A Stress-Based Theory of Disharmonic Word Orders[*]

Hisao Tokizaki and Yasutomo Kuwana
Sapporo University and Asahikawa Jitsugyo High School

1. Introduction

In this paper, we argue that the location of stress in words is a crucial factor in determining morpho-syntactic constituent orders in languages. It is argued that constituents consisting of a complement and head, in this order, have the nature of a compound. Such constituents should have the same stress location as words and compounds in the language in which they are found. We argue that the unmarked stress location determines the head-complement order of constituents from stem-affix to clause-adverbial subordinator. This stress-based theory of word orders explains languages with disharmonic word orders as well as those with harmonic orders.

In Section 2, we discuss the typology of head-complement orders based on Dryer (2005a, b, c, d, e). In Section 3, we argue that constituents with a complement-head order behave as compounds because of the short juncture in left-branching structure. In Section 4, we argue that unmarked word-stress location (Goedemans and van der Hulst 2005a, b) determines the head-complement orders in the language. Section 5 concludes the discussion.

2. Deriving disharmonic word orders by complement-movement

2.1 Head-complement orders

First, following Dryer (1992), we define *head* as a non-branching category and *complement* as a (potentially) branching category. We call a category 'branching' if it is made by merging two syntactic objects, including affixes (cf. Williams 1981). For example, A is the head and B is the complement in X in (1a) and (1b), where X stands for any level of category from X^0 to XP.

(1) a. $[_X$ A $[_B$ …]]

b. $[_X [_B$ …] A]

Then, (1a) is a head-complement order and (1b) is a complement-head order.[1]

Next, let us consider various morpho-syntactic categories, each of which is illustrated with an example in (2).

(2) a. Affix-Stem (*un-realistic*)

b. Word-Word (*spaghetti bolognese*)

c. Noun-Modifier (*something very important*)

d. Verb-Object (*read books*)

e. Adposition-Object (*in the mood*)

f. Adverbial Subordinator-Clause (*before you go*)

In each pair of categories in (2), the left one is the head because it is non-branching. The right categories are complements because they are branching or potentially branching. For example, *books* in (2d) could be *interesting books*. Note that a branching complement means the sister of a head, derived by merging two syntactic objects, which may be morphemes. Thus, *realistic* in (2a) is also branching in that it can be analyzed as *real* + *istic*; *bolognese* in (2b) can be analyzed as *bologna* + *ese*.

Following Kayne (1994), we assume that the head-complement orders in (2) are the universal base orders, which can be changed into complement-head orders by movement of the complements to a specifier position on the left of a head. The derivation of the complement-head order from the head-complement order is shown in (3).

(3) a. $[_{XP}$ X $[_{YP}$ .. Y ..]] →

b. $[_{XP}$ $[_{YP}$ .. Y ..] X *t*]

YP moves to the specifier position of X in (3b).[2] The trace *t* in (3b) is invisible at the syntax-phonology interface because it does not have phonetic features. As Holmberg (2000) points out, from a phonological point of view the complement movement changes the right-branching structure (3a) into the left-branching structure (3b), as shown in (4) (and see Section 3.2 below).

(4) $[_{XP}$ $[_{YP}$ .. Y ..] X]

Then, the complement movement changes the head-complement order in (2) into the complement-head order shown in (5).

(5) a. Stem-Affix (*stabiliz-ation*)

b. Word-Word (*vacation-land*)

c. Modifier-Noun (*very important things*)

d. Object-Verb (*Bücher lesen* 'read books'(German))

e. Adposition-Object (*huoneese-en* 'into room' (Finnish))

f. Clause-Adverbial Subordinator (*anata-ga iku maeni* (you-Nom go before) 'before you go'(Japanese))

In (5a) and (5b), *stabilize* and *vacation* can be analyzed as *stable* + *-ize* and *vacate* + *ion*.

2.2 Universal LF hypothesis

One question that needs to be answered is what motivates the complement movement to a specifier position. We argue that complements must move to a specifier position to make a semantic unit with their heads.

According to Kayne (1994), the base structure is universally Specifier-Head-Complement across languages. Kayne argues that head-final languages such as Japanese move complements to a specifier position as we have seen in the previous section. If we assume that there is no lowering operation in syntax, including LF, then the constituent order in LF is the same as the one at Spell-Out, i.e. complement-head, in head-final languages. Then, what kind of LF representation do head-initial languages have for head-complement pairs?

We argue that LF is universal across languages. It seems to be plausible that semantic representation is the same in all languages. Moreover, the idea of a universal LF has been proposed for operator movement by Huang (1982), as shown in (6).

(6)		English	Chinese
	Base	$[_{IP}$ .. Op ..]	$[_{IP}$ .. Op ..]
	Spell-Out	$[_{CP}$ Op $[_{IP}$ .. *t* ..]]	$[_{IP}$ .. Op ..]
	LF	$[_{CP}$ Op $[_{IP}$ .. *t* ..]]	$[_{CP}$ Op $[_{IP}$ .. *t* ..]]

English moves operators including *wh*-phrases to the specifier position of C by the time of Spell-Out, while Chinese does so after Spell-Out. Thus, the surface orders in these two languages are different, but the LF representations are the same. Similarly, complement-movement in head-initial languages and head-final languages can be shown as in (7).

(7)		head-final languages	head-initial languages
	Base	$[_{XP}$ X $[_{YP}$ .. Y ..]]	$[_{XP}$ X $[_{YP}$ .. Y ..]]
	Spell-Out	$[_{XP}$ $[_{YP}$ .. Y ..] X *t*]	$[_{XP}$ X $[_{YP}$ .. Y ..]]
	LF	$[_{XP}$ $[_{YP}$ .. Y ..] X *t*]	$[_{XP}$ $[_{YP}$ .. Y ..] X *t*]

Complement-movement occurs by the time of Spell-Out in head-final languages and after Spell-Out in head-initial languages.

The basic assumption behind the universal LF hypothesis is that constituents with head-complement order are not interpretable at LF. Let us consider this point in detail. First, it has been argued in generative syntax that the selectional relation needs to be checked in some way. For example, Holmberg (2000) lists three ways for c(ategorial)-selection shown in (8).

(8) a. Pure f(eature)-movement (i.e. covert movement; see Chomsky 1995: Ch.4)
 b. Head movement, i.e. f-movement pied-piping the minimal word containing the relevant feature.
 c. XP-movement, i.e. f-movement pied-piping the minimal maximal category containing the relevant feature.

The distinction between (8a), (8b), and (8c) reflects two criteria for the classification of c-selection. First, the movement is covert (8a) or overt (8b, c). Second, the moved category is X^0 (8b) or XP (8c).

However, if we assume the universal LF hypothesis, the distinction between covert movement (8a) and overt movement (8b, c) is reduced to the order of movement and Spell-Out. Overt movement (8b, c) occurs by the time of Spell-Out. After Spell-Out, movement may apply to X^0 or XP, which has semantic features but not phonetic features. This movement after Spell-Out is equivalent to pure feature-movement. Thus, we do not have to specify pure feature-movement (8a) to check c-selection.

Let us return to the assumption that head-complement order is not interpretable while complement-head order is interpretable. This assumption is based on the observation that constituents with left-branching structure are more tightly connected to each other than constituents with right-branching structure. Put phonologically, the juncture between constituents is shorter in left-branching structure than in right-branching structure, as we will argue in Section 3.1. Following American structuralist linguistics and Selkirk (1984), we define 'juncture' as the relations between the segments in a sequence. Juncture shows the degrees of connectedness between segments of phonological representation, which may affect the application of phonological rules.[3] We propose that constituents must be connected to each other in order to be interpreted in LF. Then, head-complement order with right-branching structure need to have been changed into complement-head order with left-branching structure by the time the derivation reaches the LF output where semantic interpretation takes place. In the next section, we show some evidence that the juncture between constituents is shorter in left-branching structure than in right-branching structure. After that, we will propose a reason for why head-initial languages defer the movement until after Spell-Out while head-final languages move complements before Spell-Out.

3. Complement-head order as compound

3.1 Short juncture in left-branching structure

Tokizaki (2008) argues that juncture between constituents is shorter in left-branching structure than in right-branching structure. Let us consider the structures in (9).

(9) a. [[X Y] Z]
    b. [X [Y Z]]

The two structures in (9a) and (9b) are symmetrical in constituency, but they differ in phonological realization. The evidence is based on phonological changes in compounds. Japanese Sequential Voicing (*Rendaku*) and Korean *n*-Insertion occur between constituents in left-branching structure but not in right-branching structure, as shown in (10) and (11) (cf. Otsu 1980, Han 1994).[4]

(10) a. [[*nise tanuki*] *<u>sh</u>iru*] → *nise danuki <u>j</u>iru*
mock badger soup
'mock-badger soup'

b. [*nise* [*<u>t</u>anuki shiru*]] → *nise <u>t</u>anuki jiru (*<u>d</u>anuki)*
mock badger soup
'mock badger-soup'

(11) a. [[*on chən*] *yok*] → *on chən <u>n</u>yok*
hot spring bathe
'bathing in a hot spring'

b. [*kyəŋ* [*yaŋ sik*]] → *kyəŋ yaŋ sik (*<u>n</u>yaŋ)*
light Western food
'a light Western meal'

Left-branching structure in (10a) and (11a) allows phonological rules to apply between constituents, and right-branching structure in (10b) and (11b) does not. This fact can be explained by the assumption that left-branching structure has shorter juncture between its constituents than right-branching structure. The long juncture between constituents blocks phonological rules in (10b) and (11b).

An alternative explanation of the contrast in (10) and (11) is to assume that phonological rules do not change the phonology of the compound made by Merge. In (10a) and (11a), phonological rules apply to the third word in each example (*shiru* and *yok*), when it is merged with the result of the first merge (of the first and the second words, i.e. *nise tanuki* and *on chən*). In (10b) and (11b), phonological rules do not apply to the second word in each example (*tanuki* and *yaŋ*), because they have been merged syntactically with the third word and become a part of a compound (*tanuki-jiru* and *yaŋ-sik*) at the time of the second Merge.[5]

This alternative idea based on the derivational cycle is an interesting one, but I will not adopt it here because of the following data. In Japanese, Accent Deletion applies to one of the two constituents when they are merged, as shown in (12).

(12) *míso* + *shíru* → *misoshíru*
miso soup 'miso soup'

Accent Deletion applies to the first constituent *miso* and deletes its accent when merged with the second constituent *shiru*. Accent Deletion applies to the first constituent in left-branching compounds as in (13a), but not in right-branching compounds as in (13b).

(13) a. [[*nihon búyoo*] *kúrabu*] → *nihon buyoo kúrabu*

Japan dance club

'Japanese-dance club'

b. [*nihón* [*hoosoo kyóokai*]]→*nihón hoosoo kyóokai (*nihon hoosoo kyóokai)*

Japan broadcasting association

'broadcasting association in Japan'

In the first constituent *nihon búyoo*, which is the result of Merging the two words, the accent is deleted in (13a). This fact cannot be explained if we assume the derivational cycle we have seen above. The derivational cycle claims that phonological rules do not change the phonology of the compound made by Merge. The derivational cycle predicts that once *nihon búyoo* is created with its accent in the first cycle, the accent on *búyoo* should not be deleted in the next cycle that makes the whole compound *nihon buyoo kurabu*. However, this is not the case: the merge of *nihon búyoo* and *kúrabu* makes the whole compound *nihon buyoo kúrabu,* deleting the accent on *buyoo*. On the other hand, stress deletion in (13a) and non-deletion in (13b) are naturally explained with our hypothesis that juncture is shorter in left-branching structure than in right-branching structure. The constituents in the left-branching structure in (13a) are closely connected to each other and lose the left-hand stress. The constituents in (13b) are in a right-branching structure with a looser connection and thus keep their own stress. Thus, we conclude that the different phonological behaviors between left-branching compounds and right-branching compounds, shown in (10), (11), and (13) are due to junctural asymmetry, not to the derivational cycle.

The role of junctural asymmetry in compounds is further supported by data in Dutch. Krott et al. (2004) show that in Dutch, the occurrence of interfixes including *-s-* in tri-constituent compounds matches the major constituent boundary better in right-branching compounds than in left-branching compounds. They counted the occurrences of tri-constituent compounds in the Dutch section of the CELEX lexical database, which is based on a corpus of approximately 42 million words. In (14) and (15), the number of compounds with *-s-* and all interfixes are shown in parentheses after the examples.

(14) a. Interfixes at the constituent break in right-branching compounds (*-s-* 38; all 60)

[*arbeid-s-*[*vraag stuk*]]

employment+question-issue

b. Interfixes within the inner compound in right-branching compounds (*-s-* 3; all 11)

[*hoofd* [*verkeer-s-weg*]]

main+traffic-road

(15) a. Interfixes at the constituent break in left-branching compounds (-*s*- 25; all 39)

[[*grond wet*]-*s-aartikel*]

ground-law+article, constitution

b. Interfixes within the inner compound in left-branching compounds (-*s*- 13; all 50)

[[*scheep-s-bouw*] *maatschappij*]

ship-building+company

The ratio of the unmarked interfix position (14a) and (15a) to the marked interfix position (14b) and (15b) is higher in right-branching (14) (-*s*- 38÷3=12.7; all 60÷11=5.5) than in left-branching (15) (-*s*- 25÷13=1.9; all 39÷50=0.8). Comparing the ratios between unmarked and marked interfixes of right-branching 12.7 (-*s*-) and 5.5 (all) with those of left-branching 1.9 (-*s*-) and 0.8 (all), we conclude that interfixes in the unmarked position are more likely to occur in right-branching compounds than in left-branching ones. That is, interfixes occur at the constituent break more often in right-branching compounds than in left-branching compounds. In other words, interfixes are more likely to occur within the inner compound in left-branching compounds than in right-branching compounds. This result is expected if we assume that the juncture between constituents in right-branching structure is long enough for interfixes to intervene there. In left-branching structure (15), the juncture between the second word and the third is about as short as the juncture between the first word and the second. Thus, marked interfixes (15b) can occur more frequently in left-branching structure (15b) than in right-branching structure (14b). This fact supports our junctural asymmetry hypothesis.[6]

Moreover, junctural asymmetry can also be seen in morphology. Hyman (2008: 323) argues that suffixes tend to be more tightly bound to their stem than prefixes. This observation also supports the asymmetry in juncture hypothesis because $[_{\text{Word}}$ prefix $[_{\text{Stem}}$ ...]] is right-branching while $[_{\text{Word}}$ $[_{\text{Stem}}$ ...] suffix] is left-branching. Thus, cross-linguistic facts show that juncture between constituents is longer in right-branching structure than in left-branching structure.

Furthermore, the typology of adverbial subordinators also supports the junctural asymmetry hypothesis. Investigating 611 languages in the world, Dryer (2005e) points out that all clear instances of affixal adverbial subordinators (Sb) are suffixes on the verb, with no clear instances of prefixes on the verb.[7]

(16) a. $[_{\text{CP}}$ Sb $[_{\text{IP}}$ ...]] (367 languages)

b. *$[_{\text{CP}}$ Sb-$[_{\text{IP}}$ ...]] (0 languages)

(17) a. $[_{\text{CP}}$ $[_{\text{IP}}$ ...] Sb] (90 languages)

b. $[_{CP}\ [_{IP}$ ...]-Sb] (59 languages)

The fact that (16b) does not exist shows that clause-initial adverbial subordinators must be separated from IP as in (16a). This is because CP is right-branching in (16) with its immediate constituents, adverbial subordinator and IP, separated from each other by the long juncture between them. It is impossible to attach an adverbial subordinator to the following IP as a prefix. Clause-final adverbial subordinators can be attached to the preceding IP as a suffix as shown in (17b) because they merge with the IP on its left to make a left-branching structure. This is possible because the juncture between constituents in left-branching structure is short enough for adverbial subordinators to attach to the preceding clause. Thus, the data in (16) and (17) support the junctural asymmetry hypothesis.[8]

3.2 Complement-movement as compounding

Given that left-branching structure has short juncture between its constituents, we can argue that complement-movement serves as compactization. By way of illustration, complement-movement changes (18a) into (18b), where the silent copy of the moved YP (i.e. formerly *trace*) is shown in italics.

(18) a. $[_{XP}$ X $[_{YP}$ .. Y ..]]

b. $[_{XP}\ [_{YP}$ .. Y ..] $[_{X'}$ X $\mathit{[_{YP}}$ *.. Y ..]*]]

Syntactically, (18b) still has right-branching structure in X' with YP branching. However, phonologically, XP in (18b) is left-branching, assuming that silent categories and the constituent made by merging them to another constituent are invisible at PF. Since X' and the original copy of YP in (18b) are invisible as shown in (19a), XP in the phonological representation (19) is left-branching.

(19) a. $[_{XP}\ [_{YP}$ .. Y ..] $\mathit{[_{X'}}$ X $\mathit{[_{YP}}$ *.. Y ..]]*]

b. $[_{XP}\ [_{YP}$ .. Y ..] X]

Then, complement movement changes right-branching PF (18a) into left-branching PF (19b), which has short juncture between YP and X.[9] We expect that constituents with complement-head order (19b) will behave as compounds. In the next section, we argue that the main stress position in constituents with complement-head order needs to match the unmarked word-stress location in the language.

4. Typology of stress location and head-complement orders

4.1 Word-stress locations

Goedemans and van der Hulst (2005a, b) classify languages into two classes, namely languages with fixed stress location and languages with weight-sensitive stress. The two classes are divided into several subcategories according to the stress locations. The lists in (20) and (21) show the classes and subcategories and the number of corresponding languages.

(20) Fixed stress location

a.	No fixed stress (mostly weight-sensitive stress)	219
b.	Initial: stress is on the first syllable	92
c.	Second: stress is on the second syllable	16
d.	Third: stress is on the third syllable	1
e.	Antepenultimate: stress is on the antepenultimate (third from the right) syllable	12
f.	Penultimate: stress is on the penultimate (second from the right) syllable	110
g.	Ultimate: stress is on the ultimate (last) syllable	50
	Total	500

(21) Weight-sensitive stress

a.	Left-edge: Stress is on the first or second syllable	37
b.	Left-oriented: The third syllable is involved	2
c.	Right-edge: Stress on ultimate or penultimate syllable	65
d.	Right-oriented: The antepenultimate is involved	27
e.	Unbounded: Stress can be anywhere in the word	54
f.	Combined: Both Right-edge and unbounded	8
g.	Not predictable	26
h.	Fixed stress (no weight-sensitivity)	281
	Total	500

Note that the number of fixed stress languages in (20b-h: 92+16+1+12+110+50=281) corresponds to that of fixed stress languages in (21h: 281), and the number of languages with no fixed stress in (21a-g: 37+2+65+27+54+8+26=219) corresponds to that in (20a: 219). Thus, the total number of languages listed in (20) and (21) is 500 altogether.

4.2 Fixed stress locations

In this and the next section, we will outline our theory of how word-stress location determines head-complement orders. First, let us consider the languages with fixed stress locations. For example, languages with penultimate stress have words with the syllable structure in (22), where the underscore represents stress.

(22) $[_{\text{Word}} \cdots \underline{\sigma}\ \sigma]$

A phrase with head-complement order has the structure in (23), where X is the head word of the phrase and Z is the last word in the complement YP.

(23) $[_{XP}\ [_{X} \ldots \underline{\sigma}\ \sigma]\ [_{YP} \ldots [_{Z} \ldots \underline{\sigma}\ \sigma]]]$

This right-branching structure does not pose any problem in phonology: each word (e.g. X and Z) has penultimate stress in (23). However, if the complement moves to the specifier position of X the resulting structure is (24a), which is phonologically left-branching as shown in (24b), as we have argued in Section 3.2 above.

(24) a. $[_{XP}\ [_{YP} \ldots [_{Z} \ldots \underline{\sigma}\ \sigma]]\ [_{X'}\ [_{X} \ldots \underline{\sigma}\ \sigma]\ t]]$

b. $[_{XP}\ [_{YP} \ldots [_{Z} \ldots \underline{\sigma}\ \sigma]]\ [_{X} \ldots \underline{\sigma}\ \sigma]]$

We expect that the left-branching XP in (24b) will have the same stress position as a (compound) word because of the short juncture between YP and X.

We will follow Cinque's (1993) idea that the most deeply embedded element in the recursive side of a structure has the primary stress in the structure. If this is right, the primary stress goes on to Z in (23) and (24b). However, this primary stress causes a problem in (24b) because (24b) is left-branching and compound-like. If the whole XP in (24b) is considered to be a (compound) word, its main stress location, represented with bold underlining, is far back from the penultimate position in XP, as shown in (25).

(25) $[_{XP}\ [_{YP} \ldots \underline{\boldsymbol{\sigma}}\ \sigma]]\ [_{X} \ldots \underline{\sigma}\ \sigma]]$

This stress location deviates from the stress template (20) in this language. Thus, moving complement YP to the specifier position of the head X to make (24b) violates the phonological constraint on stress location. We expect that complement movement does not occur in languages with penultimate word-stress to make the complement-head order in (24b). This prediction is borne out, as we will see in Section 4.5.2. The same argument applies to fixed-stress languages with ultimate and antepenultimate stress.

On the other hand, we expect that languages with word-initial stress allow complement movement to make complement-head orders. The stress template of these languages can be represented as in (26).

(26) $[_{Word}\ \underline{\sigma}\ \sigma \ldots]$

A phrase with head-complement order has the structure in (27).

(27) $[_{XP}\ [_{X}\ \underline{\sigma}\ \sigma \ldots]\ [_{YP}\ \underline{\sigma}\ \sigma \ldots]]$

The phonology of (27) does not have a problem with respect to the stress location (26) because each word has initial stress in its own domain, X and YP. Complement YP may also move to the specifier position of X to make the complement-head order shown in (28).

(28) $[_{XP}\ [_{YP}\ \underline{\boldsymbol{\sigma}}\ \sigma \ldots]\ [_{X}\ \underline{\sigma}\ \sigma \ldots]]$

Assuming that YP is also left-branching, the main stress falls on the leftmost syllable in YP in the compound-like structure XP in (28). This stress location in XP matches the word-stress template (26) because the heaviest stress in XP falls on the initial syllable in XP in (28). Thus, as well as the head-complement order, the complement-head order is an option for initial stress languages.

To sum up, we expect that languages with word-stress fixed on the ultimate, penultimate or antepenultimate syllable will not allow complements to move to the specifier position of the head. This is because the complement movement would make the maximal projection a left-branching compound with the primary stress on the complement. This leftward stress of the derived compound would not correspond to the righthand word-stress location in the language. In languages with word-stress fixed on the initial syllable, the primary stress on the moved complement matches their word-stress template. These points are summarized in (29) and (30).

(29) Languages with penultimate (ultimate, or antepenultimate) stress: $[_{\text{Word}} \ldots \underline{\sigma}\ \sigma]$

a. Head-Complement: $[_{XP}\ [_{X} \ldots \underline{\sigma}\ \sigma]\ [_{YP} \ldots \underline{\sigma}\ \sigma]]$

b. *Complement-Head: * $[_{XP}\ [_{YP} \ldots \underline{\sigma}\ \sigma]]\ [_{X} \ldots \underline{\sigma}\ \sigma]]$

(30) Languages with initial stress: $[_{\text{Word}}\ \underline{\sigma}\ \sigma \ldots]$

a. Head-Complement: $[_{XP}\ [_{X}\ \underline{\sigma}\ \sigma \ldots]\ [_{YP}\ \underline{\sigma}\ \sigma \ldots]]$

b. Complement-Head: $[_{XP}\ [_{YP}\ \underline{\sigma}\ \sigma \ldots]\ [_{X}\ \underline{\sigma}\ \sigma \ldots]]$

The crucial assumption here is that constituents with complement-head order are a kind of compound which must have the same main-stress location as simple words in the language.

4.3 Weight-sensitive stress

Next, let us consider languages with weight-sensitive stress. For example, languages with right-edge stress have stress on the ultimate or penultimate syllable in a word, as shown in (31) (Goedemans and van der Hulst 2005b), where H/L stands for a heavy/light syllable and stressed syllables are in bold face.[10]

(31) a. (**H** L)]

b. (L **H**)]

Right-edge stress languages are different from fixed-stress languages with ultimate stress in that they allow a light syllable on the right of the stressed heavy syllable as in (31a). We argue that this flexibility of stress position allows a monosyllabic complement to move to the specifier position of the head. For example, let us consider a word with the stress pattern in (31b).

(32) $[_{\text{Word}} \ldots \sigma_L\ \underline{\sigma_H}]$

This word can serve as a stem when it is combined with an affix as in (33), where the affix is the head and the stem is its complement.

(33) $[_{Word}\ [_{Affix}\ \sigma]\ [_{Stem} \ldots \sigma_L\ \underline{\sigma}_H]]$

Suppose that this affix consists of a light syllable σ_L. Then the complement, Stem, may move to the specifier position of Affix to make the complement-head order in (34).

(34) $[_{Word}\ [_{Stem} \ldots \sigma_L\ \underline{\sigma}_H]\ [_{Affix}\ \sigma_L]\ t]$

In this structure, Stem and Affix are closely connected with each other to make a word as a whole. The stress falls on the penultimate syllable in the word (34). The penultimate stress in (34) may well occur in languages with weight-sensitive stress on the right-edge, i.e. on the ultimate or penultimate syllable. For example, Spanish has ultimate stress and penultimate stress as shown in (35).[11]

(35) a. ingl**é**s 'English'

b. c**a**sa 'house'

Penultimate stress corresponds to the stress pattern of $[_{Word}$ Stem-Affix] in (34). In fact, such languages are categorized as 'strongly suffixing' languages by Dryer (2005a). Some examples of Spanish are shown in (36).[12]

(36) a. entr**a**r 'enter' → entr**a**-da 'entrance'

b. tard**a**r 'delay (v.)' → tard**a**-nza 'delay (n.)'

In (36), the stem on the left has ultimate stress, and the derived word on the right has penultimate stress. Both stress locations are permitted in Spanish because it is a right-edge stress language.[13] This movement of Stem is not possible for languages with fixed stress locations such as the ultimate or penultimate syllable (e.g. Bantu), as we saw in Section 4.2.

4.4 Levels of complement-movement

As we saw in section 2.1, we assume that complement-movement applies at various levels of morpho-syntactic structure, from words to subordinate clauses ((2) repeated here as (37)).

(37) a. Affix-Stem (*un-realistic*)

b. Word (Head)-Word (Complement) (*spaghetti bolognese*)

c. Noun-Modifier (*something very important*)

d. Verb-Object (*read books*)

e. Adposition-Object (*in the mood*)

f. Adverbial Subordinator-Clause (*before you go*)

The head-complement orders in (37) are changed into the complement-head orders in (38) by complement-movement ((5) repeated here as (38)).

(38) a. Stem-Affix (*stabiliz-ation*)

b. Word (Complement)-Word (Head) (*vacation-land*)

c. Modifier-Noun (*very important things*)

d. Object-Verb (*Bücher lesen* 'read books'(German))

e. Object-Adposition (*huoneese-en* 'into room'(Finnish))

f. Clause-Adverbial Subordinator (*anata-ga iku maeni* (you-Nom go before) 'before you go' (Japanese))

Examination of the data in Dryer (2005a, b, c, d, e) shows that complement-movement can apply cyclically from the smallest domain (Affix-Stem → Stem Affix) to the widest domain (Adverbial Subordinator-Clause → Clause-Adverbial Subordinator) (See Kuwana and Tokizaki 2009, Tokizaki and Kuwana (in press)). Each language has a point at which it stops complement-movement. For example, Romance languages have complement-head order in Stem-Affix (38a) and head-complement order at the other levels of morpho-syntactic structure: in Word (Head)-Word (Complement) (37b), Noun-Modifier (37c), Verb-Object (37d), Adposition-Object (37e), and Adverbial Subordinator-Clause (37f). Uralic languages such as Finnish and Hungarian have complement-head order in Stem-Affix (38a), Word (Complement)-Word (Head) (38b), Modifier-Noun (38c), Object-Verb (38d) and Object-Adposition (38e), and head-complement order only in Adverbial Subordinator-Clause (37f). The table in (39) shows the complement-head orders (+) and Head-complement orders (–) in a range of languages (Jap/Kor=Japanese and Korean).

(39)

		Bantu	Romance	English	Germanic	Uralic	Jap/Kor
a.	Stem-Affix	–+	+	+	+	+	+
b.	Word (C)-Word (H)	–	–	+	+	+	+
c.	Modifier-Noun	–	–	–+	–+	+	+
d.	Object-Verb	–	–	–	–+	+	+
e.	Object-Adposition	–	–	–	–	+	+
f.	Clause-Subordinator	–	–	–	–	–	+

In (39a), Bantu languages have –+ value for Stem-Affix because Swahili is classified as a weakly prefixing language and Chichewa as a strongly prefixing language. In (39c), English and Germanic languages are assigned –+ value for Modifier-Noun because they show different word orders with respect to the kind of modifiers. These languages have complement-head orders in most modifiers and nouns: Adjective-Noun, Demonstrative-Noun and Numeral-Noun. However, according to Dryer (2005d), English has no dominant order in the order of genitive and noun. Germanic languages such as German and Dutch have Noun-Genitive order. In (39d), the –+ in Germanic languages indicates that they have Verb-Object order in main clauses and Object-Verb order in subordinate clauses.

Note that most languages, including English, German, and Finnish, are disharmonic with respect to head-complement orders, as shown in (39). It is implausible to argue that children need to learn the value of a head parameter for each category listed in (39). In Section 4.5 we argue that the stress pattern determines (dis)harmonic word orders and that children have only to learn the unmarked word-stress location in the language.[14]

The chart in (39) also shows a gradation of word orders among levels of morpho-syntactic constituents across languages. Note that the order of the example languages in (39) corresponds to their geographical location: from Africa to Asia through Europe. The geographical distribution of word orders is an interesting topic, but we will leave it for future research.

4.5 Correlation between word orders and stress location

4.5.1 Word-stress location and head-complement orders

The order of languages in the chart in (39) corresponds to word-stress location classified by Goedemans and van der Hulst (2005a, b), as shown in (40).

(40)
- a. Bantu (Swahili; Chichewa): penultimate
- b. Romance (Italian; Spanish): right-edge (ultimate or penultimate)
- c. English: right-oriented (ultimate, penultimate or antepenultimate)
- d. Germanic (German; Dutch): right-oriented (ultimate, penultimate or antepenultimate)
- e. Uralic (Finnish; Hungarian): initial
- f. Japanese/Korean: no stress

Generally speaking, stress location moves from right to left as we go down the list of languages in (40). Right-edge stress languages have ultimate or penultimate stress depending on the syllable weight. Thus, strictly speaking, penultimate (40a) is not the rightmost stress in the list (40). However, penultimate-stress languages in (40a) are less flexible than right-edge languages in (40b) in not permitting an extra weak syllable to be attached to the right end of a word, as we have seen in Section 4.2 and 4.3.

No-stress languages are listed at the bottom in (40f) because their word orders are all complement-head as shown in (39). Languages with no stress allow complement-movement in all the constituents, from words to subordinate clauses. This is possible because these languages do not have stress whose location determines whether complements can move or not.

Let us now consider why languages with right-hand stress do not allow complement-movement of large constituents such as objects and clauses. The chart in (41) is a combination of (39) and (40), with the languages in order of stress-location from left to right.

(41)

		Jap/Kor	Uralic	German	English	Romance	Bantu
	Word stress	no stress	Initial	R-ori	R-ori	R-edge	penult
a.	Stem-Affix	+	+	+	+	+	+–
b.	Word (C)-Word (H)	+	+	+	+	–	–
c.	Modifier-Noun	+	+	+–	+–	–	–
d.	Object-Verb	+	+	+–	–	–	–
e.	Object-Adposition	+	+	–	–	–	–
f.	Clause-Subordinator	+	–	–	–	–	–

The chart in (41) shows the correlation between word-stress location and complement-head orders. As stress moves to the right end of the word (from Japanese and Korean to Bantu), the domain of complement-movement, represented as +, becomes smaller (from all (a)-(f) to only (a)). The gradation of complement-head orders and its correlation with word-stress location can be explained by the size of complements and the number of syllables after the stressed syllable. We will discuss each stress type in detail below.

4.5.2 Penultimate stress

If a language has penultimate stress, it does not allow any complement to move to the specifier position. If it were to move there, a complement and the head would make a left-branching structure, which would have short juncture between the complement and the head. The whole constituent complement-head would behave like a word, and its stress location should conform to the word-stress location of the language, i.e. penultimate stress. However, complement-movement would leave the head as the last unstressed element in the constituent. As a result, stress would fall on the syllable before the penultimate syllable in the constituent, as shown in (42b).

(42) a. $[_{XP}\ [_{X} \ldots \underline{\sigma}\ \sigma]\ [_{YP} \ldots \underline{\sigma}\ \sigma]]$

b. $[_{XP}\ [_{YP} \ldots \underline{\sigma}\ \sigma]\ [_{X} \ldots \underline{\sigma}\ \sigma]]$

This dislocation of stress from the fixed position would occur even if the head is a monosyllabic element such as an affix or clitic in (43).

(43) a. $[_{XP}\ [_{X}\ \sigma]\ [_{YP} \ldots \underline{\sigma}\ \sigma]]$

b. $[_{XP}\ [_{YP} \ldots \underline{\sigma}\ \sigma]\ [_{X}\ \sigma]]$

In (43b), derived from (43a) by complement-movement, the stress would fall on the antepenultimate syllable in XP. Thus, in penultimate-stress languages, a stem cannot move to the left of an affix by complement-movement to make the Stem-Affix order; these languages have the Affix-Stem (i.e. head-complement) order as shown in (44).[15]

(44) a. *m-wia* (Swahili)

person-debt

'debtor'

b. *m-sunga* (Chichewa)

person-keep

'keeper'

In penultimate-stress languages, moving any size complement results in a marked stress location, and the complement-head order is avoided at all levels as shown in the Bantu column in (41a-f).[16]

4.5.3 Right-edge stress

Consider next languages with right-edge stress. As we saw in section 4.3, these languages may have complement-movement in the case of Stem-Affix. These languages allow penultimate stress as well as ultimate stress. Thus, suffixing is allowed as shown in (45) (=(36)).

(45) a. entr**a**r 'enter' → entr**a**-da 'entrance'

b. tard**a**r 'delay (v.)' → tard**a-**nza 'delay (n.)'

The stems *entrar* and *tardar* have ultimate stress, and the derived forms *entra-da* and *tarda-nza*, made by movement of the stem to the specifier position of affix, have penultimate stress, which is allowed in right-edge stress languages.

However, languages with right-edge stress do not allow complement-head order in compounds, NP, VP, PP, and CP as shown in the Romance column in (41b-f). For example, consider the Italian compounds in (46) where the head words are disyllabic (*capo* and *campo*).

(46) a. *capo stazi**o**ne* → **stazi**o**ne capo*

head station

'station master'

b. *campo s**a**nto* → **s**a**nto campo*

field holy

'cemetery'

Complement-head order on the right of the arrow is ruled out because stress is expected to fall on the fourth syllable from the end of the compound, which is a marked stress location for right-edge stress languages. As we saw in Section 4.2, we assume that stress should fall on the complement, on the assumption that it falls on the most deeply embedded element on the recursive side of a tree (cf. Cinque 1993).

The examples in (46) show that if the head is disyllabic, stress falls on the fourth syllable from the end of the compound. The longer the head, the earlier syllable the stress falls on, violating the constraint on stress location, right-edge. The heads listed in (41b-f), words, nouns, verbs, adpositions, and adverbial subordinators (e.g. *window, decide, into*, and *before*), are likely to be longer than monosyllables. Thus, we can correctly predict that languages with right-edge stress, such as Romance languages, do not have complement-head orders except for Stem-Affix, as shown in (41a-f).

However, what rules out complement-head orders in these languages if the head is monosyllabic? The heads listed in (41b-f), words, nouns, verbs, adpositions, and adverbial subordinators, can be monosyllabic (e.g. *desk, put, in*, and *when*). Monosyllabic heads add only one syllable to the resulting structure by phonologically attaching to the complement, as shown in (47).

(47) a. $[_{PP}$ *a* $[_{DP}$ *la sómbra*]]
at the shade

b. *$[_{PP}$ $[_{DP}$ *la sómbra*] *a*]

c. *$[_{PP}$ $[_{DP}$ *sómbra-la*] *a*]

We might assume the Final-Over-Final Constraint (FOFC) proposed by Holmberg (2000) and Biberauer et al. (2008). FOFC rules out a complement-head structure whose complement has the head-complement order.[17] (47b) is ruled out by FOFC because it is a PP with the complement-head order, which dominates a DP with head-complement order (D-NP/N). Thus, the base form (47a) cannot be changed into (47b) by complement-movement. Alternatively, we can claim that (47b) is ruled out because of its marked stress location. The (phrasal) compound *la sómbra-a* in (47b) has the stress on the antepenultimate syllable. This is marked in languages with right-edge stress, i.e. ultimate or penultimate stress, such as Spanish and Italian. In this view, (47b) may be constructed in syntax, but it is ruled out in the PF-component. The base (47a) cannot be changed into (47c) by successive complement-movement, which first moves the N *sómbra* to the specifier position of the D *la* and then moves the resulting DP *sómbra-la* to the specifier position of the adposition *a*. The harmonic complement-head structure in (47c) observes FOFC. However, its main stress falls on the fourth syllable from the end, which does not conform to the stress pattern of the languages with right-hand stress, including ultimate, penultimate, right-edge, and right-oriented stress. Thus, (47c) is not allowed in these languages. This is also the case with other heads. Then, languages with right-hand stress cannot have complement-head orders except for Stem-Affix even when the head is monosyllabic. They would violate FOFC or the constraint on the stress location in the language, as shown in (47b) and (47c).

4.5.4 Right-oriented stress

Next, let us consider why languages with right-oriented stress such as English and German allow complement-head orders in compounds, as shown in (41b). Right-oriented stress differs from right-edge stress in that only the former allows antepenultimate stress. Consider the example in (48), where the head noun *rack* is monosyllabic and its complement *towel* is disyllabic.

(48) a. rack (for) towel(s)

b. tówel rack

The base structure in (48a) is changed into a compound (48b) by complement-movement. The resulting compound (48b) has antepenultimate stress, which is allowed in English and other Germanic languages. Similarly, a monosyllabic complement (e.g. *rack*) can be moved to the specifier position of a disyllabic head (e.g. *railway*) without violating the right-oriented stress constraint, as shown in (49).

(49) a. railway (of) rack

b. ráck railway

Stress falls on the antepenultimate syllable in the compound (49b).

One might argue that there are compounds with more than three syllables, which violate the right-oriented stress constraint, as shown in (50).

(50) a. towel (in) kitchen

b. kítchen towel

In (50b), stress falls on the fourth syllable from the right end of the compound. However, English, German, and Dutch have weakening of vowels including weak vowels and vowel reduction. The example in (50b) has phonetic representations as shown in (51).

(51) a. kɪtʃən taʊəl

b. kɪtʃn taʊl

If the unstressed vowels are deleted, the compound has stress on the penultimate syllable, as in (51b). Thus, Germanic languages can observe right-oriented stress in compounds even if the resulting compound has stress on the (pre-)fourth syllable from its right end.[18] Weakening of vowels is not very common in Romance languages such as Italian and Spanish, which have no vowel reduction and no weak vowels in their phonological inventories.[19] [20]

Germanic languages other than English allow complement-head order in a VP, i.e. O-V, if it occurs with an auxiliary verb or it is in a subordinate clause, as shown in (52) (taken from Dryer 2005b).

(52) a. *Anna trink-t Wasser.* [V O]

Anna drink-3SG water

'Anna is drinking water.'

b. *Anna ha-t Wasser getrunken.* [Aux O V]

Anna have-3SG water drink.PST.PTCP

'Anna has drunk water.'

c. *Hans sag-t, dass Anna Wasser trink-t.* [C .. OV]

Hans say-3SG that Anna water drink-3SG

'Hans says that Anna is drinking water.'

Basically, object-verbs are the same as the compounds we have just seen above. Objects move to the specifier position of verbs to make derived compounds. The resulting compounds may have right-oriented stress. However, complement-movement is more likely to result in a marked stress location in O-V sequences than in Word (C)-Word (H) compounds because objects may well consist of more than one word. The stress position of O-V sequences can be too far to the left of the antepenultimate syllable to be rescued by vowel reduction. Thus, O-V order is not allowed in main clauses in German and Dutch, and in any clauses in English.

Two questions arise here. The first is why German and Dutch allow complement-movement in subordinate clauses. The second is why English is different from German and Dutch in disallowing complement-movement in subordinate clauses. A possible answer to the second question is to assume that stress in English falls on a syllable closer to the right end of a word than in German and Dutch. This is a plausible assumption if we consider the fact that English is influenced by language contact with French, which has right-edge stress. Thus, English has the V-O order both in main and subordinate clauses.

The first question is more difficult to answer. However, we can generalize the two cases in which German and Dutch have O-V order as follows: O-V if and only if VP is c-commanded by an overt head. Candidates for the overt head c-commanding VP are auxiliaries and complementizers. A possible explanation for the O-V order in German and Dutch is to assume that complement-movement needs to occur for compactization of constituents dominated by a higher overt head. We will not go into this issue in detail here, but see Tokizaki and Kuwana (in press).

4.5.5 Initial stress

Now let us consider languages with initial stress such as Uralic languages, including Hungarian and Finnish. These languages have complement-head order from word level to the clause level, except for adverbial subordinator-clause, as shown in (41). Word-initial stress does not conflict with complement-movement as shown in (53).

(53) a. $[_{XP}\ [_{X}\ \underline{\sigma}\ \sigma\ ...]\ [_{YP}\ \underline{\sigma}\ \sigma\ ...]]$

b. $[_{XP}\ [_{YP}\ \underline{\sigma}\ \sigma\ ...]\ [_{X}\ \underline{\sigma}\ \sigma\ ...]]$

In the resulting compound XP in (53b), the heaviest stress falls on the first syllable because it is the most deeply embedded element in XP. This stress location in compounds matches the unmarked word-stress location in these languages. Thus, we can explain why complement-movement occurs to make the complement-head orders Stem-Suffix, Word (C)-Word (H), Modifier-Noun, Object-Verb, and NP-P.[21]

The remaining question is why these languages have head-complement order only in subordinate clauses, i.e. adverbial subordinator-clause. A possible answer is that clauses are the only type of constituent that have an overt specifier (subject). Consider the structure of subordinate clauses in SOV languages shown in (54), where Sub stands for an adverbial subordinator.

(54) a. $[_{CP}\ \text{Sub}\ [_{IP}\ [_{Subj}\ \underline{\sigma}\ \sigma\ ...]\ \text{I}\ [_{VP}\ [_{Object}\ \underline{\sigma}\ \sigma\ ...]\ [_{V}\ \underline{\sigma}\ \sigma\ ...]]]]$

b. $[_{CP}\ [_{IP}\ [_{Subj}\ \underline{\sigma}\ \sigma\ ...]\ \text{I}\ [_{VP}\ [_{Object}\ \underline{\sigma}\ \sigma\ ...]\ [_{V}\ \underline{\sigma}\ \sigma\ ...]]\ \text{Sub}]$

As we have seen in (53b), (54a) is allowed in initial stress languages. If complement-movement applied to IP in (54a), the resulting structure would have a clause-adverbial subordinate order in (54b). However, the resulting compound CP in (54b) does not have initial stress: the heaviest stress falls on the object, which is the most deeply embedded element in CP, and not on the subject that is the initial element in the whole CP. CP in (54b) violates the constraints on word-stress location in initial-stress languages. Thus, initial-stress languages have head-complement order in adverbial subordinator-clause and complement-head orders in the other constituents.

4.5.6 No stress

Finally, let us consider languages with no stress such as Japanese and Korean. These languages do not have any problems in complement-movement in any of the constituents because there is no chance of stress-mismatch between words and derived compounds. Complement-movement applies to (55a) freely to make compounds in (55b).

(55) a. $[_{XP}\ [_{X}\ \sigma\ \sigma\ ...]\ [_{YP}\ \sigma\ \sigma\ ...]]$

b. $[_{XP}\ [_{YP}\ \sigma\ \sigma\ ...]\ [_{X}\ \sigma\ \sigma\ ...]]$

The presence of a subject in clauses is not a problem in making the clause-adverbial subordinator order, as shown in (56b).

(56) a. $[_{CP}\ \text{Sub}\ [_{IP}\ [_{Subj}\ \sigma\ \sigma\ ...]\ \text{I}\ [_{VP}\ [_{Object}\ \sigma\ \sigma\ ...]\ [_{V}\ \sigma\ \sigma\ ...]]]]$

b. $[_{CP}\ [_{IP}\ [_{Subj}\ \sigma\ \sigma\ ...]\ \text{I}\ [_{VP}\ [_{Object}\ \sigma\ \sigma\ ...]\ [_{V}\ \sigma\ \sigma\ ...]]\ \text{Sub}]$

Thus, we correctly predict that languages with no stress have complement-head order in any kind of constituent, as shown in (41).

5. Conclusion

We have argued that word-stress location matches the main stress position in constituents with complement-head order, which are left-branching and compound-like because of the short juncture between their elements. Complement moves to the specifier position of head in order to be interpreted with head at LF. This movement occurs in overt syntax only if the resulting constituent with complement-head order has the same stress position as a word. We have shown that fixed stress positions and weight-sensitive stress positions allow certain kinds of complements to move to the specifier position according to the number of syllables in heads and complements.

This stress-based theory of disharmonic word orders explains a fine correlation between stress position and head-complement orders in a number of languages. The next step is to show with statistical data that this theory correctly predicts word orders in more of the world's languages (see Kuwana and Tokizaki 2009). We will leave this for another article.

References

Biberauer, Theresa, Anders Holmberg and Ian Roberts (2008). Structure and linearization in disharmonic word orders. *Proceedings of the West Coast Conference on Formal Linguistics* 26, 96-104.

Chomsky, Noam (1995). *The minimalist program*. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo (1993). A null theory of phrase and compound stress. *Linguistic Inquiry* 24, 239-298.

Dryer, Matthew (1992). The Greenbergian word order correlations. *Language* 68. 81–138.

Dryer, Matthew (2005a). Prefixing vs. Suffixing in Inflectional Morphology. In Haspelmath, et al. (eds.), 110-113.

Dryer, Matthew (2005b). Order of Object and Verb. In Haspelmath, et al. (eds.), 338-341.

Dryer, Matthew (2005c). Order of Adposition and Noun Phrase. In Haspelmath, et al. (eds.), 346-349.

Dryer, Matthew (2005d). Order of Genitive and Noun. In Haspelmath, et al. (eds.), 350-353.

Dryer, Matthew (2005e). Order of adverbial subordinator and clause. In Haspelmath, et al. (eds.), 382-385.

Ernestus, Mirjam and Anneke Neijt (2008). Word length and the location of primary word stress in Dutch, German, and English. *Linguistics* 46, 507-540.

Goedemans, Rob and Harry van der Hulst (2005a). Fixed stress locations. In Haspelmath, et al. (eds.), 62-65.

Goedemans, Rob and Harry van der Hulst (2005b). Weight-sensitive stress. In Haspelmath, et al. (eds.), 66-69.

Han, Eunjoo (1994). Prosodic Structure in Compounds. Doctoral dissertation, Stanford University.

Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie (eds.) (2005). *The world atlas of language structures*, Oxford: Oxford University Press.

Holmberg, Anders (2000). Deriving OV order in Finnish. In Svenonius, Peter. (ed.) *The Derivation of VO and OV*, Amsterdam: Benjamins, 123-152.

Huang, C.-T. James (1982). Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT. [Published by Garland Publishers, New York, 1998]

Hyman, Larry M. (2008). Directional asymmetry in the morphology and phonology of words, with special reference to Bantu. *Linguistics* 46, 309-350.

Kayne, Richard S. (1994). *The antisymmetry of syntax*. MIT Press.

Krott, Andrea, Gary Libben, Gonia Jarema, Wolfgang Dressler, Robert Schreuder, and Harald Baaven (2004). Probability in the grammar of German and Dutch: Interfixation in triconsonstituent compounds. *Language and Speech* 47, 83-106.

Kuwana, Yasutomo and Hisao Tokizaki (2009). The co-variation of head-complement order and the position of word stress. Paper presented at the 8th international conference of the Association for Linguistic Typology. University of California, Berkeley July 23-26.

Mchombo, Sam (2004). *The syntax of Chichewa*. Cambridge: Cambridge University Press.

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

Scalise, Sergio (1984). *Generative morphonology*. Dordrecht: Foris Publications.

Selkirk, Elisabeth O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, Mass.: MIT Press.

Tokizaki, Hisao (1999). Prosodic phrasing and bare phrase structure. *Proceedings of the North East Linguistic Society* 29, vol. 1, 381-395.

Tokizaki, Hisao (2006). *Syntactic structure and silence: A minimalist theory of syntax-phonology interface*. Tokyo: Hitsuji syobo.

Tokizaki, Hisao (2008). Symmetry and asymmetry in the syntax-phonology interface. *Phonological Studies* 11, 123-130.

Tokizaki, Hisao and Yasutomo Kuwana (in press). Unattested word orders and left-branching structure. In Biberauer, Theresa and Ian Roberts (eds.). *Principles of linearization*. Berlin: Mouton de Gruyter.

Williams, Edwin (1981). On the notions ‘lexically related’ and ‘head of a word’. *Linguistic Inquiry* 12, 245-74.

---

* We would like to thank the organizers and participants of Theoretical Approaches to Disharmonic Word Orders, Newcastle University, May 30th-June 1st 2009. Special thanks go to Theresa Biberauer, Matthew Dryer, Anders Holmberg, Ian Roberts, and Michelle Sheehan. This work is supported by Sapporo University and Grant-in-Aid for Scientific Research (C18520388; A20242010), JSPS.

1 In addition to branchingness, we must also include a standard definition of head on the basis of projection: a head is an element that projects its morphosyntactic and semantic properties onto the category made by Merging it with its complement. This is necessary to exclude non-branching specifiers and derived categories by movement such as (i) and (ii).

(i) [He [likes coffee]]

(ii) [[*Marie* [[*chante*+T] [*souvent t*]]]
   Marie sings often

The subject *he* or *Marie* is not a head but a specifier even though it is non-branching. In (ii) *chante* is arguably branching but should be considered as the head of VP because it projects its syntactic and semantic properties to the VP.

2 It might also be possible for the complement YP to move to the specifier position of a higher functional head F as in (i).

(i) a. $[_{FP}\ F\ [_{XP}\ ZP\ X\ [_{YP}\ ..\ Y\ ..]]] \rightarrow$
  b. $[_{FP}\ [_{YP}\ ..\ Y\ ..]\ F\ [_{XP}\ ZP\ X\ t]]$

This option in (i) seems to be necessary when the specifier position of X is occupied by a trace of ZP, which has moved to a higher specifier position than FP as shown in (ii).

(ii) $..\ ZP\ ..\ [_{FP}\ [_{YP}\ ..\ Y\ ..]\ F\ [_{XP}\ t_{ZP}\ X\ t]]$

The FP in (ii) is also left-branching structure in that only YP and X are visible at the syntax-phonology interface as shown in (iii)

(iii) $..\ ZP\ ..\ [_{FP}\ [_{YP}\ ..\ Y\ ..]\ X]$

3 Selkirk (1984) proposes to represent the degree of juncture in terms of the number of "silent demibeats," which are shown by x. For example, the examples in (ia) and (ib) are distinguished by the presence of a 'rest' or 'pause' between *man* and *eating* in (ib) (p.324).

(i) a. This is a $[_{AP}$ $[_{N}$ man-] $[_{A}$ eating]] fish. $[_{N}$ man-] x $[_{A}$ eating]
  b. This is $[_{NP}$ a [man]] $[_{S}$ eating fish]. $[_{NP}$ a [man]] xx $[_{S}$ eating …]

4 Insertion of *n* is possible in (11b) in Kyungsan dialect (cf. Han 1994). For this dialectal difference, see Tokizaki (2008).

5 We thank Kimihiro Ohno and Yoshihito Dobashi for pointing out this alternative.

7 In addition to the types of adverbial subordinators in (16) and (17), Dryer (2005e) lists "clause-internal adverbial subordinators" (8 languages) and "more than one type of adverbial subordinator with none dominant" (87 languages). Dryer (2005e) does not give the diagnostics for affixhood. As suffixal adverbial subordinators, he gives examples in Kiowa

(Kiowa-Tanoan; central United States) and in Hunzib (Daghestanian; eastern Caucasus, Russia).

(i) a. Kiowa

*à-dę̀·k'ɔ́·-**àl*** *hɔ́n* *àn* *à-dę̀·hę́·m-ô*

1sg-lie-although neg hab 1sg-sleep-neg

'Although I lie down, I can't fall asleep.'

b. Hunzib

zaƛe *n-ex-**áyd**,* *xõx-ƛ'o* *ƛibu* *zuq'u-r*

wind nc5-strike-**before** tree-sup leaf be-pret

'Before the wind blew, there were leaves on the tree.'

Dryer gives examples of adverbial subordinators which are separate words that appear at the end of the clause, from Kombai (Trans-New Guinea; Papua, Indonesia) and Kolyma Yukaghir (isolate; Siberia, Russia).

(ii) a. Kombai

khe-khino *rerakharu* ***rofode***

his-legs swollen **because**

'because his legs are swollen'

b. Kolyma Yukaghir

ulum *gud-uj-l'ie-t* ***tit***

mad become-iter-ingr-ss.impf **although**

'although he was going mad'

The verb in (ia) and (ib) has a suffixal adverbial subordinator at the right and other affixes at the left. The verb in (iib) has suffixes at the right which are followed by an adverbial subordinator. We think that this positional difference with respect to other affixes could be a diagnostic for affixhood of adverbial subordinators.

[8] We argue that the facts we have seen in this section can be explained by asymmetric juncture straightforwardly. We do not mean to exclude the possibility that the facts are due to factors other than juncture. However, there have been no alternative explanations for the asymmetry facts presented here.

[9] See Tokizaki (1999, 2008) for invisible categories. Holmberg (2000) also argues that Complement-movement makes phonological left-branching structure. However, he does not consider the compound nature of left-branching structure.

[10] In addition to (31a) and (31b), there are two cases for weight-sensitive stress.

(i) a. (H **H**)]/(**H** H)]

b. (**L** L)]/(L **L**)]

In (ia) and (ib), there are two options for stress location. These are irrelevant to our discussion here.

[11] In fact, Spanish also has antepenult stress in a number of words, e.g. ***bolí**grafo* 'pen.' Note also that Spanish has no long vowels. Here we consider stressed syllables as 'heavy' because they may be pronounced with lengthening.

[12] Suffixes are closely connected to the stem to make a prosodic word, which is the domain of stress placement. In most cases, stress falls on the fixed stress location counting both affix and stem. Thus, stress may fall on the affix part in a word as in Spanish (i) and Italian (ii) (cf. Scalise 1984: 87, 99).

(i) a. or**a**r 'pray' → ora-ci**ó**n 'prayer'
b. barr**e**r 'clean (v.)' → barre-d**u**ra 'cleaning'

(ii) a. b**e**llo 'beautiful' → bell-**e**zza 'beauty'
b. aut**o**re 'author' → autor-izz**a**re 'authorize'

[13] Italian also has a variety of stress locations as shown in (i).

(i) a. citt**à** 'city'
b. mont**a**gna 'mountain'
c. t**a**vola 'table'
d. c**a**pitano (< *capitare*) 'captain'

In Italian, most words have penultimate stress (ib), and some words have ultimate stress (ia). There are also words with antepenult stress (ic) and stress on the fourth to the last syllable of the word (id). However, we do not find parallel examples to Spanish (36) because most suffixes in Italian are disyllabic.

[14] The list of weight-sensitive stress in (21) contains 26 languages with unpredictable stress location. We also need to consider how word orders are determined in tone languages without stress. We leave these matters for future research.

[15] In (44b), the base form of *sungi* is *sunga* 'keep.' The last vowel changes from *a* to *i* for agentive nouns when a prefix is attached to the word. See Mchombo (2004: 113).

[16] Penultimate stress is also seen in Welsh, in which the usual order is noun+adj, e.g. *llyfrau trwm* 'heavy books' (literally 'books heavy'), but may be adj+noun, e.g. *hen lyfrau* 'old books.' However, pronominal adjectives are limited in number (e.g. *hoff* 'favorite'; *prif* 'main'). We can consider these adjectives as lexical exceptions.

[17] Biberauer et al. (2008) formulates FOFC as (i).

(i) If α is a head-initial phrase and β is a phrase immediately dominating α, then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α, then β can be head-initial or head-final.

FOFC (i) rules out structures like that in (ii).

(ii) * $[_{\beta P}\ [_{\alpha P}\ \alpha\ \gamma P]\ \beta]$ where αP is the complement of β and γP is the complement of α.

[18] Some compounds may consist of words lacking vowel reduction as shown in (i).

(i) a. kɪtʃən kliːnə

b. kɪtʃn kliːnə

In this example, the main stress is on the first syllable of *kitchen*, which is the third from the last syllable (antepenult) in the whole compound, *kitchen cleaner,* in (ib). This is still allowed in the right-oriented (penult or antepenult) stress system in Germanic. Note also that as the compound word gets longer with more syllables, the weak syllables are pronounced even more weakly and are likely to be omitted. This is the factor that makes languages with a weight-sensitive stress system have a long complement in front of the head.

[19] Ernestus and Neijt (2008) point out that Germanic languages prefer word-initial stress. This preference seems to go well with lefthand stress in compounds.

[20] Other Romance languages such as French and Portuguese may have vowel reduction. We might expect that languages with vowel reduction have more flexible stress position and more complement-head orders than languages without vowel reduction. One difference we have found between French and other Romance languages is the order of the adjective phrase and noun in the equivalents of the English phrase, ‘a very old lady’: *une très vieille dame* vs. *una signora molto vecchia* (Italian); *una señora muy vieja* (Spanish); *uma senhora muito velha* (Portuguese). French has the AP-N order (complement-head), similar to Germanic languages, while other Romance languages have N-AP order (head-complement). We will leave the detailed examination of word order in these languages for future study.

[21] Hungarian and Finnish have variable word order. Dryer (2005) describes the order of object and verb in Hungarian as “no dominant order” and the order in Finnish as VO. Our theory predicts that complements can move to the specifier position of a head if the resulting structure observes the unmarked stress pattern. Languages with both complement-head and head-complement order give us the interesting problem of what the requirements are for complements to move to the specifier position of the head overtly rather than covertly. We will leave this problem for future research.