ivan Sag and Anna Szabolcsi, 131-181. Stanford: CSLI.

Namai, Kenichi. 2002. "The word status of Japanese adjectives." *Linguistic Inquiry* 33, 340-349.

Nishiyama, Kunio. 1999. "Adjectives and the copulas in Japanese." *Journal of East Asian Linguistics* 8, 183-222.

Okada, Sadayuki. 1999. "On the conjoinability of affixal morphemes in English." *Word* 50, 339-363.

Otsu, Yukio. 1980. "Some aspects of rendaku in Japanese and related problems." *MIT Working Papers in Linguistics, Vol.2, Theoretical Issues in Japanese Linguistics*, 207-227.

Roeper, Thomas. 1988. "Compound syntax and head movement." In *Yearbook of Morphology*, ed. Geert Booij and Jaap van Marle, 187-228.

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

Sadler, Louisa and Douglas J. Arnold. 1994. "Prenominal adjectives and the phrasal/lexical distinction." *Journal of Linguistics* 30, 187-226.

Shimamura, Reiko. 2001. "The A-N expression within the compound and the phrase/word distinction." *Grant-in-Aid for COE Research Report* (5), 163-175. Kanda University of International Studies.

Sproat, Richard. 1993. "Morphological non-separation revisited: A review of R. Lieber's *Deconstructing*

Morphology." In *Yearbook of Morphology 1992*, ed. Geert Booij and Jaap van Marle, 235-258. Dordrecht: Kluwer.

Vance, Timothy, 1987. *An Introduction to Japanese Phonology*. Albany, New York: State University of New York Press.

Yamakido, Hiroko. 2001. "Japanese attributive adjectives are not (all) relative clauses." *WCCFL* 19, 588-602.

The Department of English
Tsuda College
2-1-1 Tsuda-machi, Kodaira-shi
Tokyo 187-8577
Japan
[simamura@tsuda.ac.jp]