

Word plus: The intersection of words and phrases

Taro Kageyama, Kwansei Gakuin University

0. Introduction

The autonomy of morphology with respect to the other modules of the grammar poses a perennial problem in morphological research. In particular, the relation between morphology and syntax, repeatedly debated under the rubric of various versions of the lexicalist hypothesis, is so complex that it is necessary to accurately pinpoint the real issues involved in it before we enter into fruitful discussions of any particular topic. To my mind, the major issues are roughly divided into two groups. One is what we may call the “component problem”, namely whether or to what extent syntax and morphology interact with each other in the system of grammar. There are a variety of proposals on this issue (see Borer (1998) for an overview), ranging from the strong lexicalist position which grants morphology its complete autonomy (Di Sciullo & Williams 1987) to the strictly syntactic position which integrates all productive morphology into syntax (Lieber 1992, Halle & Marantz 1993), with mixed theories in between allowing word formation both in the syntax and in the lexical component (Baker 1988, Shibatani & Kageyama 1988). The other major issue concerning the morphology-syntax interface is the problem of “lexical integrity” or “syntactic atomicity”, namely whether or to what extent syntactic operations can have access to the internal structure of a word. Here again, diverse views have been expounded (Di Sciullo & Williams 1987, Bresnan & Mchombo 1992).

The present paper attempts to shed new light on these two problems by closely examining a certain type of complex expressions in Japanese which has been largely ignored in the literature. The expressions we are going to probe are exemplified by (1).

- (1) a. zen | gaimu-daizin ‘ex-foreign minister’, han | taisei ‘anti-establishment’
b. booeki-gaisya | syatyoo ‘president of a trading company’, tokkyuu | teisya-eki
‘stations at which special express trains stop’
c. moto | booeki-gaisya | syatyoo ‘former president of a trading company’

(1a) illustrates prefixed expressions and (1b) compound expressions, and the two types of expressions can be mixed up as in (1c). What is most striking is that unlike regular lexical compounds which are pronounced with a sequence of high-pitched moras, the expressions in (1) bear a phrase-like accent with a slight pause or minor phrase boundary (indicated by “ | ”) after the prefixes or between the members of compounds. An obvious question is whether such

expressions are identified as words or phrases.

In section 1, I will argue, along with Poser (1990), Kubozono (1995), and Kageyama (1982; 1993), that those phrase-like expressions indeed constitute words because of their syntactic atomicity. Despite their word-status, section 2 demonstrates that those expressions are highly similar to phrases in terms of the internal semantic relations and syntactic behavior allowing sentence-level anaphora. In a nutshell, expressions like those in (1) are identified as words on the morphological side but behave like phrases otherwise. To resolve this contradiction, section 3 proposes a novel category called “Word Plus” (W^+), which is larger than Word or X^0 but nonetheless constitutes a morphological object distinct from a syntactic phrase like X' . Section 4 suggests the universal validity of the proposed W^+ category by considering English compounds with rear stress like *looker-on* and *well-known*.

The proposed W^+ category will have non-trivial implications for the architecture of morphology. It has been customarily assumed that words constitute the category of X^0 , and that morphology concerns the internal structures of X^0 or below whereas syntax builds up phrases and sentences on the basis of X^0 . In other words, X^0 is held to be the boundary that separates morphology from syntax. I will call this the “category problem”. The category problem has not been questioned seriously even by researchers who have attempted to make the sublexical categories more precise. For example, Selkirk (1982), who proposes the category “Root” which occurs below “Word”, and Ackema (1999), who develops an X-bar morphology postulating X^{-2} and X^{-1} below X^0 , nonetheless hold X^0 or Word to be the maximal domain of morphology. This paper will show that the proposed W^+ category rather than X^0 is the maximal projection of morphological structure.

1. The lexical integrity of phrase-like words

In this section it is shown that phrase-like expressions like those illustrated in (1) are equipped with syntactic atomicity, thus qualifying as genuine “words”. Our argument starts with prefixed expressions like (1a) and then moves on to compound and mixed expressions like (1b) and (1c).

It is well known that Japanese has a peculiar set of prefixes which exhibit a phrasal pitch contour instead of the usual lexical accent (Kageyama 1982, Poser 1990). Representative examples of these prefixes (dubbed “Aoyagi prefixes” by Poser (1990) after the Japanese grammarian S. Aoyagi) are given in (2), where the moras indicated by capital letters are pronounced with a high pitch.

(2) Phrase-like prefixes:

hon- ‘this, the present’: HOn | daIGAKU ‘this university’

moto- ‘former, ex-’: MOto | soORIDAizin ‘a one-time Prime Minister’

zen- ‘previous, immediately preceding’: ZEn | soORIDAizin ‘the ex-Prime Minister’

gen- ‘current’: GEn | kaITYOO ‘the current president’

kaku- 'each': KAKu | daIGAKU 'each university'
 boo- 'a certain': BOO | tyoMEESAKka 'a certain famous writer'
 doo- 'above-mentioned': DOO | daIGAKU 'the above mentioned university'
 zen- 'all, whole': ZEn | koKUMIN 'the whole nation'
 ryoo- 'both': RYOo | daIGAKU 'both universities'
 ko- 'deceased': KO | uEMURA-si 'the late Mr. Uemura'
 han- 'anti-': HAn | taISEE 'anti-establishment'
 hi- 'non-': HI | yoOROPPA-GEngo 'non-European languages'

These prefixes are phonologically peculiar not only in bearing an accent on their own but also in being followed by a slight pause or minor phrase boundary, indicated by the symbol " | ".

Notwithstanding the phrase-like accent contour, there are cogent reasons to assume that such complex expressions are full-fledged words. First, the prefixes under discussion cannot be attached to NPs involving adjectival or other modifiers (Poser 1990, Kageyama 1993).

- (3) a. *boo-[_{NP} yumeina haiyuu] 'a certain famous actor'
 certain-[famous actor]
 b. *zen-[_{NP} syuugi'in-no giin] 'a former member of the House of Representatives'
 former-[House-of-Representatives of member]
 c. *kaku-[_{NP} tihoo-no tosi] 'each provincial city'
 each-[province-of city]

The expressions in (3) are ungrammatical because of the violation of the "No Phrase Constraint" (Roeper & Siegel 1978), which is a manifestation of the general principle of lexical integrity or syntactic atomicity. (The only exception I am aware of is the prefix moto- 'formerly, ex-', which may take a genitive phrase, as in moto [puro-yakyuu no sensyu] 'formerly a professional baseball player'. The idiosyncrasies of this prefix will be delineated in section 3.) Particularly noteworthy about (3) is the fact that the ungrammaticality is removed if their NP parts are changed into compound nouns, as shown in (4).

- (4) a. boo-[_N yumei-haiyuu] (same meaning as (3a))
 b. zen-[_N syuugi'in-giin] (same meaning as (3b))
 c. kaku-[_N tihoo-tosi] (same meaning as (3c))

Second, the lexical integrity of the expressions at issue is shown by the rule of identity deletion in coordinate sentences. This rule can delete up to part of phrases but does not intrude into words, as illustrated in (5) (Kageyama 1989; 1999).

- (5) a. Ken-wa sukiyaki-o ~~tabe~~, Naomi-wa susi-o ~~tabe~~-ta.
 K-TOP sukiyaki-ACC eat N-TOP sushi-ACC eat-PAST
 'Ken ate sukiyaki, and Naomi sushi.'
 b. Ken-wa [_{NP} Furansu-no kuruma]-o ~~kai~~, Naomi-wa Itaria-no kuruma-o kat-ta.
 K-TOP France-GEN car-ACC buy N-TOP Italy-GEN car-ACC buy-PAST
 'Ken bought a car made in France, and Naomi, one made in Italy.'

- c. *Ken-wa [_N inu-goya]-o tukuri, Naomi-wa [usagi-goya]-o tukut-ta.
 K.-TOP [dog-house]-ACC make N.-TOP [rabbit-house]-ACC make-PAST
 'Ken made a doghouse, and Naomi a rabbit-house.'

The deletion rule, which can erase the head noun of the NP structure in (5b), is forbidden to affect only the second member of the compound noun inu-goya 'doghouse' in (5c). In the same vein, this deletion cannot destroy the integrity of the complex expressions involving the phrase-like prefixes. Thus the same degree of ungrammaticality as (5c) is exhibited by (6a) below, where only the base is deleted leaving the prefix behind. This should be contrasted with the full grammaticality of (6b), where the prefixed words in (6a) are replaced by semantically similar NPs.

- (6) a. *A-san-wa [_N gen-kaityoo]-to ~~siriai-de~~, B-san-wa [_N zen-kaityoo]-to
 A-Mr-TOP [current-president]-with acquainted-is, B-Mr-TOP [ex-president]-with
 siriai-da.
 acquainted-is
 'Mr. A is acquainted with the current president, and Mr. B is acquainted with
 the ex-president.'
- b. A-san-wa [_{NP} genzai-no kaityoo]-to ~~siriai-de~~, B-san-wa
 A-Mr-TOP [current president]-with acquainted-is, B-Mr-TOP
 [_{NP} mae-no kaityoo]-to siriai-da.
 [former GEN president]-with acquainted-is

A third indicator of the syntactic atomicity of the phrase-like expressions is the impossibility of modifying a word-internal element with an external adjective or relative clause. Compare (7a) with (7b).

- (7) a. Tookyoo-ni iru [moto | booeeki-gaisya | syatyoo]
 Tokyo-in is(animate) ex- trading-company president
 'the ex-president of a trading company, who is in Tokyo'
- b. *Tookyoo-ni aru [moto | booeeki-gaisya | syatyoo]
 Tokyo-in is(inanimate) ex- trading-company president

intended meaning: 'the ex-president of a trading company which is located in Tokyo'

(7a) is impeccable, since the relative clause with the verb iru 'be', said of an animate subject, modifies the following expression as a whole. In contrast, (7b) is completely ruled out because the relative clause with the verb aru 'be', selecting an inanimate subject, is intended to modify only the internal element booeeki-gaisya 'trading company'. This is parallel to the infelicity of, for example, *a very [darkroom], where the adverb very cannot modify the adjective dark in the compound noun darkroom.

The preceding three observations clearly indicate that the prefixed expressions like those in (2) make up syntactically inviolable units and are therefore qualified as words. Furthermore, the complex expressions in question are amenable to lexical or morphological conditions that

are generally observed with word formation.

As is well known, word formation processes are lexically governed, restricted notably by the vocabulary strata determined by the origin of individual words. In Japanese word formation, there is a pronounced tendency for a native morpheme to combine with another morpheme of native origin and for a Sino-Japanese morpheme to combine with another Sino-Japanese morpheme, although words composed of mixed vocabulary strata are occasionally encountered. This combinative restriction is illustrated by (8).

- | | | |
|--------|---------------------------------|--------------------------------|
| (8) a. | roodoo-sya 'laborer' | *roodoo-mono |
| | labor [S-J] - person [S-J] | labor [S-J] - person [native] |
| a'. | hataraki-mono 'hard worker' | *hataraki-sya |
| | work [native] - person [native] | work [native] - person [S-J] |
| b. | dai-kyoositu 'big classroom' | *oo-kyoositu 'big classroom' |
| | big [S-J] - classroom [S-J] | big [native] - classroom [S-J] |
| b'. | oo-beya 'big room' | *dai-beya 'big room' |
| | big [native] - room [native] | big [S-J] - room [native] |

Now it is found that the same restriction pertains to the phrase-like expressions as well. Since all of the phrase-like prefixes except the native moto- 'former' are Sino-Japanese, they are most comfortably attached to Sino-Japanese (or Western) loan words.

- | | |
|--------|---|
| (9) a. | kaku gakkoo 'each school' |
| | each [S-J] school [S-J] |
| a'. | *kaku manabiya 'each school' |
| | each [S-J] learning-place [native] |
| b. | doo syuukai 'the above-mentioned assembly' |
| | above-mentioned [S-J] assembly [S-J] |
| b'. | *doo atumari 'the above-mentioned assembly' |
| | above-mentioned [S-J] gathering [native] |

It seems that only the native moto- exceptionally meshes with both Sino-Japanese and native bases.

- | | |
|---------|--|
| (10) a. | moto kyoositu 'formerly a classroom' |
| | former classroom [S-J] |
| b. | moto sigotobeya 'formerly a work-room' |
| | former work-room [native] |

It will be observed in section 3 that this exceptional behavior of moto- has syntactic repercussions.

The lexical governedness of phrase-like words is further illuminated by the idiosyncratic restrictions on the selection of combined words.

- | | |
|---------|-----------------------------------|
| (11) a. | zen syokuin 'all the staff' |
| b. | itibu syokuin 'part of the staff' |

c. *hanbun | syokuin ‘half the staff’

cf. hanbu no syokuin ‘half of the staff’ (no=genitive particle)

d. *kahansuu | syokuin ‘majority of the staff’

cf. kahansuu no syokuin ‘majority of the staff’

In (11), only zen- ‘all’ and itibu- ‘part’ are eligible for word formation; other semantically similar words like hanbun ‘half’ and kahansuu ‘majority’ can fit in only with a phrase structure supported by the genitive particle no.

All in all, the phrase-like prefixes in (2) impose lexically idiosyncratic restraints on the bases to which they attach—a phenomenon which is characteristic of word formation but is alien to the formation of syntactic phrases or sentences.

Additional support for the wordhood of the phrase-like expressions derives from the fact that they can be embedded in larger compounds, as exemplified by (12).

(12) a. [[kokuritu-daigaku] | [zen | gakutyoo]]

national-university former president

‘former president of a national university’

b. [boo | [[kokuritu-daigaku] | [zen | gakutyoo]]]

certain [national-university former president]

‘a certain ex-president of a national university’

Note the iteration of minor phrase boundaries at the loci signified by “|”.

The fact that prefixed and compound words are freely embedded inside each other indicates that the prefixation and compound formation are of the same morphological nature, most probably being situated at the same level (in the sense of level-ordered morphology). While the prefixes with phrasal accent are naturally limited to a particular set of morphemes, compound words with the same characteristic appear to be created with unlimited productivity. Below in (13) are given additional examples of phrase-like compounds.

(13) Compound words pronounced with phrasal accent

ryooriten | keieisya ‘restaurant proprietor’, tokkyuu | teisya-eki ‘station at which limited express trains stop’, (kono hikooki no) Heathrow-kuukoo | tootyaku-zikoku ‘the scheduled time of (this airplane’s) arriving at Heathrow Airport’, syugo | kuriage-kisoku ‘the subject raising rule’, naizyu | kakudai-seisaku ‘policy to boost domestic demand’, California-syuu | Sacramento ‘Sacramento, California’, mission-kei | siritu-daigaku ‘private universities founded by missionaries’, gengogakka | kyoodoo-kenkyuusitu ‘common room of the linguistics department’, kigenzen | 226-nen ‘the year 226 B.C.’

The minor phrase boundary observed between the two members of each of these compounds is reminiscent of an analogous phonological break (indicated by “:”) inherent in what Shibatani and Kageyama (1988) call “post-syntactic compounds”, namely N-VN compounds which are created most plausibly at S-structure after the transformational cycle.

Kageyama (1993) points out that despite the phonological similarity, compounds like those in (13) are different from post-syntactic compounds. First, subject honorification can be applied to the head of post-syntactic compounds, but never to that of the (13)-type compounds.

(14) subject honorifics on the head

a. post-syntactic:

sensei-ga [yooroppa : go-taizai] no ori

teacher-NOM [Europe : HONORIFIC-stay] GEN occasion

‘while the teacher stayed in Europe’

b. (13)-type compound:

*yooroppa | go-taizai-kan

Europe HONORIFIC-stay-period

‘the period of the stay in Europe’

Second, post-syntactic compounds may sometimes allow phrases in the first member, whereas (13)-type compounds strictly exclude phrases from appearing inside.

(15) phrasal expressions inside

a. post-syntactic:

[[kono zikken] : syuuryoo]-go ni (Shibatani & Kageyama 1988)

[[this experiment] : finish] after DAT

‘after this experiment is completed’

b. (13)-type compound:

*[sono ryooriten] | keieisya

that restaurant manager

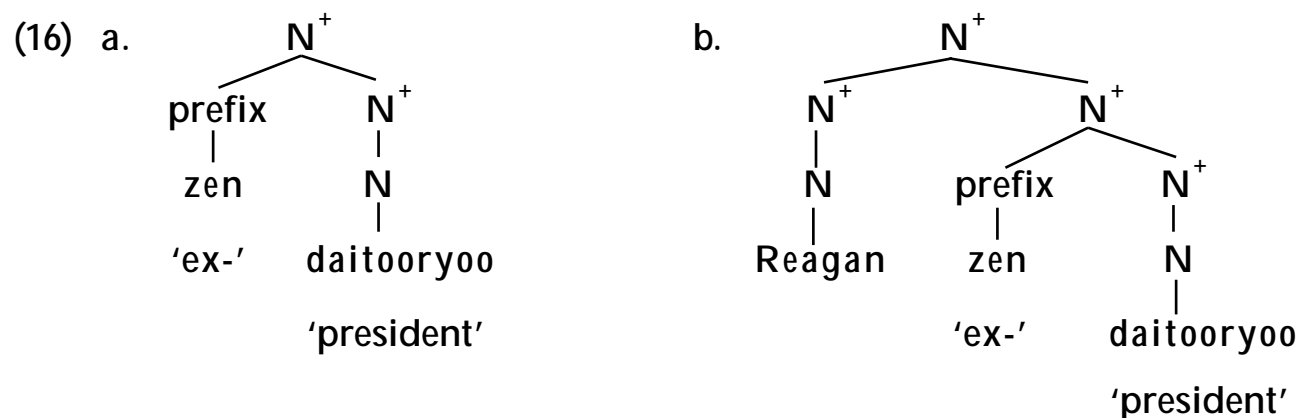
‘manager of that restaurant’

(15b) is ungrammatical on the intended reading where the determiner sono ‘that’ modifies ryooriten ‘restaurant’, although the same string will be perfectly acceptable if the determiner is detached from the compound structure and interpreted as modifying the whole compound, as in “that restaurant manager”.

The occurrence of honorific markers and determiners inside post-syntactic compounds was part of the motivation for Shibatani & Kageyama (1988) to locate their formation after the syntax rather than in the lexicon. Viewed in this light, the rejection of honorifics and determiners in (14b) and (15b) indicates that the (13)-type compounds are not (post-)syntactic but lexical.

This section has argued for the word status of the peculiar phrase-like expressions on the basis of phenomena involving lexical integrity and lexical idiosyncrasies. Because of the word status, we will refer to the phrase-like words, both prefixed and compounded, as Word⁺ (word plus). This is intended as a new morphological category that is larger than ordinary words but still belongs to the morphological as opposed to the syntactic domain. For example, the morphological structures of zen | daitooryoo ‘ex-President’ and Reagan | zen | daitooryoo ‘ex-

President Reagan' will be represented as follows.



In (16a), the phrase-like prefix zen- 'ex-' selects N^+ as its base. This c-selection property need be specified in the lexical entry of this particular prefix, along the lines of Lieber's (1980) subcategorization theory of affixes. The structure in (16b) illustrates a more complex W^+ involving the compounding of Reagan and zen | daitooryoo. Here, Reagan, which normally has the status of N, is projected to N^+ to be compounded with another word of the same category.

To summarize the foregoing discussion, this section has established the "Word" part of "Word⁺". In the subsequent section, we will turn our attention to the "Plus" part of Word⁺ and elucidate what extra properties W^+ has which are not shared by ordinary words.

2. The phrasal character of Word plus

Having established the lexical integrity of W^+ expressions, we now encounter a seeming contradiction that those expressions, albeit morphologically identified as words, show phrasal behavior with respect to their internal phonology, semantics, and syntax.

First, the phonological properties. As has been repeatedly mentioned, Word⁺ expressions carry phrasal accent, with a slight pause inside. This phrasal accent, as opposed to the so-called lexical accent observed with regular derived or compound words, provides an ostensible indication that W^+ might belong to the syntax. Rendaku or sequential voicing is another phonological phenomenon that is observed inside lexical words (17a) but can never apply to what we identify as W^+ (17b).

- (17) a. lexical compound: boeki-gaisya [<-- kaisya 'company']
 b. Word⁺: kaku | kaisya 'each company', zen | kaisya 'all companies'

The phrase-like accent contour and the absence of sequential voicing suggest that the members of W^+ expressions are independent units connected with each other only loosely.

The phonologically loose connection of W^+ is correlated with the internal semantics. Compare the internal semantic relations of a regular compound (18a) and a W^+ expression (18b).

- (18) a. regular compound: sokkyuu-toosyu 'fast-ball pitcher'
 b. W^+ : moto | kyozingun | toosyu 'formerly a pitcher of the Giants team'

Regular compound nouns designate a general type of noun with the modifier providing a

further specification of the head noun. Thus, *sokkyuu-toosyu* ‘fast-ball pitcher’, (18a), is a type of baseball pitcher characterized by the high speed of the balls he throws. On the other hand, the W^+ expression in (18b) does not represent a general type of pitcher but a particular token satisfying the description, with the modifier *moto | kyozingun* ‘former Giant’ predicating the head noun. This is more or less parallel to a relative clause “a pitcher who formerly belonged to the Giants”, where the predication “formerly belonged to the Gaints” factually holds for the head noun “pitcher”. Similarly in (18b), it is necessary that the “pitcher” formerly belonged to the Giants team. It is an obvious contradiction to say, for example, that Mr. A is *moto | kyozingun | toosyu* but he never belonged to the Giants. However, it is not contradictory to say, using the lexical compound in (18a), that Mr. B, though *sokkyuu-toosyu* ‘fast-ball pitcher’, throws slow balls in today’s game.

The preceding cursory observations will be sufficient to indicate that W^+ has internal semantics and phonology comparable to phrases. An even more striking demonstration of the phrase-like character of W^+ is provided by the prefix *doo-* ‘the aforementioned’, which is shown to participate in sentence-level coreference.

For English, it is well established that the applicabilities of Postal’s (1969) Anaphoric Island Constraint differ in two types of anaphora, inbound and outbound. As shown by Ward, Sproat, and McKoon (1991), outbound anaphora of sublexical elements, as in (19), is made possible by pragmatically evoking an appropriate referent corresponding to a noun contained in a complex word, whereas inbound anaphora like (20) is strictly disallowed by general morphosyntactic principles.

(19) outbound anaphora: Patty is a definite Kal Kan_i cat. Every day she waits for it_i.

(Ward, Sproat, and McKoon 1991:451)

(20) inbound anaphora: *Bill was a MaCarthy_iite, and Fred was also a him_iite. (Sproat 1988: 293)

We are concerned only with inbound anaphora here. Sproat (1988) and Ward, Sproat, and McKoon (1991) resort to different morphosyntactic principles to account for the impossibility of inbound anaphora. For Sproat (1988), inbound anaphora is ruled out because of the categorial mismatch between the DP structure of English pronouns and the morphological categories below X^0 ; for Ward, Sproat, and McKoon (1991: 450), the prohibition stems from the fact that pronouns form a closed class and therefore cannot participate freely in the formation of new words. As far as English is concerned, the two views appear equally plausible. If we turn our attention to Japanese, however, the categorial approach seems more to the point.

A conspicuous case of inbound anaphora is presented by the prefix *doo-*. This prefix has two meanings, illustrated by (21a) and (21b), but only the latter meaning concerns us here.

(21) a. A-san to B-san wa doo-nenrei desu.

‘Mr. A and Mr. B are of the same age.’

- b. Nagai-san wa Washington-syuu no kookoo de 3-nen-kan osie,
 Nagai-Mr TOP Washington-state GEN high-school at 3-years teach,
 ima wa doo|syuu no daigaku de manande iru.
 now TOP said-state GEN university at studying-is
 ‘Mr. Nagai taught at a high school in the state of Washington for three years,
 and is now studying at a university in the aforementioned state.’

The *doo*- in (21a) means ‘same, identical’ and the complex word involving it carries the regular lexical accent. On the other hand, the homophonous prefix in (21b), our present concern, has entirely different properties: it is pronounced with a phrasal accent and has the function of referring to a previously mentioned entity. Thus, *doo* | *syuu* ‘the said state’ is bound by *Washington-syuu* ‘state of Washington’ in the preceding clause.

Before delineating the referential properties of the second instance of *doo*-, we should first show that this prefix is truly referential rather than deictic. This point is significant, because if the function of this prefix were deictic, then its apparent referentiality could be reduced to pragmatic inferences just like the English outbound anaphora in (19). We prove the point by comparing *doo*- with two other phrase-like prefixes which also have referential functions: *hon*- ‘this, the present’ and *too*- ‘the said, in question’.

- (22) a. [In a book preface] hon|syo no mokuteki wa . . .
 this | book GEN purpose TOP
 ‘The purpose of this book is’
 b. [a cashier’s notice] too|ten wa syoohi-zei o itadaite-orimasen.
 this | store TOP sales-tax ACC collect-not
 ‘We (i.e. this store) do not collect sales tax.’

These two prefixes are purely deictic, as is shown by the fact that sentences like (22a, b) can be used merely by pointing to the entities in question, without reference to any linguistic antecedent. If we follow Sag & Hankamer’s (1984) distinction between “deep anaphora” and “surface anaphora”, the referential function of these two prefixes is characterized as deep anaphora, which is determined pragmatically or deictically. In stark contrast to these two is the prefix *doo*-, which turns out to participate in surface anaphora. This prefix invariably requires an overt linguistic antecedent and cannot be used to point deictically to an entity placed before the speaker’s eyes or pragmatically to an entity evoked in the brain.

- (23) [Watching the police arresting a congressman on TV news] *Doo|daigisi no sizisya
 wa okotte iru daroo.

‘I’m sure the supporters of this congressman are angry.’

The ill-formedness of (23) might be partly due to the stylistic incompatibility where the literary and written usage of *doo*- conflicts with the colloquial spoken language. However, even if written in literary style, *doo*- is totally inappropriate if there is no linguistic antecedent. Observe a decided contrast between the deictic *hon*- (24a) and the pronominal *doo*- (24b).

- (24) a. [At the conclusion of an article] hon|ronbun no simekukuri to site ...
 'In concluding this present article, ...'
- b. [At the conclusion of an article] *doo|ronbun no simekukuri to site ...
 'In concluding the aforementioned article, ...'

It is thus found that the W^+ prefix *doo-* is endowed with the function of creating a referential expression which presumably corresponds to pronominals. While the English personal pronouns are distinguished by gender, person, and number, the Japanese *doo-* can be prefixed to an unlimited number of Sino-Japanese head nouns to give rise to pronominal expressions of various semantic types. For example, *doo|sya* 'the said company' may refer to any antecedent NP identifiable as a company or business firm, *doo|daigaku* 'the said university' to any NP identifiable as a university, and *doo|koku* 'the said country' to any NP identifiable as a country.

Having seen the pronominal function of the prefix *doo-*, we now provide examples of inbound anaphora involving this prefix inside.

- (25) a. New South Wales-syuu_i o osotta yamakazi wa yooyaku tinka no kizasi o misete-kita. [Doo|syuu_i | syooboosyo-honbu] wa kinkyuuzitai-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 aforementioned state has lifted the emergency measures.'
- b. Daitooryoo wa asu yuukoo-zyooyaku_i ni tyoo'in-suru. [Doo|zyooyaku_i | saisyuuan] ni yoruto ...
 'The President is going to sign the amity treaty. According to the final version of the said treaty, ...'

In these examples, adapted from newspaper articles, the complex expressions in brackets make up W^+ compounds, the first members of which are anaphoric to the preceding antecedents. Such complex compounds involving *doo-* inside are extremely common in journalistic writings, and in the same style, examples of outbound anaphora and of mixtures of outbound and inbound anaphora are observed no less commonly.

- (26) a. Kodai no iseki ga [Nara-zyosi-daigaku_i | koonai] de mitukatta to, doo|daigaku_i no hakkutu-tyoosa-iinkai ga happyoo-sita.
 '"Ancient ruins were discovered on the campus of Nara Women's University," the excavation committee of the said university announced.'
- b. [Egypt_i | kaku-si] ni yoruto, doo|koku_i no gyoosei-saibansyo wa ...
 'Every Egyptian newspaper reports that the administrative courthouse of the said country ...'
- (27) [K.-syoten_i | syatoo] niyoru kokain-mituyu-ziken de, [moto | doo|sya_i | kameraman] ga beikoku ni syuttyoo-si ...
 'Concerning the case of cocaine smuggling by the president of K. Publishing Company,

an ex-cameraman of the said company went on a business trip to America and ...'

The interpretations of the *doo*-pronominals in (26) and (27) are not achieved by pragmatic inferences like Ward, Sproat, and McKoon's English example (19) but directly by the linguistic antecedents that precede them. This means that in (26a), for example, the antecedent (Nara Women's University) contained in a complex W^+ is syntactically active, and consequently that the whole W^+ does not constitute an anaphoric island. The waiver of the Anaphoric Island Constraint here is due solely to the referential function inherent in the prefix *doo*-, which itself belongs to the W^+ category. Thus, if ordinary DP pronouns are employed, the islandhood of W^+ expressions is restored with respect to both inbound and outbound anaphora. Witness the following examples where the *doo*-words in (26a) and (25b) are replaced by ordinary pronouns.

- (28) a. Kodai no iseki ga [Nara-zyosi-daigaku | koonai] de mitukatta to, sono daigaku no / sono hakkutu-tyoosa-iinkai ga happyoo-sita.

'"Ancient ruins were discovered on the campus of Nara Women's University," the excavation committee of the university announced.'

- b. Daitooryoo wa asu yuukoo-zyooyaku ni tyoo'in-suru. *[Sore / Sono zyooyaku | saisyuuan] ni yoruto . . .

'The President is going to sign the amity treaty. According to its final version . . .'

In the case of outbound anaphora in (28a), the referent of *sono daigaku* 'that university' is vague, and the possibility of designating Nara Women's University comes about only inferentially in the same way as Ward et al.'s English examples. As for inbound anaphora, *sore* 'it' and *sono-zyooyaku* 'that treaty' are totally impossible inside the W^+ expression in (28b).

The question we now ask is: To what extent is the internal structure of W^+ visible to the syntax? The subsequent observations will reveal that the internal structure of W^+ is available to the construal of the *doo*-pronominals in a manner parallel to the internal structure of noun phrases.

We begin by noting that *doo*-derived words are not "anaphors" in the sense of Condition A of the Binding Theory. Like the personal pronouns and unlike the reflexive pronouns, *doo*-pronominals must not be bound in a single clause.

- (29) a. *Gaimu-daizin_i wa doo | daizin_i o semeta.

lit. 'The Foreign Minister_i blamed the said minister_i.'

- b. Gaimu-daizini wa {zibun-zisini/*karei} o semeta.

'The Foreign Minister_i blamed {himself_i/*him_i}.'

- (30) a. *Koosei-syoo_i wa doo | syoo_i o dame ni sita.

lit. 'The Ministry of Welfare_i spoiled the said ministry_i.'

- b. Koosei-syoo_i wa {zibun-zisin_i/*sore_i} o dame ni sita.

'The Ministry of Welfare_i spoiled {itself_i/*it_i}.'

Rather, *doo*-words behave like pronominals in complying with Condition B of the Binding Theory (Pronouns must be free within a governing category). In point of fact, *doo*-words are capable of appearing in ungoverned positions in a sentence.

- (31) a. Gaimu-daizin_i wa [_{NP} doo | daizin_j no hisyo] o semeta.
 Foreign-Minister_i TOP [said-minister_i GEN secretary] ACC blamed
 lit. 'The Foreign Minister_i blamed the said minister_i's secretary.'
- b. Koosei-syoo_i wa [_{NP} doo | syoo_j no syokuin] ni tuutatu o dasita.
 Welfare-Ministry_i TOP [said-ministry_i GEN staff] DAT notice ACC sent
 lit. 'The Ministry of Welfare_i sent notice to the staff of the said ministry_i.'

The sentences in (31) are well-formed because the *doo*-words are contained in NPs and are free there. Interestingly enough, parallel structural relations obtain when *doo*-pronominals are embedded within a larger W^+ (without the mediation of the genitive marker).

- (32) a. Gaimu-daizin_i wa [_N moto | [doo | daizin_j | hisyo]] o semeta.
 Foreign-Minister_i TOP [former [said-minister_i | secretary] ACC blamed
 lit. 'The Foreign Minister_i blamed the former secretary of the said minister_i.'
- b. Koosei-syoo_i wa [_N doo | syoo_j | syokuin] ni tuutatu o dasita.
 Welfare-Ministry_i TOP [said-ministry_i | staff] DAT notice ACC sent
 lit. 'The Ministry of Welfare_i sent notice to the staff of the said ministry_i.'

The two sets of examples in (31) and (32) demonstrate the parallelism between NP structure and W^+ structure. Regardless of the presence or absence of the genitive particle, *doo*-pronominals are accepted if they are embedded within a larger branching structure and are not directly governed by the antecedent. This is a strong indication that the morphological branching inside W^+ is visible to the syntax, at least for the purpose of interpreting the reference of *doo*-words. It should be stressed, however, that we are not claiming that the syntax can freely have access to the internal structure of W^+ . Were this the case, our arguments in section 1 for the lexical integrity of W^+ would lose ground. Rather, our claim is that the islandhood is removed by virtue of the inherent property of the prefix *doo*-.

3. The morphological structure of W^+

Sections 1 and 2 brought to light two conflicting sets of properties observed with W^+ . On the one hand, W^+ expressions are shown to possess lexical integrity; on the other hand, they are found to share with noun phrases certain phrasal properties related to accentuation and the construal of *doo*-. To resolve this conflict, I propose a new morphological category "Word⁺". Before doing so, however, it will be in order to touch upon two other plausible analyses that have been proposed in the literature, one semantic and the other prosodic. Kubozono (1995) analyzes compounds of the (13)-type in terms of the semantic relationships between the two members, including (33a-e).

- (33) a. organization + official position

- kaisya | syatyoo 'company president'
- b. personal name
miki | takesi 'Takesi (first name) Miki (family name) '
- c. name + status
Katoo | kyoozyu 'Professor Kato'
- d. region + part
Kyuusyuu | nanbu 'southern part of Kyushu'
- e. argument relation
naikaku | kaizoo 'reshuffle of the Cabinet'

It is immediately apparent that such a semantic characterization is far from adequate. For one thing, it is next to impossible to enumerate the exact semantic relations that exhaustively comprise the phrase-like compounds, as there are miscellaneous adverbial relations.

- (34) a. time: 2000-nendo | nyuugaku-siken
'entrance examination for the year 2000'
- b. locative: zenbei | gorufu-toonamento
'All U.S. golf tournament'

For another, there are a myriad of examples which have one and the same semantic relation and yet bear different accentual patterns. The examples in (33d) and (33e) should be compared with those in (35), which have the same semantic relations but display the regular lexical accent.

- (35) a. region+part: niHON-ZEnkoku 'all parts of Japan'
- b. argument relation: iSIKI-TYOosa 'attitude survey'

We thus conclude that it is difficult to delimit the range of phrase-like compounds in terms of their internal semantic relations.

Another plausible analysis is suggested by Poser (1990: 286). Drawing on Inkelas's (1989) theory of dual prosodic and morphosyntactic subcategorization, Poser assumes that the phrase-like prefixes morphologically subcategorize a stem but prosodically subcategorize a minor phrase. The notion of subcategorization, though tenable for affixes, does not hold for compound words like those we brought out above. In order to account for those compounds, it will be necessary to set up a prosodic rule that combine two minor phrases into a compound. This, however, begs the fundamental question of exactly what causes the difference between regular compounds and minor-phrase compounds.

Our approach is neither semantic nor prosodic, but morphological. It is proposed that the accent patterns are direct reflections of the morphological architecture of compound and affixed words. To prove this point, we leave out semantic and prosodic considerations, concentrating our attention on the morphological composition.

It is known that Japanese words are grouped into several classes, depending on their internal makeup (Kageyama 1982). The basic units of word formation are single morphemes,

which may be either native or Sino-Japanese. In the case of Sino-Japanese morphemes, the smallest building block is represented by single Chinese characters like hoo 'visit' and bei 'America', which in turn are combined to form two-character words like hoo-bei 'a visit to America', which may be further expanded to bigger words like hoo-bei yotei 'a schedule of a visit to America'. The orderly processes of word formation observed in such examples will be easily formalized in terms of the level-ordered morphology. For example, the process of compounding single Sino-Japanese morphemes is located at an earlier level than that of compounding two-character words.

(36) level 1: [hoo]+[bei]

visit America

level 2: [hoo-bei]+ [yo-tei]

visit-America schedule

Separation of these two levels accounts for the fact that the level 1 compounding is strictly limited to single morphemes and cannot apply to words written with two or more Chinese characters. Compare hoo-bei with the ungrammatical makeups in (37).

(37) a. *hoo-[bei-koku] 'a visit to America'

visit-[America-country]

b. *hoo-[Amerika-gasyuu-koku] 'a visit to the United States of America'

visit-[America-united.states-country]

The preceding is a simplified view of how Japanese word formation is processed: word formation proceeds from smaller units (which phonologically and morphologically have a higher degree of internal cohesion) to larger units (which have a morphologically and phonologically looser connection inside). Technically, this compositional arrangement will be formalized in a variety of ways such as the strict level ordering as suggested above, the use of different types of boundary markers (Kageyama 1982), or X-bar morphology using certain diacritics.

Elaborating on Selkirk's (1982) idea of distinguishing Word and Stem, Kageyama (1993) explores a system of Japanese word formation which sets up three types of building blocks: Root, Stem, and Word. To illustrate the system, I use only Sino-Japanese words here.

(38) a. Root: single morphemes like hoo- 'visit' and bei 'America'

b. Stem: two-morpheme words like hoo-bei 'a visit to America'

c. Word: words that can be used independently, typically with more than two morphemes

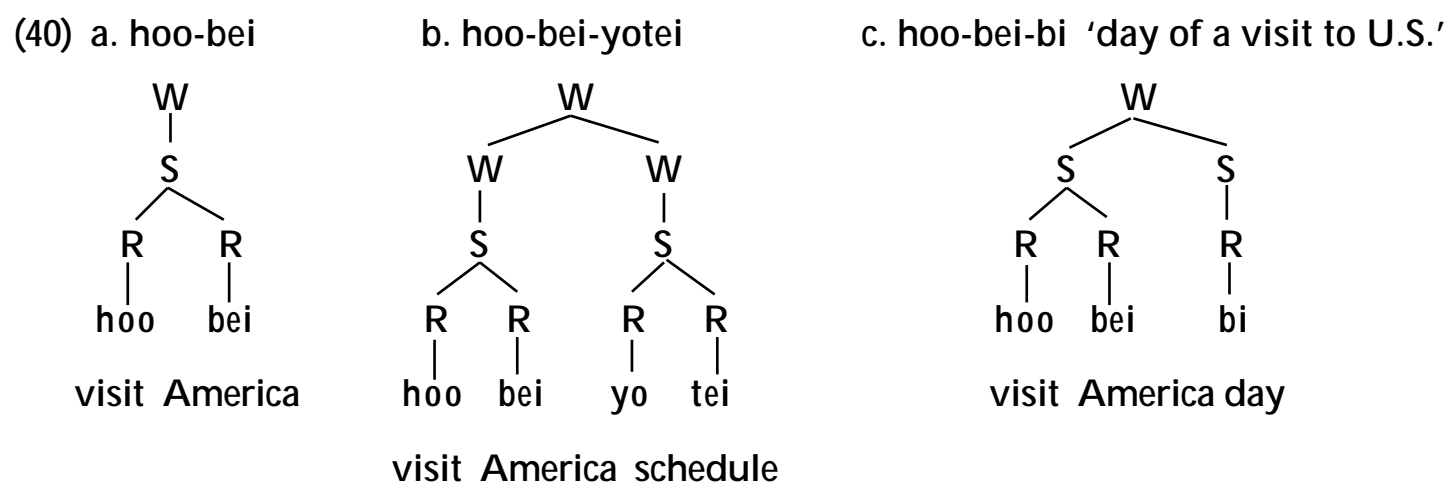
These building blocks are combined with each other by rules like (39) to make larger and larger units.

(39) a. Root (+Root) --> Stem

b. Stem (+Stem) ---> Word

c. Word (+Word) ---> Word

According to these rules, the internal composition of hoo-bei ‘visit to America’ and hoo-bei-yotei ‘schedule for a visit to America’ will be represented as in (40a) and (40b). (Word, Stem, and Root are respectively abbreviated as W, S, and R.)



We need not concern ourselves with the technical details of these representations. What is important is that the difference in ‘size’ among W, S, and R is not just phonological or semantic but has a morphological reality. This is substantiated by the possibility of inserting certain coordinate conjunctions inside the complex words. As observed by Kageyama (1993: 17), the conjunctions *naisi* ‘or’ and *oyobi* ‘and’ have a peculiar capacity to show up inside W-level complex words, but not inside R- or S-level words.

- (41) Word-level complex words with conjunctions inside
 [kokuritu naisi kooritu]-daigaku ‘national or prefectural/municipal universities’,
 [geruman-kei oyobi romansu-kei]-gengo ‘Germanic and Romance languages’,
 gaikokuzin-[kyoosi naisi koosi] ‘foreign teacher or lecturer’, [hoo-bei oyobi hoo-tyuu]
 -yotei ‘schedules of a visit to U.S. and a visit to China’
- (42) Stem-level complex words with conjunctions inside
 *[tyuu-goku naisi kan-koku]-zin ‘Chinese or Korean people’, *[nihon oyobi taiwan]
 -zaru ‘Japanese and Taiwan monkeys’, *hu-[yoozyoo naisi sessei] ‘carelessness or
 negligence about one’s own health’
- (43) Root-level complex words with conjunctions inside
 *hoo-[bei naisi tyuu] ‘a visit to U.S. or China’, *[koo oyobi ryoku]-tya ‘black tea and
 green tea’

Faced with examples like (41), one might wonder whether the coordinated elements as a whole are embedded inside the complex words or they are derived by conjunction reduction from a fuller structure as in (44).

- (44) [kokuritu-daigaku] naisi [kooritu-daigaku]
 national university or prefectural/municipal university
 ---> [kokuritu- ø] naisi [kooritu-daigaku]

Although the reduction analysis cannot be dismissed off hand, it has some unfavorable consequences. First, the suggested reduction rule deleting part of a word results in a blatant violation of the lexical integrity principle. Second, some additional conditions will have to be

introduced to account for the fact that R- and S-level words do not allow the deletion. Third, the deletion analysis claims that the resultant structure is something like *kokuritu-|naisi kooritu-daigaku* where the morphophonological break falls right after the first element *kokuritu-* ‘national’, but the native intuition dictates that the coordinated part [*kokuritu- naisi kooritu-*] ‘national or prefectural/municipal’ is a cohesive unit.

Besides *naisi* and *oyobi*, which are of native origin, there is a Sino-Japanese morpheme which typically concatenates two or more words: *ken* ‘and, cum’, as in *soori-daizin ken gaimu-daizin* ‘Prime Minister and (concurrently) Minister of Foreign Affairs’ and *sinsitu ken ima* ‘a bed-cum-sitting room’. In the light of this conjunctive morpheme, it is not unreasonable to treat *naisi* ‘or’ and *oyobi* ‘and’ as sublexical coordinators that directly conjoin words.

As expected, these coordinators are able to appear inside the W^+ expressions as well. Below are some examples from Kageyama (1993).

- (45) *zen | [soori-daizin oyobi gaimu-daizin]* ‘ex-Prime Minister and ex-Foreign Minister’,
kaku | [kokuritu-daigaku oyobi kooritu-daigaku] ‘each national university and each prefectural/municipal university’, *gaikokuzin | [kyoosi naisi koosi]* ‘foreign teacher or foreign lecturer’

Insofar as the examples in (41) and (45) are concerned, the W^+ expressions are no different from ordinary compounds with lexical accent.

However, a notable difference is detected if another conjunction, *to* ‘and’, is taken into account. The canonical function of this conjunction is to connect NPs rather than mere words. Substitution of *to* for *naisi* and *oyobi* in W-level compounds (41) yields outrageous results like (46).

- (46) **[kokuritu to kooritu]-daigaku* ‘national and prefectural/municipal universities’,
 **[geruman-kei to romansu-kei]-gengo* ‘Germanic and Romance languages’,
 **gaikokuzin-[kyoosi to koosi]* ‘foreign teacher and foreign lecturer’

The last example in (46), *gaikokuzin kyoosi to koosi*, is acceptable only on the reading ‘a foreign teacher and a lecturer’, where the second conjunct (‘lecturer’) is not included in the scope of the modifier (‘foreign’).

Now the ungrammaticality of (46) should be contrasted with the example in (47) which I found in a newspaper headline.

- (47) *ko | Hasegawa-Kazuo-si to Uemura-Naomi-si, kokumin-eeyosyoo ni kimaru.*

‘The late Mr. Kazuo Hasegawa and Mr. Naomi Uemura are to be given the National Honor Award.’ (The Asahi, April 1984)

Poser (1990: 283) briefly mentions this phenomenon, judging a similar example (48) to be unacceptable in the intended meaning where the prefix takes the whole coordination in scope.

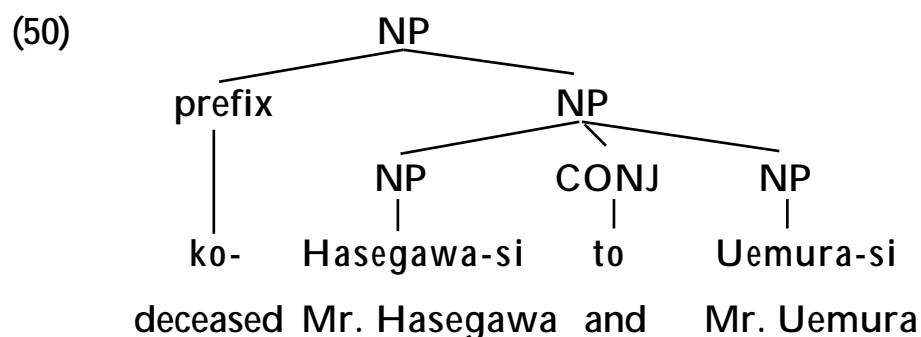
- (48) *moto syusyoo to daitooryoo* (Poser 1990: 283)
 former Premier and President

In view of the attested example (47), however, I am inclined to assume that the coordination

by t_0 is actually well-formed inside W^+ . The analogous examples below seem fully acceptable.

- (49) zen | [bun-soo to hoo-soo] 'ex-Education Minister and Justice Minister', gen | [kaityoo to huku-kaityoo] 'current president and vice-president', dai | [3-syoo to 4-syoo] 'Chapters 3 and 4', boo | [kyoozyu to zyokyoozyu] 'a certain professor and associate professor', syuusyoku | [1-nenme to 2-nenme] 'first and second years of employment', kookyookyoku | [dai-5ban to dai-6ban] 'Symphonies No. 5 and No. 6',

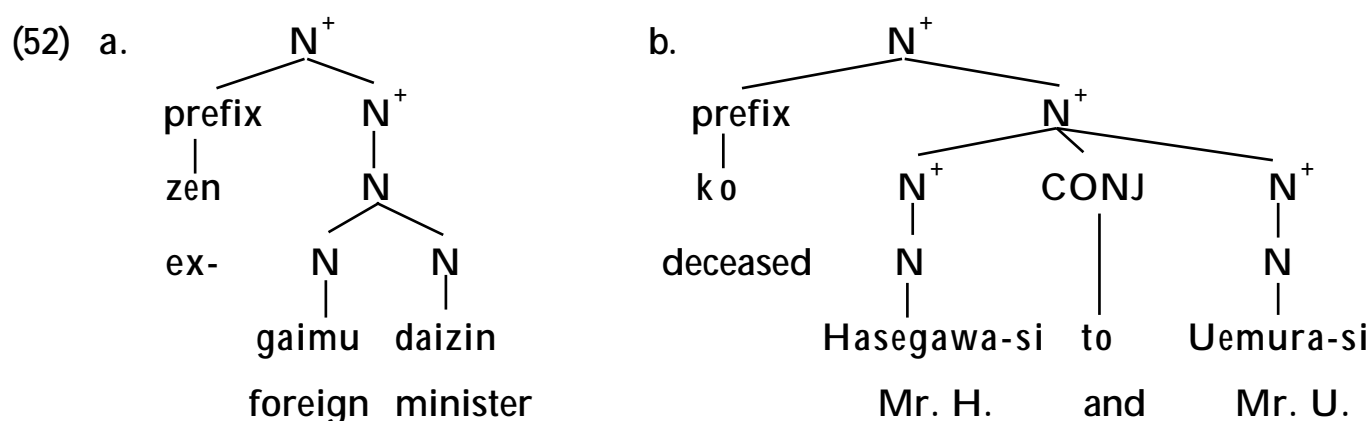
Since to 'and' normally conjoins NPs, it might well be suspected that the underlined expressions in (47) has an NP structure like the following.



Once the phrase-like prefixes are allowed to be attached to NPs, however, it is extremely difficult to prevent examples like (51) involving full NPs with determiners, adjectives, genitive phrases, and relative clauses.

- (51) *ko | [sono haiyuu] ‘that deceased actor’, *zen | [ookina kaisya no syatyoo] ‘ex-president of a big company’, *gaikokuzin | [yoku hataraku kyoosi] ‘foreign teacher who works hard’

The ungrammaticality of (51) reconfirms our assumption made in section 1 that the complex expressions at hand are not phrases but make up words. Coordination appears to be the only exception to the No Phrase Constraint, but the exceptionality will disappear if we assume that the coordinated elements constitute some kind of morphological object, namely a W^+ structure like the following.



Following Lieber (1980), we assume that each affix specifies in its lexical entry what kind of morphological category it selects as the base. For example, the prefix *hu-* 'in-, un-', which can be attached to a Root (e.g. *hu-an* 'uneasiness') and a Stem or Word (e.g. *hu-giri* 'ingratitude'), has a lexical entry like (53a), whereas *ko-* 'deceased', always accompanying a minor phrase boundary, is specified as (53b).

- (53) a. hu- : [___ [Root/Stem/Word]]
b. ko- : [___ [Word⁺]]

Given morphological structures like (52), the appropriate accent patterns can be obtained by assuming that the lexical accent rule assigning a sequence of high pitches applies inside the Word category and does not extend to W^+ . This ensures that a W^+ carries two or more layers of lexical pitch contours, with a minor phrase boundary in between.

Although we have established that W^+ obeys the No Phrase Constraint, a caveat will be necessary on the idiosyncratic behavior of the prefix *moto-* ‘former, ex-’. As previously mentioned, this prefix is the only native morpheme among the W^+ prefixes. It is also peculiar in that it can be used as an independent word, as in (54).

- (54) Kare wa moto no monbu-daizin desu.
 he TOP one-time GEN education-minister is
 ‘He was formerly a Minister of Education.’

Somehow related to these peculiarities is the fact that *moto-* can take what appears to be a genitive NP. Observe the following contrast between the two semantically similar prefixes, *zen-* and *moto-*.

- (55) a. zen | [keizai-kikakutyoo | tyookan]
 ex- [Economic Planning Agency | director general]
 a'. *zen | [keizai-kikakutyoo no tyookan]
 ex- [Economic Planning Agency GEN director general]
 b. moto | [keizai-kikakutyoo | tyookan]
 ex- [Economic Planning Agency | director general]
 b'. moto | [keizai-kikakutyoo no tyookan]
 ex- [Economic Planning Agency GEN director general]

Only *moto-* allows a genitive phrase after it. More examples follow in (56), where the genitive particle *no* seems relatively acceptable, with the same meaning as the compound versions without the genitive.

- (56) a. moto | [massaazi-ten no keieisya]
 former [massage-parlor GEN manager]
 b. moto | [gaikoku-zin no kyoryuuti]
 former [foreign-people GEN settlement]
 c. moto | [kyozingun no toosyu]
 former [the-Giants GEN pitcher]

It is not the case, however, that *moto-* can attach freely to any type of NPs. Descriptive modifiers like adjectives and relative clauses are totally ruled out.

- (57) a. *moto | [yuumei-na toosyu]
 former [famous pitcher] ‘formerly a famous pitcher’
 b. *moto | [[ima Oosaka ni sundeiru] Koobe-simin]
 former [[now Osaka IN living] Kobe-citizen]
 ‘formerly a citizen of Kobe, who now lives in Osaka’

Moreover, there seems to be a semantic restriction on the genitive phrases that are allowed after *moto-*. The well-formed examples in (56) above present a sharp contrast to the ill-formed ones below.

- (58) a. **moto* | [*kamemoti* no *keieisya*]
 former [rich GEN manager]
 ‘formerly a rich manager’
 b. **moto* | [*gaikoku kara kita hitobito* no *kyoryuuti*]
 former [foreign-country FROM came people GEN settlement]
 ‘formerly a settlement of the people coming from foreign countries’

Comparison of (56) with (58) demonstrates that the eligible genitive phrases are limited to those which denote conventionalized concepts or names more or less comparable to lexicalized compound nouns (without the genitive). Just as the compound *gaikoku-zin kyoryuuti* ‘foreign settlement’ is a name designating a particular historical site, so is the genitive phrase *gaikoku-zin no kyoryuuti* when it is used after *moto-*. This semantic characterization strongly suggests that the whole complex with *moto-* is a word rather than a phrase, because only words have a naming function.

Note that the genitive phrases in (56) are clearly distinct from lexicalized expressions like (59)—the Japanese equivalents of what Taylor (1996) calls “possessive compounds” (e.g. a girls’ school and a woman’s magazine).

- (59) *haha* no *hi* ‘Mother’s Day’, *nomi* no *iti* ‘flea market’, *ama* no *zyaku* ‘devil’s advocate’,
ama no *kawa* ‘the Milky Way’, *mago* no *te* ‘back scratcher’

Despite the superficial occurrence of the genitive particle, those expressions are completely lexicalized and enjoy the freedom to participate in lexical word formation, yielding compounds pronounced with a lexical accent, like *haha no hi purezento* ‘Mother’s Day present’ and *nomi no iti biyori* ‘fine weather suitable for flea markets’. Such lexical compounding is not possible with the genitive phrases in (56). This leads us to conjecture that the exceptional occurrence of the genitive in (56) might be due to the contamination of the two different usages of the prefix *moto-* as a W^+ prefix and as the phrasal determiner taking an NP structure (54). Put another way, the *no*-phrases in (56) are attached directly to *moto-* (by mistaken analogy) and are consequently subjected to the general principles of word formation like the naming function.

This conclusion implies that there is no necessity to postulate movement either on the part of the prefix *moto-* or on the part of its base at any level of derivation, like the LF movement of affixes proposed by Pesetsky (1985) or the syntactic incorporation of N to D suggested by Miller (1993:104). The absence of movement will be ascertained by the scope relation of *moto-*. It is a general rule that the prefix takes the base (i.e. the material it attaches to) in its scope. In (60a), *moto-* takes scope over *syusyoo* ‘Prime Minister’, whereby the proper name (Thatcher) that appears to the left of *moto-* is outside the scope. This is reasonable, since ‘*ex-*’ designates

only the official position.

(60) a. Thatcher | moto | syusyoo 'ex-Prime Minister Thatcher'

b. *moto | Thatcher | syusyoo 'formerly Prime Minister Thatcher'

The scope relation makes (60b) meaningless. This example could make sense only in a hypothetical situation where Prime Minister Thatcher changed her family name after leaving her office.

To sum up, we have observed that probably due to its native status, moto- has the exceptional property of expanding its base from Word⁺ to a genitive N'. The N', however, must be restricted to conventionalized nameworthy concepts and cannot include descriptive modifiers, because the whole moto- complex counts as a 'word'.

4. Possible candidates for W⁺ in English

The discussion so far has been confined to Japanese. In this section, we will briefly touch on analogous examples in English to suggest the universal applicability of our proposed W⁺ category.

It has been frequently observed in English that most compound words have primary stress on the first members, while some carry primary stress on the rear (Fudge 1984, Ladd 1984). Obviously, the former compounds correspond to the Japanese compounds with a lexical accent, and the latter to the Japanese W⁺ compounds with a phrasal accent (Kubozono 1995).

(61) world fámous, crystal cléar, waist hígh, student rébellion, government fúnding,
police investigátion, Arnold Pálmer, Mr. Jónes, President Clínton

The phrase-like stress pattern is also observed with such prefixes as non-, ex-, and anti- which have their own stress (Bates 1988).

(62) pseúdo-scíentific, nón-Indo-Européan, ánti-[government intervénition],
éx-[electrical enginéer]

The similarity with the Japanese W⁺ becomes even more striking when we notice that those complex expressions allow internal coordination.

(63) socio-linguistics and economics, anti-abortion and segregation, ex-housewife and homemaker

Expressions like these, which have hitherto been analyzed simply as compound words on a par with regular word-level compounds, are now qualified as candidates for the category W⁺.

The Word⁺ analysis can provide a natural explanation for the occurrence of inflectional elements inside. Consider the position of the plural -s in two types of compound words.

(64) a. púshups (*pushes-up) (Williams 1997), chéckups (*checks-up), breakthroughs
(*breaks-through)
b. lookers-ón, passers-by, hangers-on, diners-out, whippers-in, runners-up,
washers-up, brothers-in-law

The compounds in (64a), which undergo the regular Compound Stress Rule, have the plural -s

on the right edge, whereas those in (64b), carrying a phrase-like stress pattern, allow the plural inflection inside them. Interestingly enough, the reversal of order brings back the regular front stress and sets the plural marker on the right edge.

(65) ónlookers

If we assume that the plural inflection creates the W^+ category, the internal plural marker in the compounds in (64b) can be explained by the following derivation.

(66) pass + -er --> [passer]_N --> [[passer]_N -s]_{N+} --> [[[passer]_N -s]_{N+} [by]_{N+}]_{N+}

A similar phenomenon is observed with complex words involving the suffix -ful. As pointed out by Allen (1978), this suffix has two manifestations in (67a) and (67b).

(67) a. careful, powerful, delightful, thoughtful

b. mouthful, spoonful, cupful, bagful

The adjective-forming suffix -ful in (67a) is stressless with its vowel reduced to a schwa, whereas the -ful in (67b), denoting the amount held by a container, carries a stress on its own and its vowel is fully pronounced as [-ful]. Allen (1978) accounts for the distinction of the two types of -ful by assigning the adjectival -ful to Level 2 and the fully pronounced -ful to Level 3 in her extended level-ordering hypothesis. Our alternative analysis is to treat Allen's Level 3 as constituting a distinct category of W^+ . This analysis is supported by the fact that the expressions in (67b) can have the plural inflection inside them, although a recent tendency is to put the plural on the right edge.

(68) spoonsful, bagsful (cf. spoonfuls, bagfuls)

Another candidate for W^+ in English is found in certain types of compound adjectives which realize comparative and superlative inflections on their first members.

(69) a. widespread, more widespread, *wider-spread/*widest-spread

easy-going, more easy-going, *easier-going

b. well-known, better-known/best-known, *more well-known/*most well-known

good-looking, better-looking/best-looking, *more/*most good-looking

well off, better off/best off, *more well off/*most well off

long-running, the longest-running musical comedy (Namiki 1985: 144)

kind-hearted, kinder-hearted (Namiki 1985: 144)

The examples in (64b), (67b), and (69b) have their internal structures accessible to the syntax, so that the plural or comparative/superlative inflections, which are generally held to take place in syntactic structure, sneak into the branching W^+ structure.

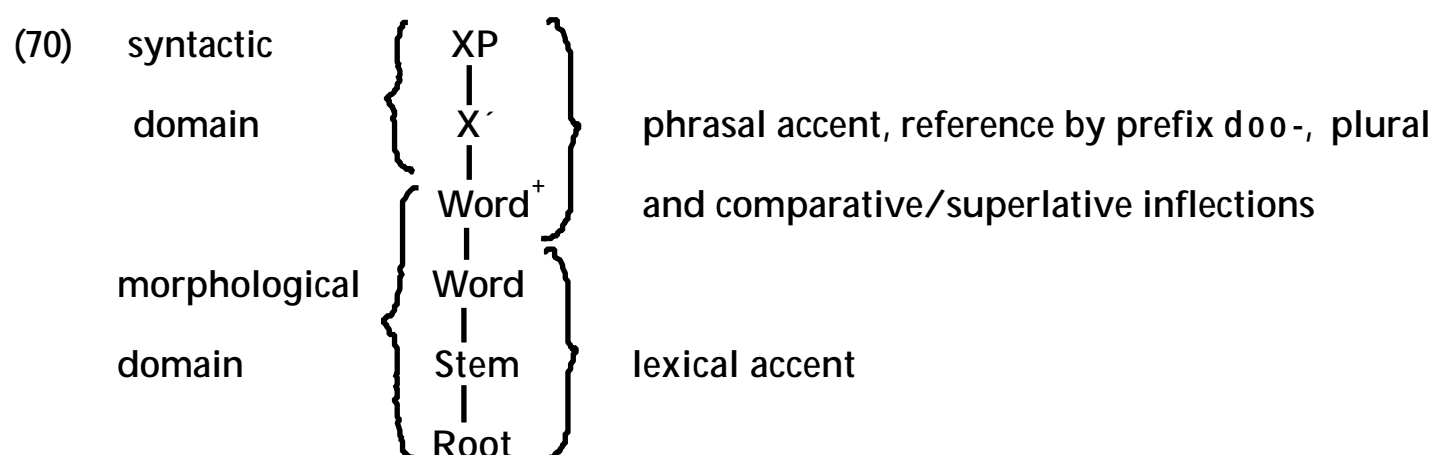
5. Conclusion

Recently, Bisetto & Scalise (1999) have taken up Italian complex expressions which look like compound words but are analyzed as syntactic phrases because of the insertability of lexical material. From this they conclude that morphology and syntax can be distinguished as different components. The phrase-like expressions discussed in this paper pose a similar

problem of the morphology-syntax boundary but lead us to a different conclusion from Bisetto & Scalise's.

We have argued that different kinds of building blocks participate in forming various 'sizes' of words in Japanese. In particular, the new morphological category of W^+ is proposed which is larger than Word but nonetheless belongs to the realm of morphology. It might appear that the formation of W^+ can be accommodated in the framework of level-ordered morphology by allocating it to the last level of derivation in the lexical component, just as Kiparsky (1982) proposed that the regular inflections take place at the last stratum in his lexical phonology theory. However, if the lexical and syntactic components are sharply separated from each other, with the bracket erasure convention wiping out all morphological structures before the syntax, the behavior of the referential prefix *doo-* observed in section 2 remains unaccounted for.

Instead of level ordering, we postulate different types of morphological categories to capture the real nature of the phenomena. The repertoire of morphological and syntactic categories will now be hierarchically represented as follows.



Here the category $Word^+$ is situated at the intersection of morphological and syntactic structures. As shown by various tests in section 1, $Word^+$ is qualified as a morphological object in terms of lexical integrity. The syntax-morphology distinction in the sense of the "component problem" posed at the outset of the paper, however, does not match the "atomicity problem" perfectly, because the internal composition of W^+ can be made visible for particular syntactic and phonological purposes.

On the phonological side, the lexical accentuation in Japanese, applying inside Word, allows W^+ to have two or more lexical accents, thereby giving W^+ a phrase-like appearance. On the syntactic side, the referential function of the prefix *doo-* serves to cancel the anaphoric islandhood, rendering the internal structure of W^+ available to the interpretation of anaphoric relations at the sentence level. In the same way, the plural and comparative/superlative inflections in English, which are normally attached to the right edge of Word, can be embodied inside W^+ , as in *lookers-on* and *better-known*.

We have thus seen that the component problem (the lefthand side of the diagram in (70)) and the atomicity problem (the righthand side in (70)) are not strictly co-extensive but diverge

at the border indicated by the category W^+ . In a way, W^+ is a special category which might have originated from phrase structure but is now grammaticized as a morphological object. The reason we regard W^+ as special is that it manifests itself only when the designated prefixation or compounding is executed, and in unmarked cases the maximal projection in word structure should be deemed Word or X^0 as has been customarily assumed. For, if the Word category were automatically projected to $Word^+$, then all words would have to be pronounced with a phrasal accent. This makes a decided contrast to the category projections in syntax, where all phrases are assumed to automatically project to the maximal projection XP. We thus claim that the way of projection in word structure is different from the one in syntactic structure, contrary to Lieber (1992) and Ackema (1999), who maintain that identical principles apply to both structures.

Though strange at first glance, the existence of W^+ will not be so unnatural if we recall that the reverse case is also attested, namely syntactic incorporation. Since the unmarked locus of compounding is the lexicon, syntactic incorporation can be viewed as a process which is lexically originated but is established as a grammatical operation in syntax. Viewed from this perspective, the $Word^+$ category and syntactic incorporation can be looked upon as epiphenomena deriving from a more fundamental principle of modularity in grammar: the modules that comprise the grammar are basically independent as the lexicalist hypothesis assumes, but they are given leeway to have access to each other under certain conditions.

References

- Ackema, P. (1999). *Issues in morphosyntax*. Amsterdam: John Benjamins.
- Allen, M. (1978). *Morphological investigations*. Ph.D. dissertation, University of Connecticut.
- Baker, M. (1988). *Incorporation*. Chicago: University of Chicago Press.
- Bates, D. (1988). *Prominence relations and structure in English compound morphology*. Ph.D. dissertation, University of Washington.
- Bisetto, A. & S. Scalise (1999). Compounding: morphology and/or syntax? In L. Mereu (ed.), *Boundaries of morphology and syntax*. Amsterdam: John Benjamins. 31--48.
- Borer, H. (1998). Morphology and syntax. In A. Spencer & A.M. Zwicky (eds.), *The handbook of morphology*. Oxford: Blackwell. 151--90.
- Bresnan, J. & S. Mchombo (1995). The lexical integrity principle. *Natural Language & Linguistic Theory* 13, 181--254.
- Di Sciullo, A.-M. & E. Williams (1987). *On the definition of word*. Cambridge, Mass.: MIT Press.
- Fudge, E. (1984). *English word-stress*. London: George Allen & Unwin.
- Halle, M. & A. Marantz (1993). Distributed morphology and the pieces of inflection. In K. Hale & S.J. Keyser (eds.), *The view from building 20*. Cambridge, Mass.: MIT Press. 111--76.

- Inkelas, S. (1989). Prosodic constituency in the lexicon. Ph.D. dissertation, Stanford University.
- Kageyama, T. (1982). Word formation in Japanese. *Lingua* 57, 215--58.
- Kageyama, T. (1989). The place of morphology in the grammar. In G. Booij & J. van Marle (eds.) *Yearbook of Morphology* 2. Dordrecht: Foris. 73--94.
- Kageyama, T. (1993). *Bunpoo to gokeisei* [Grammar and word formation]. Tokyo: Hituzi Press.
- Kageyama, T. (1999). Word formation. In N. Tsujimura (ed.), *The handbook of Japanese linguistics*. Oxford: Blackwell. 297--325.
- Kiparsky, P. (1982). Lexical morphology and phonology. In *The Linguistic Society of Korea* (ed.), *Linguistics in the morning calm*. Seoul: Hanshin Pub. Co. 3--92.
- Kubozono, H. (1995). *Gokeisei to on'in-koozoo* [Word formation and phonological structure]. Tokyo: Kurosio Publishers.
- Ladd, D. R. (1984). English compound stress. In D. Gibbon & H. Richter (eds.), *Intonation, accent and rhythm*. Berlin: Mouton de Gruyter. 253--66.
- Lieber, R. (1980). On the organization of the lexicon. Ph.D. dissertation, MIT.
- Lieber, R. (1992). *Deconstructing morphology*. Chicago: University of Chicago Press.
- Miller, D. G. (1993). *Complex verb formation*. Amsterdam: John Benjamins.
- Namiki, T. (1985). *Gokeisei* [Word formation]. Tokyo: Taishukan.
- Pesetsky, D. (1985). Morphology and logical form. *Linguistic Inquiry* 16, 193--246.
- Poser, W. (1990) Word-internal phrase boundary in Japanese. In S. Inkelas & D. Zec (eds.), *The phonology-syntax connection*. Chicago: University of Chicago Press. 279--88.
- Postal, P. (1969). Anaphoric islands. *CLS* 5, 205--39.
- Roeper, T. & M. Siegel (1978). A lexical transformation for verbal compounds. *Linguistic Inquiry* 9, 199--260.
- Sag, I. & J. Hankamer (1984). Towards a theory of anaphoric processing. *Linguistics and Philosophy* 7, 325--45.
- Selkirk, E. (1982). *The syntax of words*. Cambridge, Mass.: MIT Press.
- Shibatani, M. & T. Kageyama (1988). Word formation in a modular theory of grammar. *Language* 64, 451--84.
- Sproat, R. (1988). On anaphoric islandhood. In M. Hammond & M. Noonan (eds.), *Theoretical morphology*. San Diego: Academic Press. 291--301.
- Taylor, J. (1996). *Possessives in English*. Oxford: Clarendon Press.
- Ward, G., R. Sproat & G. McKoon (1991). A pragmatic analysis of so-called anaphoric islands. *Language* 67, 439--72.
- Williams, E. (1997). Lexical and syntactic complex predicates. In A. Alsina et al. (eds.), *Complex predicates*. Stanford: CSLI Publications. 13--28.