Chapter 6

Phrasal compounds in Japanese

Kunio Nishiyama
Ibaraki University

Although Japanese does not have phrasal compounds analogous to English *an over the fence gossip* or *a who's the boss wink*, it does have phrasal compounds like *kireena mati-dukuri*, literally 'clean city-making,' meaning construction of a clean city. This example illustrates one of several types of phrasal compounds in Japanese. The criteria that classify phrasal compounds in Japanese are (i) whether the head of the compound is a predicate, (ii) if a predicate, whether the head is of Sino-Japanese or of native origin, (iii) if not a predicate, whether the compound involves coordination or cliticization.

One source of phrasal compounding is noun incorporation. When an argument incorporates into a Sino-Japanese verbal noun predicate, we get what Shibatani & Kageyama (1988) refer to as post-syntactic compounds, which have phrasal accent. In contrast, when an argument incorporates into a verbal noun predicate of native origin, we get a phrasal compound with word accent. The phrasal nature is evidenced by modifier stranding, and there are some conditions (e.g., pragmatic factors like cliché) on modifier stranding. There are three other sources of phrasal compounding which do not involve noun incorporation: natural coordination, enclitics, and proclitics. The first two have word accent and the last has phrasal accent. Whether a compound has word accent or phrasal accent is predicted by its structure: right branching compounds have phrasal accent (Kubozono 1995; 2005). Kageyama's (1993; 2001; 2009) notion of Word Plus is reconsidered and reclassified into three distinct classes: right-branching compounds, constructions involving proclitics, and phrases involving genitive deletion.

1 Introduction

The term "phrasal compound" refers to compounds containing a phrase, in apparent violation of Botha (1981) No Phrase Constraint, exemplified by the following English examples:

Kunio Nishiyama. Phrasal compounds in Japanese. In Carola Trips & Jaklin Kornfilt (eds.), *Further investigations into the nature of phrasal compounding*, 149–183. Berlin: Language Science Press. DOI:10.5281/zenodo.885125

(1) a. an over the fence gossip
b. a who's the boss wink (Lieber 1992)

Apparently parallel examples in Japanese are as follows:

(2) a. Tokyo-kara-no nimotu
Tokyo-from-GEN package
'a package from Tokyo'
b. dare-ga bosu-da-teki taido
who-NOM boss-COP-like attitude
'a who's the boss attitude'

In (2a), the genitive marker *no* emerges between PP and the head noun. Thus, the example is not a compound but a phrase like its English translation. In (2b), a morpheme *-teki* 'like' attaches to the sentence 'who's the boss.' The morpheme usually attaches to a word (e.g., *hankoo-teki* 'rebellion-like, rebellious'), but it has recently acquired the ability to attach to a phrase (the example in (2b) has an innovative or substandard flavor). The attachment of *-teki* is a case of encliticization, which is discussed in §4.2. *-teki* is also discussed in §5, but in the present context, it suffices to notice that (2b) as a whole is not a compound but a phrase like 'a "who's the boss"-like attitude,' consisting of a modifier and the head noun. In short, neither of the examples in (2) is a compound. This is evidenced by the fact that (2a) and (2b) have phrasal accent. Accent in Japanese is described in §2.

Although the examples in (2) are not phrasal compounds, Japanese does have phrasal compounds like the following:

(3) kireena mati-dukuri
clean town making
'construction of a clean town' (Kageyama 2009: 518)

Here, *mati-dukuri* is a compound, and *kireena* modifies a part of the compound, resulting in a syntactic bracketing of [*kireena mati*]*-dukuri*. In other words, we have modifier stranding. This is a case of a bracketing paradox, for the bracketing in terms of phonological words is [*kireena*] *mati-dukuri*. It is reminiscent of *criminal lawyer*, with the meaning of a lawyer who practices criminal law (cf. Beard 1991). With this meaning, the syntactic bracketing is [*criminal*] *law*]*yer*. The difference is that (3) does not involve bound derivational morphemes but compounding.

The purpose of this paper is to describe and analyze phrasal compounds in Japanese. Most of the examples discussed in this paper are reproduced from previous studies. However, in those previous studies, such examples of phrasal compounds are not discussed within an explicit perspective of phrasal compounds. This paper integrates several types of compounds within such a perspective. In addition to the type illustrated in (3), Japanese has a number of other types of phrasal compounds. The criteria used for classifying phrasal compounds in Japanese are as follows: (i) Whether the head of the compound is a predicate; (ii) if the head is a predicate, whether it is of Sino-Japanese origin (i.e. whether it is a vocabulary item in Japanese which is of Chinese origin), or whether it is of native origin; (iii) if the head of the compound is not a predicate, whether the compound involves coordination or cliticization.

This paper is organized as follows. §2 briefly introduces accent in Japanese, which is crucial in differentiating between words and phrases. In §3, phrasal compounds formed by noun incorporation are discussed. There are two subtypes: one type involving Sino-Japanese verbal nouns (3.1) and one type involving verbal nouns of native origin (§3.2). §4 discusses phrasal compounds without noun incorporation. There are three subtypes: one involving natural coordination (§4.1), one involving suffixes (enclitics) (§4.2), and one involving prefixes (proclitics) (§4.3). In §5, Kageyama's (1993; 2001; 2009) notion of Word Plus is reconsidered and reclassified into several existing notions.

2 Accent in words and phrases in Japanese

Just like English *green hóuse* versus *gréenhouse*, accent differentiates between words and phrases in Japanese. Key features of accent in Japanese are summarized as follows (see also Kawahara 2015) (H is for high and L for low):

(4) *Accent in Japanese*

a. Accent is defined as falling pitch (HL).
b. A word is either accented or unaccented.
c. Where the accent falls is specified for each accented word.
d. A word can have at most one accent.
e. A word starts as either LH (rising pitch) or HL (falling pitch) (the latter of which instantiates accent on the first mora).

In this paper, the feature in (4d), i.e. that a word can have at most one accent, becomes crucial. The following examples illustrate accent in words:

(5) a. inu ‘dog’ LH (unaccented)
b. nèko ‘cat’ HL (accent on the first mora)
c. huransu ‘France’ LHHH (unaccented)
d. dòitu ‘Germany’ HLL (accent on the first mora)
e. yooròppa ‘Europe’ LHHLL (accent on the third mora, segmented yo.o.ro.p.pa)

When relevant, accent is represented with a grave diacritic on the accented vowel in this paper.

Given that a compound is a word, there should be at most one accent in a compound, according to (4d). The accentuation rules of compounds are complicated (cf. Kubozono 2008; Nishiyama 2010), but typically accent falls on the first mora of the right-hand element, regardless of how each element in the compound is accented independently.[1] This is illustrated in the following examples:

(6) a. dòitu + bùngaku → doitu-bùngaku, *dòitu-bùngaku
‘German literature’
b. LHHH LHH LHHH HLL LHHH LHH
booeki + kaisya → booeki- gàisya *booeki- gaisya
‘a trading company’

(6a) is a case of compounding of *dòitu* ‘Germany’ and *bùngaku* ‘literature’, both of which are accented on the first mora. **dòitu-bùngaku*, which has two accent positions, is ruled out by (4d). The correct form *doitu-bùngaku* bears accent on the first mora of the right-hand element. Thus, the right-hand element seems to retain the position of its accent in the compound. But this is not the case in (6b), where both of the elements *booeki* and *kaisya* are unaccented originally. Here, the resulting compound *booeki-gàisya* is likewise accented on the first mora of the right-hand element, and thus has the pitch contour LHHH-HLL. The alternative **booeki-gaisya* (LHHH-LHH) is ruled out, because there cannot be an instance of rising pitch after falling pitch in a word. On the assumption that a compound is a word that obeys the word accent rules, the pitch contour LHHH-LHH is ruled out, for the H-LH part instantiates rising pitch after falling pitch.

In this paper, when there is at most one accent in a word, I refer to it as *word accent*. In contrast, *phrasal accent* refers to independent accent for each word in a phrase. This typically happens when a phrase includes the genitive *no*:

[1] The rationale behind this accentuation is to mark the root boundary (cf. Kubozono 2008).

(7) dòitu-no bùngaku
Germany-GEN literature
'literature of Germany' (cf. (6a))

Here, the accent of each element, *dòitu* and *bùngaku*, is retained. This is because (7) is a phrase. (7) is to be compared to the compound *doitu-bùngaku* in (6a), where there is only one accent. Given that words are either accented or unaccented, *phrasal accent* refers to not only multiple accent but also to instances of falling pitch followed by rising pitch, which is prohibited in a word.

Another feature of accent in Japanese crucial in this paper is its sensitivity to the internal structure of compounds. Concretely, when three elements are involved, while left-branching compounds obey the compound accentuation rule (having at most one accent), right-branching compounds violate it, resulting in multiple accent. This is illustrated in the following examples (cf. Kubozono 2005: 13):

(8) *Left-branching vs. right-branching*

a. [dòitu + bùngaku] + kyookai → doitu-bungaku-kyòokai
Germany literature association
'Association of German Literature'

b. dòitu + [bùngaku + kyookai] → dòitu : bungaku-kyòokai
Germany literature association
'German Association of Literature'

(8a) is a compound consisting of [*dòitu* + *bùngaku*] and (inherently unaccented) *kyookai*. This means that the compound has the left-branching structure, and the resulting *doitu-bungaku-kyòokai* 'Association of German Literature' has only one accent, namely word accent. In contrast, (8b) is a compound consisting of *dòitu* and [*bùngaku* + *kyookai*] (i.e., right-branching), and the resulting *dòitu : bungaku-kyòokai* 'German Association of Literature' has multiple accent, namely phrasal accent, which is reflected by the colon (:).

The distinction between word accent and phrasal accent is crucial throughout this paper. Basically, we can identify the word/phrasal status of word strings by the accent pattern. Thus, when the string [A B] has word accent, A and B are taken to form a compound, and are cited as "A-B".

In the present context, the behavior of right-branching compounds is exceptional: they have phrasal accent, but are *not* phrases syntactically. That they are not syntactic phrases is shown by the following example:

(9) *doitu to huransu : bungaku-kyookai
Germany and France literature association
'associations of literature in Germany and France'

Here, the left-hand element is a coordination of proper nouns, and thus is a phrase. The ungrammaticality of (9) shows that right-branching compounds are not phrasal compounds, despite having phrasal accent. (We return to coordination in §4.1.) This is a case where a phonological notion and a syntactic notion do not match: phrasal accent is a phonological notion, and does not always reflect the syntactic status of a phrase.[2] In this sense, phrasal accent in itself is not helpful in deciding whether a compound is a phrasal compound or not. Note, however, that whether the accent is word-like or phrasal is crucial in determining whether compounding is involved or not, as mentioned above.

Some notes on notations and terminology in this paper are in order. "A-B" represents compounds with word accent, which are called *real compounds*. In contrast, "A : B" represents compounds with phrasal accent (like 8b), which are called *pseudo compounds*.[3]

To recap, the following premise is crucial in the following sections:

(10) Right-branching compounds have phrasal accent.

Before concluding this section, let us see why right-branching compounds are exceptional. Kubozono (1995: 107) notes that right-branching A+[B+C] is harder to process than left-branching [A+B]+C (see also Hawkins 1990 and Sugioka 2008). To remedy the processing difficulty, right-branching compounds are exceptionally multiple-accented (phrasal-accented), making constituency easy to identify.

3 Noun-incorporated phrasal compounds

This section discusses phrasal compounds formed by noun incorporation. Unlike noun incorporation familiar from polysynthetic languages, noun incorporation

[2] The term "phrasal compounds" is used in Ito & Mester (2007) to refer to compounds with phrasal accent. Their term is based on the *phonological* notion of "phrase" that comes between "intonational group" and "word" in the prosodic hierarchy. Crucially, "phrasal compounds" in Ito & Mester (2007) are *not* phrasal compounds as defined in this paper (and in this volume as well) as XP-X, namely utilizing the *syntactic* notion of "phrase".

[3] Kageyama (1993; 2001; 2009) uses the colon : for what he terms post-syntactic compounds (discussed in §3) and | for what he terms Word Plus (including (8b), discussed in §s 4 and 5). As far as accent is concerned, they all have phrasal accent. Moreover, I argue in §5 that there is no need to postulate Word Plus as a novel concept. Therefore, I use only the colon notation.

in Japanese is limited to verbal noun predicates (or nominalized verbs).[4] Depending on whether the predicate is of Sino-Japanese or of native origin, the resulting phrasal compounds behave differently with respect to accent, and this led previous studies to treat them separately. I claim that this dichotomy is theoretically unmotivated. Phrasal compounds involving Sino-Japanese predicates are discussed in §3.1, and those involving predicates of native origin are discussed in §3.2.

3.1 Noun incorporation resulting in phrasal accent: Sino-Japanese verbal nouns

This section discusses "post-syntactic compounds" in the sense of Shibatani & Kageyama (1988) (henceforth S&K). The analysis in S&K is extended in Kageyama & Shibatani (1989) (K&S) and Kageyama (1993) and is also mentioned in Kageyama (2009). The following summarizes the key features of the compounds analyzed in S&K:

(11) *Features of noun-incorporated pseudo compounds*

a. They have phrasal-accent.
b. The right-hand element is a Sino-Japanese verbal noun predicate.[5]
c. The left-hand element is the complement of the right-hand predicate.
d. The complement is in a case-marked position before incorporation.

Due to the feature in (11a), the examples discussed in this section are called pseudo compounds.

Consider first the following three examples:

(12) a. yooroppa-ryòkoo
Europe-traveling
'Europe-traveling'
(real compound, with word accent)

b. yooròppa-o ryokoo-tyùu
Europe-ACC traveling-while
'while traveling in Europe'
(the temporal suffix *-tyuu* 'while' attached to a VP)

[4] With the exception of several (lexicalized) verbs like *tabi-datu* 'trip-set.out,' where the verb remains non-nominalized, noun incorporation resulting in a verb is quite limited and unproductive, unlike noun incorporation involving verbal nouns as discussed in this section.

[5] Kageyama (1993: 240f) notes that an adjectival noun can also be the right-hand element of a noun-incorporated pseudo compound. We will return to adjectival nouns in note 13.

c. yooròppa : ryokoo-tyùu
Europe traveling-while
'while traveling in Europe'
(pseudo compound, with phrasal accent)

(12a) is a case of a real compound; it has word accent, i.e., only one accent position on the first mora of the right-hand element. (12b) is obtained by attaching a temporal suffix *-tyuu* 'while' to the VP 'travel Europe.' Note that the object is Accusative-marked. (12c) is a case of noun-incorporated pseudo compound. It is a pseudo compound because it is phrasal-accented (i.e. multiple-accented).

As noted by S&K (p. 462), a manner adverb can intervene between the object and the predicate in (12b), but not in (12c):

(13) a. yooròppa-o nonbiri ryokoo-tyùu
Europe-ACC leisurely traveling-while
'while traveling in Europe leisurely'

b. *yooròppa : nonbiri ryokoo-tyùu
Europe leisurely traveling-while
'while traveling in Europe leisurely'

This shows that (12c) is not simply derived from (12b) by case deletion. More specifically, (12c) is not a phrase but a word (compound).

One phenomenon that points to the involvement of noun incorporation is modifier stranding (cf. Baker 1988). Modifier stranding also indicates that a phrase is involved in the compounding. As demonstrated by S&K, the compounds in question allow modifier stranding:[6]

[6] Kageyama (2009: 525) says that noun-incorporated pseudo compounds (post-syntactic compounds in his terminology) do not tolerate modifier stranding, but this refers to a different type of modifier stranding. Kageyama's example is as follows:

(i) a. hidari-asi-o kos-setu
left-leg-ACC bone-break
'to break the bone of the left leg'.

Here, *hidari-asi* is supposed to be modifying *kos*, but this is not literally the case. *Kos* is a Sino-Japanese lexical item for 'bone' and can be used only in Sino-Japanese compounds. When modified by *hidaro-asi* independently, the correct word for 'bone' is a native word *hone*, as:

(ii) a. hidaro-asi-no hone
left-leg-GEN bone
'the bone of the left leg'

(14) a. kono zikken : syuuryoo-go
this experiment finish-after
'After this experiment finishes,' (S&K: 471)

b. [watasi-ga ima yatteiru] zikken: syuuryoo-go
I-NOM now doing experiment finish-after
'After the experiment that I am now doing finishes,' (S&K: 472, adapted)

Note that a modifier (a demonstrative in (14a) and a relative clause in (14b)) of *zikken* 'experiment' is stranded.[7]

As in (13), an adverb can intervene between *zikken* and *syuuryoo* in a clause as in (15a), but not in a compound as shown in (15b):

(15) a. kono zikken-ga yooyaku syuuryoo-go
this experiment-NOM finally finish-after
'After this experiment finally finishes'

b. *kono zikken : yooyaku syuuryoo-go
this experiment finally finish-after
'After this experiment finally finishes,'

Compare (15b) with (14a). This shows that there is no phrasal boundary between *zikken* and *syuuryoo* in (14); they form a compound, as S&K argue.

In addition to an accusative NP (12) and a nominative NP (14), a genitive NP can also incorporate:[8,9]

---

[7] In this paper, I use the term "modifier" loosely as "being a part of the argument DP." Thus, it is immaterial whether the modifier is an adjunct or a specifier in the phrase structure.

[8] A dative NP can also incorporate:

(i) butyoo-e-no syoosin
department.head-DAT-GEN promotion
'promotion to the department head'

(ii) butyoo : syoosin (K&S: 154)

no here is more like a linker, as we saw in (2a).

[9] I leave open the exact theoretical mechanism of noun incorporation. It is generally assumed (cf. Baker 1988) that noun incorporation is restricted to internal arguments. Therefore, one might think that (16)) as well as (14)) involve incorporation of an unaccusative subject, which is underlyingly an object. However, Kageyama (2009: 517f, 2013) shows that an agentive noun can also incorporate:

(16) a. zyukensee-no zooka(-no riyuu)
applicant-GEN increase(-GEN reason)
'(the reason of) increase of applicants'

b. zyukensee : zooka(-no riyuu)
applicant increase(-GEN reason)
'(the reason of) increase of applicants'

(16a) is an ordinary noun phrase involving a genitive-marked argument. (16b) is a case of a pseudo compound formed by incorporation of the (originally genitive-marked) argument.

In fact, S&K (1988) are not explicit about the relevance of noun incorporation in the formation of the compounds in question, and suggest (p. 480, n. 15) that the genitive is deleted in examples like (16b). It is in K& S (1989: 155) and Kageyama (1993: 236) that the noun incorporation analysis is entertained. Concretely, they say that (16a) and (16b) have the common caseless, non-incorporated structure:

(17) a. [zyukensee zooka] $_{VNP}$ b. [zyukensee zooka] $_{VNP}$
↓ case realization ↓ compounding (noun incorporation)
[zyukensee-no zooka] [zyukensee : zooka]

With case realization, we get (17a) (=16a), and with noun incorporation, we get (17b) (=16b).

(i) Spielberg : seesaku-no eega
S. production-GEN movie
'a movie that Spielberg produced'

Although Kageyama argues that the 'internal argument constraint' is still valid, for the agent compounds in question must be used adjectivally as above, this raises the question of whether the Baker-style incorporation is involved in the compounds in question. To complicate the issue, there are counterexamples to the internal argument constraint itself (cf. Mithun 2010 and Lieber 2010, among others). Due to such considerations, one might opt for merger under adjacency (cf. Marantz 1988 and Halle & Marantz 1993) or the First Sister Principle of Roeper & Siegel (1978) as the mechanism of the compounding in question, but I leave further discussion on the issue for future research. Incidentally, Kageyama (2009: 525) notes that the incorporation in question is not a case of Pseudo Noun Incorporation in the sense of Massam (2001), a phrase structure in which an NP directly merges with a V, because the incorporated elements in Japanese are not phrases. Specifically, although a phrasal argument can be in the original structure before incorporation, only the head can participate in compounding, with the modifier stranded, as we saw in (14).

3.2 Noun incorporation resulting in word accent: verbal nouns of native origin

Japanese abounds in compounds with a nominalized verb of native origin as the right hand element and its argument as the left hand element:

(18) gomi-atume 'garbage collecting'
yuki-kaki 'snow plowing'

Unlike the compounds discussed in the previous section, the compounds in (18) have word accent, which will be illustrated in §3.2.1. This section discusses such compounds. §3.2.1 discusses compounds with a phrasal complement as evidence for phrasal compounding, and seeks an account for why they have word accent, in contrast to the phrasal-accented compounds discussed in the previous section. §3.2.2 offers conditions on modifier stranding. §3.2.3 compares noun-incorporated compounds discussed in this paper and the so-called synthetic compounds (in English) like *mountain climbing*.

3.2.1 Compounds with a phrasal complement and compound accentuation

Sugioka (2002: 496) argues that compounds of the type illustrated in (18) are formed by noun incorporation—a proposal which I basically follow here. (But I leave the exact mechanism of noun incorporation open (cf. note 9), and argue in §3.2.3 that the compounds in (18) are structurally ambiguous.) In (18), the left-hand element of the compound is a word, so it is not clear whether a phrase is involved. To make sure that phrasal compounding is involved, a modifier is called for, and indeed, with this type of compounds, sometimes (but not always, see §3.2.2) modifier stranding is possible.

(19) a. [titi-no haka]-mairi (cf. Kageyama 2009: 521)
father-GEN grave-visiting
'visiting father's grave' (lit. [father's grave]-visiting)

b. [asagao-no tane]-maki (Kageyama 1993: 334)
morning.glory-GEN seed-sowing
'sowing seeds of morning glory' (lit. seed-sowing of morning glory)

This shows that the compounds above are formed in the phrasal syntax. Since a predicate participates in this type of compounding, noun incorporation is likely to be involved, as suggested by Sugioka (2002). I return to evidence for this in §3.2.3.

That compounding is really involved in (19) is confirmed by accent. (20a) illustrates accent of a VP in a sentence, while (20b) illustrates accent of the corresponding verbal noun:

(20) a. LHLLL HL(L) HL
asagao-no tane(-o) mak-u
morning.glory-GEN seed(-ACC) sow-PRES
'to sow seeds of morning glory'

b. LHLLL LHLL
asagao-no tane-maki
morning.glory-GEN seed-sowing
'sowing seeds of morning glory'

As shown in (20a), *asàgao*, *tàne*, and *màk* are all accented, containing falling pitch (HL). (The presence of the accusative marker does not affect accent). But in (20b), *tane-maki* has word accent in that it contains only one accent position. Crucially, the accents of the original words *tàne* and *màk* are fused into one. This shows that *tane-maki* behaves as a word, confirming the presence of compounding.

Another piece of evidence for compounding comes from *rendaku* (sequential voicing), as observed in (3) (repeated below):

(21) kireena mati-dukuri
clean town making
'construction of a clean town'

Here, the verb for 'make' is originally *tukur-*, and the sound change in *dukuri* above is due to *rendaku* voicing, a hallmark of compounding (cf. Tsujimura 2007: 50ff; Ito & Mester 2003; Kubozono 2005, among others).

As a mechanism for the compounding in (3) and (22), Kageyama (1993: 335) does not endorse his own incorporation analysis that he entertains for (16b). The main reason seems to be accent: (12c), (14), and (16b) have phrasal accent but (19a) and (19b) have word accent, and Kageyama seems to be assuming that a syntactic derivation should always result in phrasal accent and cannot result in word accent. However, a syntactic derivation like incorporation *can* result in word accent. Consider the following example of a verb-verb compound:

(22) [doa-o osi]-tuzuke(ru)
door-ACC push-continue
'to keep on [pushing the door]'
(cf. Kageyama 1989; 1993; Nishiyama 2008; Fukuda 2012)

There is a consensus in the literature that the compound in (22) is formed syntactically (by verb incorporation).[10] Crucially for the current context, *osi-tuzuke(ru)* has word accent.[11] It might be that verb incorporation and noun incorporation (if Japanese has both) have different mechanisms. But to the extent that (19) and (22) share common features (i.e., a word-accented compound containing a phrase), there is no reason for analyzing them separately, i.e., forming the compound in (19) in the lexicon and forming the compound in (22) in the syntax, as Kageyama does.[12]

But a question remains: why do (12c), (14), and (16b) have phrasal accent, while (19a) and (19b) have word accent, if they are all formed by incorporation? One prominent difference between phrasal-accented phrasal compounds as in (12c), (14), and (16b) and word-accented phrasal compounds as in (3), (18), (19) is that while the verbal noun in the former is a Sino-Japanese word, the verbal noun in the latter is of native origin. But there are a few cases of phrasal-accented noun-incorporated compounds with a native verbal noun:

(23) a. tosyo : kasi-dasi
book lend-let.out
'checking out books'

b. bentoo : moti-komi
lunch.box hold-let.in
'bringing lunch box in' (Kageyama 1993: 229)

The obvious difference between (3), (18), (19) and (23) is that the latter involve a nominalized compounded verb-verb predicate. Thus, one suspects that what's

[10] See Kageyama (1989) and Nishiyama (2008) for details. One piece of evidence for the syntactic derivation is that the complement can be passivized:

(i) doa-ga os-are-tuzuke-ta
door-NOM push-Pass-continue-Past
'The door kept being pushed.'

Given that passivization happens in the syntax, the compound *os-are-tuzuke* is formed in the syntax as well.

[11] Specifically, while *os* and *tuzuke* are inherently unaccented, the compound *osi-tuzukè(ru)* has accent in the right-hand element. The accentuation rule of verb-verb compounds in Japanese is to place accent in the right-hand element, regardless of the accent pattern of the original words. *(ru)* is added at the end to derive the present/citation form of the compound, and [i] after *os* is a linking element. See Nishiyama (2016) for details.

[12] Remarkably, the observation that (19) and (22) are parallel goes back to Sakakura (1952: 114).

going on is right-branching compounding accentuation. Recall from (10) that right-branching compounds have phrasal accent.

Let us suppose, therefore, the following:

(24) When a complement incorporates into a Sino-Japanese predicate, the predicate is reanalyzed as right-branching, resulting in phrasal accent.

Intuitively, both right-branching and Sino-Japanese words are 'heavy' in a sense. Recall from §2 that right-branching compounds have phrasal accent *for ease of processing*. A similar situation might hold in Sino-Japanese verbal nouns. For our purpose it suffices to capture the "reanalysis" above as resegmentation; while *ryokoo* 'trip' is usually taken as monomorphemic, it is analyzed as bimorphemic *ryo-koo* after incorporation.

In fact, Sino-Japanese words in general consist of bound roots (like *philosophy*), but this is not the only basis for (24); recall that with Sino-Japanese nouns, we have a word-accented compound as *doitu-bùngaku* 'German literature'. Probably the predicate-argument relation is important, so that, when noun incorporation happens, phrasal accent is required to make the morpheme boundary (or *phrasal* boundary) explicit. This requirement is removed when the predicate is of native origin, for it is easier to recognize. It is well known that phonological rules in Japanese apply differently in native words and Sino-Japanese words (e.g., *rendaku* sequential voicing; cf. Tsujimura 2007: 50ff; Ito & Mester 2003;Kubozono 2005, among others). What is special about (24) is that it is limited to predicates.[13]

3.2.2 Conditions on modifier stranding

In the last subsection, we discussed compounds with a stranded modifier. But modifier stranding is not always possible, and this subsection offers conditions on modifier stranding.

[13] As mentioned in note 5, an adjectival noun can also be the right-hand element of a noun-incorporated pseudo compound. The majority of adjectival nouns are Sino-Japanese, but there are also certain instances of adjectival nouns of native origin. As Kageyama (1993: 241) notices, whether Sino-Japanese or of native origin, an adjectival noun can be the right-hand element of a noun-incorporated pseudo compound (with phrasal accent) One example of an adjectival noun of native origin is the following:

(i) hyoozyoo : yutaka-na hito
expression rich-MOD person
'a person with diverse facial expressions' (*na* is a modifier marker)

In Nishiyama (1999), I argued that adjectival nouns (nominal adjectives in the terminology of Nishiyama (1999)) are bimorphemic (like compounds), and this might be the reason why (i) has phrasal accent, though the incorporation host (i.e., *yutaka*) is of native origin.

Consider first the following examples:

(25) a. uma-nori
horse-riding
'horseback riding'

b. *[ookina uma]-nori
big horse-riding (S&K: 471)

c. *[titi-no uma]-nori
father-GEN horse-riding
'riding on father's horse'

Why is modifier stranding impossible here, in contrast to the compounds in (19)? Kageyama (1993: 334) notes that the incorporated noun in (19) (with a stranded modifier) is a *relational noun* and needs further specification. One typical case of relational nouns is a parent, i.e. a noun whose meaning is defined only in relation to a child. In the same way, a grave is so-named only when it is known that somebody is buried there, and every seed is a seed of some kind of plant. A horse, in contrast, is not such a relational noun.

There is another condition. Consider (3), repeated below:

(26) [kireena mati]-dukuri
clean town making
'construction of a clean town'

Here the noun *mati* 'town' is not a relational noun, but modifier stranding is possible. One thing to notice here is that the modifier has a limited semantic range: instead of *kireena* 'clean', one can also use *sumiyoi* 'comfortable' or *zizokukanoona* 'sustainable', but not *kyodaina* 'giant,' in this kind of compound.

Thus, we are dealing here with a construction based on a template, i.e., [*Xish town*]-*construction*, where X has a positive (or ecological) meaning. This is reminiscent of the contrast between [*American history*]*teacher* versus *[*recent history*]*teacher* (Bresnan & Mchombo 1995: 193f). Carstairs-McCarthy (2002: 81f) cites similar examples like [*open door*]*policy* versus *[*wooden door*]*policy*, and says that the left-hand element must be a *cliché* for a left-branching compound like [*open door*]*policy* to be possible.

How an expression is recognized as a cliché is purely a matter of pragmatics and beyond the scope of this paper.[14] Bresnan & Mchombo (1995) argue that

[14] For example, [*small car*]*driver* is possible while *[*green car*]*driver* is not, because *small car* is a cliché but *green car* is not. However, as Sproat (1993: 251) notes, in an imaginary world in

what looks like a phrase in phrasal compounding is lexicalized, but this 'lexicalization' can be instant or impromptu, for it accommodates "context-dependent innovation".

I take the two conditions mentioned above (one semantic, the other pragmatic) as the output conditions on the construction [$_{XP}$ Mod X]-X; when a construction with this schema does not meet these conditions, it is filtered out. In this sense, instantiations of this construction are independent of the mechanism for compounding. Whether the mechanism is syntactic incorporation or morphological merger, it produces the construction, obeying ordinary principles imposed on it. It is only after the construction [$_{XP}$ Mod X]-X is produced when the semantic and pragmatic conditions become relevant.

When there is no predicate involved in compounding, there can be no noun incorporation. Therefore, phrasal compounding should be impossible in such a case. This prediction is generally confirmed:

(27) *[doitu-no bungaku]-kyookai
Germany-GEN literature-association
'Association of German Literature'

(27) is a compound of *doitu-no bungaku* 'literature of Germany' and *kyookai* 'association,' and it cannot mean 'Association of German Literature.' It has a meaning of 'German Association of Literature,' but it is derived from a different structure *doitu-no* [*bungaku-kyookai*].

However, when a cliché is involved, phrasal compounding becomes possible even without noun incorporation:

(28) a. [tiisana sinsetu]-undoo
small kindness campaign
'campaign for doing small kindnesses'

b. [midorino hane]-bokin
green feather fund.raising
'fund raising for restoration of plants' (cf. Kubozono 1995: 129)

This is reminiscent of examples like *[open door] policy* vs. **[wooden door] policy* we saw above, and strongly suggests the relevance of cliché.[15]

which a gasoline rationing scheme is based on the color of one's vehicle, [*green car*]*driver* will be acceptable. This is a typical characteristic of pragmatics, namely how language is used in the actual world.

[15] Although this cliché account captures modifier stranding as in (28), it cannot be extended to the case with a relational noun as in (19), because relational nouns are defined semantically,

Returning to noun-incorporated compounds resulting in phrasal accent as in (12c), (14), and (16b), modifier stranding in such examples is less constrained than in the ones resulting in word accent discussed in this section. Thus, as we saw in (14), a demonstrative and a relative clause can be stranded. It is true, as S&K (p. 471) note, that the following example with an adjective is ungrammatical:

(29) ?*[utukusii yooròppa] : ryokoo-tyùu (cf. 12c)
beautiful Europe traveling-while
'while traveling in beautiful Europe'

However, with a cliché complement, adjective stranding seems possible. The following example is constructed from (3) by replacing the native words by Sino-Japanese words with a similar meaning:

(30) [kireena tosi] : kensetu
clean town construction]]
'construction of a clean town'

In contrast to (29), (30) is acceptable. So when a cliché is involved, an adjective can be stranded in the formation of noun-incorporated compounds resulting in phrasal accent. But if S&K's observations are correct, a demonstrative and a relative clause cannot be stranded, and this contrasts with the formation of noun-incorporated compounds resulting in word accent, which requires a relational noun or a cliché for modifier stranding, as we saw in the last subsection. Why the difference?

I hypothesize that the phrasal accent that results in the formation of noun-incorporated compounds involving a Sino-Japanese predicate makes the compounding less tight, as attested by a pause that can intervene between the left-hand and right-hand elements, and this renders the syntactic constituency easy to recognize. This might make modifier stranding in this case less constrained. In the last subsection I stated that noun-incorporated compounds involving a Sino-Japanese predicate are harder to process than those involving a predicate

while cliché is defined pragmatically. Jaklin Kornfilt (p.c.) suggests that, if phrasal compounding is possible even if the right-hand element is not a predicate, we can dispense with noun incorporation altogether as a mechanism to derive phrasal compounds. But it is not clear whether the two types of phrasal compounds—ones whose right-hand element is a predicate and ones whose right-hand element is not a predicate—are derived by the same mechanism. First, the former is more productive. Second and relatedly, although the pragmatic condition (being a cliché) can be relevant in both types of phrasal compounds, the former involves another condition not observed in the latter: the relational noun condition as in (19).

of native origin, and that this results in phrasal accent on the former. The conjecture in this subsection implies that the resulting phrasal accent "promotes" noun-incorporated compounds involving a Sino-Japanese predicate to an advantageous position for processing, and makes modifier stranding easier for them.

3.2.3 Noun-incorporated compounds vs. synthetic compounds

At this point, one might wonder how the noun-incorporated compounds discussed so far are related to the so-called synthetic compounds (in English) like *mountain climbing* or *truck driver*. Synthetic compounds are conventionally defined as compounds in which there seems to be a thematic relation between the two parts. As is well known, there is a long debate over whether *truck driver* has the structure/derivation of *[truck] [driver]* or *[[truck driv]er]* (cf. Roeper & Siegel 1978;Lieber 1983;Spencer 1992: 324ff;Ackema & Neeleman 2004, and Harley 2009, among others). I remain neutral regarding the situation in English, but in this subsection I argue that in Japanese, there is another type of compounds that look like synthetic compounds but are *not* formed by noun incorporation.

First, recall from (12) (adapted):

(31) a. yooroppa-ryòkoo
Europe-traveling
(real compound, with word accent)

b. yooròppa : ryokoo
Europe traveling
(pseudo compound, with phrasal accent)

Both (31a) and (31b) are formed by compounding *yooroppa* 'Europe' and *ryokoo* 'traveling.' The former results in word accent, and the latter in phrasal accent. In the terminology of this paper, the former is a real compound and the latter a pseudo compound.

Although (12a) and (12c) are synonymous, there is a case where there is a semantic difference between a real compound and a pseudo compound consisting of the same elements. Consider:

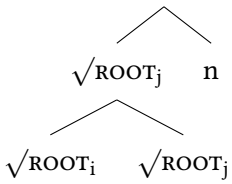
(32) a. katee-hòomon (real compound)
home-visiting
'a teacher's visit to a pupil's home' (specialized meaning)

b. katee : hoomon (pseudo compound)
home visiting
'a home visit (compositional) (S&K: 478)

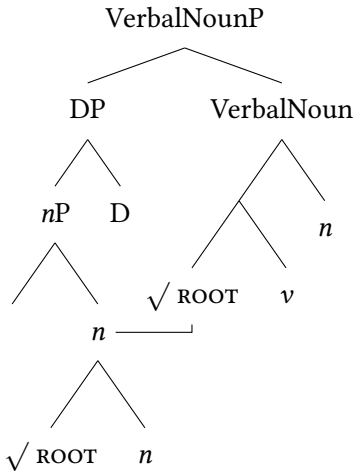
As noted by S&K, (32a) with word accent has a specialized meaning of 'a teacher's visit to a pupil's home,' but (32b) with phrasal accent has a compositional meaning.

To capture the above differences, I propose that real compounds and noun-incorporated compounds have the following structure and derivation:

(33) a. real compounds, (12a) and (32a)



b. noun-incorporated compounds, (12c) and (32b)



I assume that roots are categorially neutral and a functional head like *n* or *v* categorizes the root (cf. Marantz 1997). *n* and *v* form phases, in whose complement the semantics is fixed (cf. Embick 2010). Therefore, (33a), the structure for (32a), can have a specialized meaning.[16] In contrast, in (33b), the structure for (32b),

[16] Strictly speaking, (33a) is a structure for a dvandva like *oya-ko* 'parent-child,' and the com-

the meaning of *katee* is fixed by *n* before incorporation. Therefore, subsequent incorporation has no semantic effect and (32b) has a compositional meaning. In (33b), only the categorized root is incorporated. Therefore, any other parts of the DP (if there are any) are stranded. This is what we observed in (3), (14), (19), and (30).

One implication of the above analysis is that real compounds like *yooroppa-ryokoo* 'Europe-traveling' (12a) and *katee-hoomon* 'home-visiting' (32a), although they look like synthetic compounds like *mountain climbing*, do *not* involve noun incorporation. In other words, despite appearances, there is no thematic relation between the right-hand element (apparent predicate) and the left-hand element (apparent argument) in (12a) and (32a). In this sense, there is no structural difference between *yooroppa-ryokoo* 'Europe-traveling' and *yooroppa-rengoo* 'the European Union,' and referring to the former as a synthetic compound is in fact a misnomer. Any relationship in these compounds is established based on our world knowledge after the structure in (33a) is constructed.

Unlike (12c) and (32b), when noun-incorporation involves a verbal noun of native origin, the resulting compound has word accent. Therefore, one cannot tell whether the structure is (33a) or (33b). The only way to tell is whether there is modifier stranding. Thus, when there is modifier stranding, we can safely say that the structure is (33b), involving noun incorporation. However, without modifier stranding (like (18)), a compound with a verbal noun of native origin is simply ambiguous between (33a) and (33b). Wiese (2008) takes a similar position regarding synthetic compounds in German. What is important from a cross-linguistic point of view is that, when a Sino-Japanese verbal noun is involved, we can differentiate the two kinds of compounds, namely word-accented compounds and phrasal-accented compounds. Given that all the phrasal-accented compounds discussed in §3.1—compounds with a Sino-Japanese verbal noun with a temporal suffix attached—have a thematic relation, while many of the word-accented compounds (like *doitu-bungaku* LHH-HLLL 'German literature') do not, accentuation can be a diagnostic when the analysis can be ambivalent (as in the case involving *yooroppa* 'Europe' and *ryokoo* 'traveling' in (12) that tells us whether there is a thematic relation within a compound (as (12c)) with the structure of (33b)) or not (as (12a) with the structure of (33a)).

Although I remain agnostic about whether the Baker-style incorporation is involved for the compounding in question (cf. note 9), there is one piece of evidence for this approach. Consider the following example:

pounds in (12a) and (32a) have a more articulated structure as proposed by Ito & Mester (2003: 83f). But I abstract away from this issue and (over)simplify the structure.

(34) [(*osanai) te]-dukuri
childish hand-making
'(*children's) hand-made' (Sugioka 2005: 220)

te-dukuri 'hand-made' itself is well-formed, but it cannot have a stranded modifier. This is a case where an adjunct (the instrumental) constitutes the left-hand element of a compound. Given that only arguments can undergo (the Baker-style) noun incorporation, it is expected that a compound containing an instrument cannot have the structure in (33b); it must have the structure in (33a). Since only the structure in (33b) allows modifier stranding, it is expected that *te-dukuri* 'hand-made' does not allow modifier stranding, which is the case, as in (34).[17]

Admittedly, there is a pragmatic condition (i.e., being a cliché) for modifier stranding as discussed in the previous subsection, and (34) might be ruled out by that condition. But modifier stranding is systematically not observed with instrumental compounds. For example, another case of an instrumental compound is *enpitu-gaki* 'pencil-written, written with a pencil', but this also does not allow modifier stranding. This is expected if the Baker-style incorporation is involved.

Phases are assumed to be where not only semantics but also phonology is fixed. However, the structure in (33b) results in either word accent or phrasal accent, depending on whether the right-hand element is 'heavy' (i.e., right-branching or Sino-Japanese) or not. Besides, consider:

(35) a. tàne 'seed' (accent on the first mora)
b. [asàgao-no tanè]-maki
morning.glory-GEN seed sowing
'sowing seeds of morning glory'
(accent on the second mora of *tane*) = (19b)

tàne 'seed' has accent on the first mora by itself (35a). But when incorporated, accent is on the second mora (35b). This suggests that accentuation and incorporation go hand in hand, both applying after syntax at PF. Specifically, if we assume that accent in Japanese is not inherently specified for each word, but that accentuation applies to the structure obtained after all the morphological derivations are complete (cf. Kubozono 2008;Nishiyama 2010),[18] there is no accent shift

[17] For this kind of argument, it is immaterial whether the modifier *osanai* 'childish' is structurally an adjunct or a specifier. It is the adjunct (instrumental) status of *te* 'hand' that is crucial.

[18] In (4c), I introduced the traditional view that the position of accent is specified for each noun in Japanese for expository purposes. In Kubozono's 2008 alternative view, nouns in Japanese are accentuated by the default antepenultimate accent rule, and nouns whose accent is not

from (35a) to (35b). In (35b), *tane* receives accent on the second mora after compounding. In this sense, S&K's terminology '*post*-syntactic compounds' seems really appropriate (although their original analysis is restricted to cases involving a Sino-Japanese verbal noun). Also, this analysis lends support to Chomsky's (2001) 2001 conjecture that head movement is not part of narrow syntax.

4 Phrasal compounds without noun incorporation

As we saw in (28), there are examples of phrasal compounds whose right-hand element is not a predicate, i.e., phrasal compounding without noun incorporation. This section presents three other types of phrasal compounding without noun incorporation, namely natural coordination (§4.1), suffixes/enclitics (§4.2), and prefixes/proclitics (§4.3).

4.1 Natural coordination

The following examples contain coordination as the "phrasal" part of phrasal compounds:

(36) a. LHHHH LHH-HLLL
[karaoke to geemu]-taikai
[karaoke and game]
'contest for taikai' (Kageyama 2009: 518)

b. HL L LHH-HHHLLL
[bizyo to yazyuu]-syookoogun
[beauty and beast]-syndrome
(used in a blog as synonymous to the Stockholm Syndrome)

Specifically, the examples in (36) involve 'co-compounds' or 'natural coordination' forming a conceptual unit (e.g., *father-mother* denoting parents, cf. Wälchli 2005), again a kind of cliché. This usage of co-compound extends the original terminology (a.k.a. dvandva), which does not contain an overt conjunction.

By considering contrasts in accentuation, we can confirm that in (36a), *[karaoke to geemu]* is a phrase and *geemu-taikai* is a compound. The contour of *[karaoke to geemu]* is LHHHH LHH, with two rising pitch accents, which is typical for

antepenultimate (including unaccented nouns) are lexically specified as such. Such specifications and the default rule are realized after all the morphological derivations are complete. For accent in verbs in Japanese, see Nishiyama (2010).

phrases. The word *taikai* 'content' is inherently unaccented (LHHH), but *geemu-taikai* has the contour LHH-HLLL, showing that the accent falls on the first mora of *taikai*. As we saw in §2, this behavior is typical of compound accentuation.

As we saw in (9) (repeated below), right-branching compounds, which have phrasal accent, do not allow a coordinate phrase as the left-hand element:

(37) *doitu to huransu : bungaku-kyookai
Germany and France literature association
'associations of literature in Germany and France'

(9) was cited in §2 to show that right-branching compounds, despite having phrasal accent, are not phrases but words. Since the coordinated phrase in (9) is not a natural coordination, (9) cannot be ruled in as a phrasal compound.

4.2 Suffixes (enclitics)

In the following examples, a bound morpheme attaches to a phrase:

(38) a. LHH-HH
[dai-kigyoo-no syatyoo]-**kyuu**
big-company-GEN president equivalent
'equivalent to the president of a big company' (Kageyama 1993: 327)

b. LH-HLL
[sakunen-no ziko]-**irai**
last.year-GEN accident
'since last year's accident' (cf. Kubozono 1995: 131)

c. LH-HH
[atama-ga ookii hito]-**yoo**
head-NOM big people
'for the use by big-headed people'

Accent is specified in the last part to show that the sequence consisting of the host + *-kyuu*/*-irai*/*-yoo* has word accent, and therefore that the latter morphemes are integrated as part of the word. One can say that *-kyuu*, *-irai*, and *-yoo* are suffixes, but they may better be analyzed as clitics, which are often characterized as phrasal affixes. If so, (38) involve cliticization rather than compounding. The choice of terminology is immaterial here.

The enclitics in question are originally Sino-Japanese bound roots which have turned into clitics. As expected, there are also proclitics which originate from Sino-Japanese bound roots; those are discussed in the next section.

4.3 Prefixes (proclitics)

Previous studies of the morphemes discussed in this subsection (Poser 1990; Kageyama 2001; 2009) have referred to them as prefixes. However, based on the criterion mentioned in the previous subsection, namely that these morphemes attach to an entire phrase, it is preferable to call them proclitics. They are illustrated by the following examples:

(39) a. HL LHH-HLLL
zen : gaimu daizin
ex foreign minister

b. HL LHHH
han : taisei
anti establishment

Note that the examples have phrasal accent. Other proclitics with this property include *hòn-* 'this,' *mòto-* 'former,' *gèn-* 'current,' *kàku-* 'each,' *bòo-* 'a certain,' *dòo-* 'above-mentioned,' *ryòo-* 'both,' *ko-* 'deceased,' *hi-* 'non.'

The proclitics can attach to coordinate structures, revealing their phrasal nature:

(40) a. LHHHLL LHLLL late Hasegawa-mr and Uemura-mr
ko : [Hasegawa-si to Uemura-si]
'the late Mr. Hasegawa and Mr. Uemura' (adapted from Kageyama 2001: 265)

b. LHH H LHHH
gen : [syusyoo to gaisyoo]
current prime.minister and foreign.minister
'current prime minister and foreign minister'

As in (36), accent reveals the phrasal nature of the coordinate structure.

In addition, the inherently anaphoric proclitic *dòo-* 'above-mentioned' violates the anaphoric island constraint, again strongly suggesting its phrasal status:

(41) daitooryoo-wa asu yuukoo-zyooyaku-ni tyooinsuru doo:
president-TOP tomorrow amenity-treaty-DAT sign said
zyooyaku: saisyuu-an niyoruto
treaty final-version according.to
'The President is going to sign the amenity treaty tomorrow. According to the final version of the said treaty,…' (Kageyama 2001: 258)

Here, *doo-* in the second sentence refers to *yuukoo* 'amenity' of *yuukoo-zyooyaku* 'amenity treaty'. *doo-* itself is also a part of the compound *[doo: zyooyaku: saisyuu-an].* In other words, both the anaphor and the antecedent are a part of a word, violating the anaphoric island constraint, which says that anaphoric relations cannot be established within a word.

While natural coordination in (36) and enclitics in (38) result in word accent, proclitics in (39) and (40) result in phrasal accent. Again this may be related to the fact that proclitics tend to yield a right-branching structure (cf. (10)). Even with a binary structure as in (39b), the clitic status of the left-hand element makes the right-hand element relatively heavy, and this might induce the reanalysis of the right-hand element as bimorphemic, as with the case of noun-incorporated compounds with Sino-Japanese predicates discussed in (24).

5 Reconsidering "Word Plus"

Kageyama (1993; 2001; 2009) proposes the new term Word Plus, which covers all the phrasal-accented compounds minus what he and S&K term post-syntactic compounds as discussed in §3.1. The level of Word Plus comes between a word and a phrase, and this is meant to capture the dual (i.e., word and phrasal) nature of the examples in question.

In my view, the notion Word Plus subsumes heterogeneous examples. First of all, many instances of Kageyama's Word Plus are right-branching compounds of the type in (8b), which cannot be analyzed as involving a phrase, as we saw in (9). This leaves us with prefixes (discussed in §4.3) and non-right-branching pseudo compounds. Let us discuss them in turn.

We saw in §4.3 that the proclitics in question have phrasal nature, in that they can attach to a phrase. But Kageyama argues that they have a word-like nature as well. The evidence comes from ellipsis:

(42) *A-wa gen : ~~kaityoo-to~~ ~~suruai-de~~, B-wa zen :
A-TOP current president-with acquainted-COP B-TOP ex
kaityoo-to siriai-da
president-with acquainted-COP
'A ~~is acquainted with the current president~~, and B is acquainted with the ex-president.' (Kageyama 2001: 251)

The strikethrough indicates (cataphoric) ellipsis under identity. If *gen* is replaced with *genzai-no* 'current-GEN' and *zen* is replaced with *mae-no* 'former-

GEN', the sentence becomes grammatical. On the assumption that ellipsis is possible with phrases, Kageyama argues that (42) is evidence for the word-like nature of the proclitics in question.

However, (42) is independently ruled out, because a clitic *gen=*- – a bound morpheme – does not have a host after ellipsis. Alternatively, (42) is accounted for by assuming that the presence of the genitive is required for recovering the elided part. This is analogous to the following contrast in English:

(43) a. John's dog is bigger than Bill's ~~dog~~.
b. * John's dog is bigger than Bill~~'s dog~~.

Lobeck (1990) proposes an analysis of ellipsis based on Spec-Head agreement, but regardless of the validity of this analysis, whatever account captures the contrast in (43) would also account for (42).

Another piece of evidence that Kageyama cites for his observation that the proclitics in question are word-level (as opposed to phrasal level) entities is the following:

(44) a. yuumee-na haiyuu
famous-MOD actor
b. ??yuumee-hàiyuu
famous-actor
c. tihòo-no tòsi
province-GEN city
d. tihoo-tòsi
province-city

(45) a. bòo : [yuumee (*na) haiyuu]
certain famous MOD actor
'a certain famous actor'
b. kàku : [tihoo (*no) tòsi]
each province GEN city
'each provincial city' (Kageyama 2001: 249f)

yuumee-na haiyuu 'famous actor' (44a) and *tihòo-no tòsi* 'provincial city' (44c) are phrases, and the former cannot be a compound (44b), but the latter can (44d). The examples in (45) illustrate cases with proclitics, and the modifier marker *na* and the genitive marker *no* cannot appear here. This means that the prolicitcs

cannot attach to a phrase (with *na* or *no*) but must attach to a word (i.e. here, to a compound). The contrast between (44b) and (45a) is telling: the compound *??yuumee-hàiyuu* does not exist by itself, but with the proclitic *bòo-*, the compound must be used. If *bòo-* and *kàku-* are clitics, they should be able to attach to a full phrase, and (45) should be possible with *na/no*, contrary to fact. This, according to Kageyama, is evidence for the word-like nature of the proclitics in question.

The above point is well taken, but cross-linguistically, the distinction between clitics and affixes is often not categorial but a matter of degree. For example, Romance clitics are often analyzed as being on a grammaticalization path towards agreement markers (namely suffixes) (cf. Suñer 1988, among others). Thus, the hybrid nature of the morphemes in question might simply reflect the hybrid nature of clitics in general, and this alone is not sufficient as a motivation for postulating a novel level of Word Plus.

Non-right-branching pseudo compounds are of two types: binary compounds with phrasal accent and left-branching compounds with phrasal accent. The former is illustrated by the following example:

(46) kyùusyuu : nànbu
Kyuusyuu southern.part
'Southern Kyusuyu' (Kubozono 1995: 70, also cited in Kageyama 2001: 261)

We have been assuming that (exceptional) phrasal accent in compounds is due to a right-branching structure. So why does (46) have phrasal accent, unlike ordinary compounds (with word accent), although it is not right-branching?

One important point is that (46) optionally can have the genitive between the two parts of the construction, and when this happens, we have phrasal accent, as expected:[19]

(47) kyùusyuu-no nànbu
Kyuusyuu-GEN southern.part
'southern part of Kyuusuyuu'

It is reasonable to analyze (46) as involving genitive deletion. Therefore, (46) is not a compound in a strict sense, but is better called a phrase in disguise.

Since the genitive usually cannot be left out, I conjecture that it is the *cliché* nature of (46) that makes genitive deletion possible. Thus, Kyuusyuu is an island

[19] Kubozono (1995: 70) notes that person names also have phrasal accent. Here as well, the genitive marker used to appear between the family name and the given name; however, this usage of the genitive marker has become obsolete.

stretching from north to south, and is usually referred to as having a northern and a southern part.[20]

A similar deletion process is involved with *-teki* (repeated):

(2) dare-ga bosu-da-teki tàido
who-NOM boss-COP-like attitude
'a who's the boss attitude'

teki- usually attaches to a root and derives an adjectival noun, which requires *na* as the modifying marker as in (48a):

(48) a. hankoo-teki-na tàido
rebellion-like-MOD attitude
'rebellious attitude'
b. hankoo-teki : tàido
rebellion-like attitude
'rebellious attitude'

But as (48b) shows, the modifying marker *na* can be left out, resulting in what looks like a pseudo compound (with phrasal accent). In (2b) as well, *na* can emerge after *-teki*. The alternation between (48a) and (48b) is analogous to the presence and absence of *no* in (47) and (46).

Apparent left-branching pseudo compounds are illustrated as follows:

(49) a. booeki-gàisya : syatyoo
trading-company president
b. siritu-dàigaku : kyoozhu
private-university professor (Kageyama 2009: 518f)

The genitive deletion analysis proposed above for binary pseudo compounds can be extended to this case. These examples are also fixed expressions; they refer

[20] Kubozono's (1995: 71f) account is couched in terms of "semantic unity." Regarding this, Kageyama (2001: 261) states that "it is difficult to delimit the range of phrase-like [pseudo] compounds in term of their internal semantic relations." In the context of the current discussion, Kubozono's insight is reinterpreted as a pragmatic factor leading to cliché. It should also be noted that the examples in (46) and (47) are different from *haha no hi* 'Mother's Day' and *ama no zyaku* 'devil's advocate', which do not allow genitive deletion. Kageyama (2001: 268) cites them as Japanese equivalents of possessive compounds (e.g., *a girls' school*), and says that "those expressions are completely lexicalized." This is corroborated by the fact that they have word accent, as opposed to (46) and (47), which have phrasal accent.

to some distinguished titles, and *syatyoo* 'president' cannot be replaced by *sarariiman* 'salaried worker' and *kyoozyu* 'professor' cannot be replaced by *syokuin* 'worker' in this kind of expression.

To summarize, Kageyama's notion of Word Plus is not a natural class and should be reclassified into three distinct classes: right-branching compounds, constructions involving proclitics, and phrases involving genitive deletion.

The genitive-deletion analysis is actually suggested by Kageyama & Shibatani (1989: 163, n. 7) for right-branching pseudo compounds as in (8b). However, as we saw in (9), the right-branching pseudo compounds of the type in (8b) cannot contain a phrase. Therefore, it is unlikely that they involve genitive deletion.

In fact, in later works Kageyama (1993: 342, 2001; 2009) does not endorse his own earlier suggestion of genitive deletion mentioned above and develops the Word Plus analysis instead. In particular, he notes (2001:250f, Kageyama (2009):519) notes that partial ellipsis is impossible with pseudo compounds, although it is possible when the genitive is present.

(50) A-wa siritu-daigaku *(no) ~~kyoozyu-de~~, B-wa kokuritu-daigaku
A-TOP private-university GEN professor COP B-TOP national-university
*(no) kyoozyu desu
GEN professor COP
'A ~~is a professor~~ (of) a private university, and B is a professor of a national university.' (adapted from Kageyama 2009: 519)

This might be taken as evidence against the genitive-deletion analysis. However, as mentioned after (42), the contrast in question is accounted for by assuming that the presence of the genitive is required for recovering the elided part. Thus, the ungrammaticality of (50) without *no* is not an obstacle for postulating genitive deletion for deriving left-branching pseudo compounds as in (49).

6 Conclusions

This paper has discussed phrasal compounds in Japanese, reanalyzing and reclassifying examples discussed in the previous studies in this area. One important mechanism for phrasal compounding is noun incorporation, although I leave open the exact mechanism of this process. I have extended Shibatani & Kageyama's (1988) and Kageyama & Shibatani's (1989) analysis of post-syntactic compounds (involving Sino-Japanese verbal noun) to verbal nouns of native origin. A noun-incorporation analysis for compounds involving verbal nouns of native ori-

gin has been proposed by Sugioka (2002), but I have refined the analysis. Specifically, compounds involving verbal nouns of native origin are structurally ambiguous, with one structure involving noun incorporation and the other without noun incorporation. Only when there is modifier stranding can we be certain that noun incorporation is involved.

Through the classification of phrasal compounds, I have claimed that Kageyama's (1993; 2001; 2009) notion of Word Plus should be reclassified into three existing types, namely right-branching compounds, constructions involving proclitics, and phrases involving genitive deletion.

Here is a table summarizing the proposed analyses and classes of phrasal compounds in Japanese:

Table 1: Summary and representative examples of types of phrasal compounds in Japanese

noun incorporation Sino- Japanese verbal noun	verbal noun of native origin	
yooròppa : ryokoo 'Europe traveling' (12c)	*asagao-no tane-maki* 'sowing seeds of morning glory' (19b) relational noun	*kireena mati-dukuri* 'construction of a clean town' (3) cliché

NO noun incorporation modifying structure	coordinate structure	prefix/ proclitic	suffix/ enclitic
tiisana sinsetu-undoo 'campaign for doing small kindness' (28a) cliché	*bizyo to yazyuu-syookoogun* 'beauty and beast-syndrome' (36b) cliché	*dai-kigyoo-no syatyoo-kyuu* 'equivalent to the president of a big company' (38a)	*zèn : gaimu-dàizin* 'ex-foreign minister' (39a)

Phrasal compounds are classified primarily by whether noun incorporation is involved or not. If it is, a further division is made according to whether the predicate is of Sino-Japanese or of native origin. With a Sino-Japanese verbal noun, the resulting compound has phrasal accent. In contrast, with a verbal noun of native origin, one cannot tell whether the compound is formed by noun incorporation or not without modifier stranding. This is why the above examples have

modifier stranding, to make the case for the phrasal status of the complement of the verbal noun. There are two licensing conditions for modifier stranding: the complement of the predicate—the left-hand element of the compound—should be a relational noun or a part of a cliché.

If no noun incorporation is involved, there are four subclasses. With modifying structures and coordinate structures, the licensing condition is again cliché. Prefixes/proclitics and suffixes/enclitics originate in Sino-Japanese bound roots, but they have become clitics, so that they attach to a phrase. Given the ability of clitics to attach to entire phrases, they don't have to obey any conditions (such as cliché) in order to participate in the formation of phrasal compounds.

Lastly, I summarize and clarify my standpoint regarding the relationship between accent and syntax. As we saw in (8b), *dòitu : bungaku-kyòokai* 'German Association of Literature' has phrasal accent, but is not a phrasal compound. Conversely, there are cases of phrasal compounds with word accent. *kìreena mati-dùkuri* 'construction of a clean town' in (3) has word accent in the *mati-dùkuri* 'city-making' part, but it is a phrasal compound as a whole. Furthermore, *yooròppa : ryokoo-tyùu* 'while traveling in Europe' in (12c) has phrasal accent but is analyzed as a compound. These situations manifest a kind of syntax-phonology mismatch and might give an impression that accent is not a reliable diagnostic for determining whether a string is a word or a phrase.

However, I believe that the hypothesis that accent in Japanese reflects the syntactic status is basically correct. Specifically, whenever a string [A B] has word accent, it is always analyzed as a compound. In the case of *kìreena mati-dùkuri* 'construction of a clean town' in (3), the word status of the *mati-dùkuri* part is independently confirmed by *rendaku* sequential voicing, as we saw in §3.2.1. In this sense, the other two cases are exceptional, but not without a reason. *dòitu : bungaku-kyòokai* 'German Association of Literature' in (8b) has phrasal accent because it is a right-branching compound, which requires a special treatment for ease of processing, as we saw at the end of §2. For *yooròppa : ryokoo-tyùu* 'while traveling in Europe' in (12c), accent is really unhelpful, but the fact that an adverb cannot intervene between the two parts shows that it is not a phrase but a word, as we saw in (13). Its (exceptional) phrasal accent has been attributed to the Sino-Japanese nature of the verbal noun.

Acknowledgements

Earlier versions of this paper were presented in a meeting of the Lexicon Study Circle at Keio University and at the Workshop on Phrasal Compounds at Univer-

sity of Mannheim. I thank Carola Trips and Jaklin Kornfilt for inviting me to the workshop. For valuable comments and suggestions, I thank the audiences, Yoko Sugioka and Ichiro Yuhara. I also thank Jaklin Kornfilt for detailed and helpful comments on the earlier version of this paper. This study has been supported by grants from the Japan Society for the Promotion of Science (Grant # 23520454, 15K02470).

References

Ackema, Peter & Ad Neeleman. 2004. *Beyond morphology: Interface conditions on word-formation*. Oxford: Oxford University Press.

Baker, Mark. 1988. *Incorporation*. Chicago: University of Chicago Press.

Beard, Robert. 1991. Decompositional composition: The semantics of scope ambiguities and 'bracketing paradoxes. *Natural Language and Linguistic Theory* 9. 195–229.

Botha, Rudolf P. 1981. A base rule theory of Afrikaans synthetic compounds. In Michael Moortgat, Harry van der Hulst & Teun Hoekstra (eds.), *The scope of lexical rules*, 1–77. Dordrecht: Foris.

Bresnan, Joan & Sam Mchombo. 1995. The lexical integrity principle: Evidence from Bantu. *Natural language and linguistic theory* 13. 181–254.

Carstairs-McCarthy, Andrew. 2002. *An introduction to English morphology*. Edinburgh: Edinburgh University Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Embick, David. 2010. *Localism and globalism in morphology and phonology*. Cambridge: MIT Press.

Fukuda, Shin. 2012. Aspectual verbs as functional heads: Evidence from Japanese aspectual verbs. *Natural Language and Linguistic Theory* 30. 965–1026.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Harley, Heidi. 2009. Compounding in distributed morphology. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 129–144. Oxford: OUP.

Hawkins, John A. 1990. A parsing theory of word order universals. *Linguistic Inquiry* 21. 223–261.

Ito, Junko & Armin Mester. 2003. *Japanese morphophonemics: Markedness and word structure*. Cambridge: MIT Press.

Ito, Junko & Armin Mester. 2007. Prosodic adjunction in Japanese compounds. *MIT Working Papers in Linguistics 55: Formal Approaches to Japanese Linguistics* 4. 97–111.

Kageyama, Taro. 1989. The place of morphology in grammar: Verb-Verb compounds in Japanese. *Yearbook of Morphology* 2. 73–94.

Kageyama, Taro. 1993. *Bumpoo to gokeesee [Grammar and word formation]*. Tokyo: Hituzi Syobo.

Kageyama, Taro. 2001. Word Plus: The intersection of words and phrases. In Jeroen van de Weijer & Tetsuo Nishihara (eds.), *Issues in Japanese phonology and morphology*, 245–276. Berlin: de Gruyter.

Kageyama, Taro. 2009. Isolate: Japanese. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 512–526. Oxford: Oxford University Press.

Kageyama, Taro. 2013. Postsyntactic compounds and semantic head-marking in Japanese. *Japanese/Korean Linguistics* 20. 363–382.

Kageyama, Taro & Masayoshi Shibatani. 1989. Mozyuuru bunpoo no gokeeseeron: No meesiku kara no hukugoogo-keesee. [on word-formation in module grammar: Compounding out of genitive noun phrases]. In *Nihongogaku no sintenkai [New developments in Japanese linguistics]*, 139–166. Tokyo: Kuroshio.

Kawahara, S. 2015. The phonology of Japanese accent. In H. Kubozono (ed.), *Handbook of Japanese phonetics and phonology*, 445–492. Berlin: Mouton de Gruyter.

Kubozono, Haruo. 1995. *Gokeesee to oninkoozoo* (Word formation and phonological structures). Tokyo: Kuroshio.

Kubozono, Haruo. 2005. Rendaku: Its domain and linguistic conditions. In Kensuke Nanjo Jeroen van de Weijer & Tetsuo Nishihara (eds.), *Voicing in Japanese*, 5–24. Berlin: Mouton de Gruyter.

Kubozono, Haruo. 2008. Japanese accent. In Shigeru Miyagawa & Mamoru Saito (eds.), *The Oxford handbook of japanese linguistics*, 165–191. Oxford & New York: Oxford University Press.

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

Lieber, Rochelle. 1992. *Deconstructing morphology. Word formation in syntactic theory*. Chicago: University of Chicago Press.

Lieber, Rochelle. 2010. On the lexical semantics of compounds: Non-affixal (de)verbal compounds. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 127–144. Philadelphia: John Benjamins.

Lobeck, Ann. 1990. Functional heads as proper governors. *NELS* 20. 348–362.

Marantz, Alec. 1988. Clitics, morphological merger, and the mapping to phonological structure. In Michael Hammond & Michael Noonan (eds.), *Theoretical morphology*, 253–270. New York: Academic Press.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your lexicon. In *Proceedings of the 21st Annual Penn Linguistic Colloquium: University of Pennsylvania Working Papers in Linguistics*, vol. 4.2, 201–225.

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19. 153–197.

Mithun, Marianne. 2010. Constraints on compounds and incorporation. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 37–56. Philadelphia: John Benjamins.

Nishiyama, Kunio. 1999. Adjectives and the copulas in japanese. *Journal of East Asian Linguistics* 8. 183–222.

Nishiyama, Kunio. 2008. V-V compounds. In Shigeru Miyagawa & Mamoru Saito (eds.), *The Oxford Handbook of Japanese linguistics*, 320–347. Oxford & New York: Oxford University Press.

Nishiyama, Kunio. 2010. Penultimate accent in Japanese predicates and the verb–noun distinction. *Lingua* 120. 2353–2366.

Nishiyama, Kunio. 2016. The theoretical status of ren'yoo (stem) in Japanese verbal morphology. *Morphology* 26. 65–90.

Poser, William. 1990. Word internal phrase boundary in Japanese. In Sharon Inkelas & Draga Zec (eds.), *The Phonology-Syntax Connection*, 279–288. Chicago: University of Chicago Press.

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

Sakakura, Atsuyoshi. 1952. *Nihon-bunpoo-no hanasi [Lectures on Japanese grammar]*. Osaka: Sougensha. Revised (1974) by Tokyo: Kyouiku Shuppan.

Shibatani, Masayoshi & Taro Kageyama. 1988. Word formation in a modular theory of grammar: A case of postsyntactic compounds in Japanese. *Language* 64. 451–484.

Spencer, Andrew. 1992. *Morphological theory*. Malden: Blackwell.

Sproat, Richard. 1993. Morphological non-separation revisited: Review article: Deconstructing morphology, by rochelle lieber. *Yearbook of Morphology* 1992(1992). 235–258.

Sugioka, Yoko. 2002. Incorporation vs. Modification in deverbal compounds. *In Japanese/Korean Linguistics* 10. 495–508.

Sugioka, Yoko. 2005. Multiple mechanisms underlying morphological productivity. In Salikoko S. Mufwene, Elaine J. Francis & Rebecca S. Wheeler (eds.), *Polymorphous linguistics: Jim McCawley's legacy*, 203–223. Cambridge: MIT Press.

Sugioka, Yoko. 2008. Remarks on asymmetry and recursion in compound formation. In Tetsuya Sano, Mika Endo, Miwa Isobe, Koichi Otaki, Koji Sugisaki & Takeru Suzuki (eds.), *An enterprise in the cognitive science of language: A festschrift for Yukio Otsu*, 65–78. Tokyo: Hituzi Shobo.

Suñer, Margarita. 1988. The role of agreement in clitic doubled constructions. *Natural Language and Linguistic Theory* 6. 391–434.

Tsujimura, Natsuko. 2007. *An introduction to Japanese linguistics (second edition).* Malden: Blackwell Publishing.

Wälchli, Bernard. 2005. *Co-compounds and natural coordination.* Oxford & New York: Oxford University Press.

Wiese, Richard. 2008. Two levels of morphological structure. *Journal of Germanic Linguistics* 20(3). 243–274.