# Limited Consonant Clusters in OV languages

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

It has been claimed that the complexity of syllable structure is correlated to the order between verb and object in languages of the world: the syllable structure in OV languages is simpler than that in VO languages. However, our analysis of data in Maddieson (2005) and Dryer (2005) seems to show that a number of OV languages have (moderately) complex syllable structure. In spite of this result, we argue that the syllable structure in OV languages is simpler than has been reported, by considering the geographical gradience of coda variety, coda inventory, phonological simplification and particles attached to nouns, and complement-head orders other than OV/VO. We also discuss why OV languages have simple syllable structure: it is argued that juncture between constituents is stronger in left-branching structure (OV) than in right-branching structure (VO); strong juncture in left-branching structure makes words closely connected to each other; simple syllable structure such as CV fits nicely into the stronger juncture without making a consonant cluster.

## 1. Introduction

It has been pointed out that languages with object-verb order (OV) tend to have simple syllable structure (Lehmann 1973, Gil 1986, Plank 1998). This is the case in some OV languages such as Ijo, Yareba and Warao, whose syllable form is CV. However, examination of data in Haspelmath et al. (2005) (henceforth *WALS*) shows that a number of OV languages have (moderately) complex syllable structure.

In this paper, we argue that the syllable structure in OV languages is simpler than has been reported, by showing that consonant clusters are limited at word boundaries and between words in OV languages. We base our argument on only a small number of example languages but hope that these will be sufficient to demonstrate the viability of our research proposal. From a conceptual and theoretical point of view, we also discuss the reason why OV languages should have simple syllable structure.

In Section 2, we review the previous studies of the correlation between syllable complexity and word order. We also examine the correlation hypothesis using data from *WALS*. In Section 3, we argue that syllable structure in OV languages is simpler than it looks if we consider geographical gradation, simplification processes and limited coda inventory.

Section 4 discusses why OV languages have simple syllable structure; we argue that juncture between constituents is stronger in left-branching structure (OV) than in right-branching structure (VO). Section 5 concludes the discussion.

## 2. The correlation between syllable structure and OV order

### 2.1 Implicational universals

There have been a number of studies that try to show the correlation between phonology and syntax; Plank (1998) presents an overview of these. Here we concentrate on the relation between syllable structure and verb-object order. It has been pointed out that languages with object-verb order (OV) tend to have simple syllable structure (Lehmann 1973, Donegan and Stampe 1983, Gil 1986, Plank 1998). The Universals Archive lists two correlations, no. 196 and no. 207, with comments by Frans Plank as shown in (1).[1]

(1)
- a. OV languages tend to have simple syllable structure.
- b. IF basic order is OV, THEN syllable structure is simple (tending towards CV).
- c. Counterexamples: -
- d. Comments: Languages with flexive morphology (which tend to be OV) tend to have the ends of syllables closed, with consonant clusters occurring in this position as freely as in initial position (Lehmann 1973: 61).

This implicational relation is the case in some OV languages, such as Ijo (Niger-Congo), Yareba (Papua New Guinea) and Warao (Venezuela), whose syllable form is CV.

The Universals Archive also shows another correlation between word order and syllable structure, as shown in (2).

(2)
- a. VO languages tend to have complex syllable structure.
- b. IF basic order is VO, THEN syllable structure is complex (permitting initial and final consonant clusters).
- c. Counterexamples: Old Egyptian (Afro-Asiatic): VO,

only syllable types CV and CVC (F. Kammerzell, p.c.).

d. Comments: -

The two observations in (1) and (2) predict that there will be considerable differences between SOV and SVO languages with respect to syllable complexity.

Gil (1986) tests the correlation between OV/VO order and syllable structure with his 170 sample languages. He reports that the average number of segments in the syllable structure templates: SOV 4.04 < SVO 4.93. However, this result is not very convincing because the difference between SOV and SVO is less than 0.9 (0.89). Moreover, the number of sample languages is not large enough to claim (1) and (2) as universals across languages; it is necessary, therefore, to test the hypothesis with more data.

## 2.2 Testing the correlation with data from *WALS*

Let us try to show the correlation between OV/VO order using data from *WALS*, which lists 2,561 languages, including 359 languages with data on both syllable structure and OV/VO order.

Maddieson (2005) in *WALS* (chapter 12) divides languages into three categories according to their syllable structure: simple, moderately complex and complex, as shown in (3).

(3) a. Simple

CV Hawaiian and Mba (Adamawa-Ubangian, Niger-Congo; Democratic Republic of Congo)

(C)V Fijian, Igbo (Niger-Congo; Nigeria), and Yareba (Yareban; Papua New Guinea)

b. Moderately complex

CVC

$CC_2V$ $C_2$=liquids (r/l) or glides (w/j)

$CC_2VC$ $C_2$=w in Darai (Indo-Aryan; Nepal)

c. Complex

(C)(C)(C)V(C)(C)(C)(C) English

Categorizing syllable complexity into three groups is effective in showing typological differences between languages. However, we will argue that, as Plank (2009) points out, the categorization is not fine enough to enable correlations between syllable complexity and other features to be identified.[2]

Dryer (2005) in *WALS* (Chapter 83) distinguishes three types of languages with respect to the order of object and verb: OV, VO and no dominant order. The third type, languages in which neither OV nor VO is dominant, falls into two classes. The first class is of languages with flexible word order, where both orders are common and the choice is determined by extragrammatical factors, such as many Australian languages (e.g. Ngandi (Gunwinyguan; Northern Territory, Australia). In the second class are languages in which word order is primarily determined syntactically, but in which there are competing OV and VO constructions. This class includes German in which VO order is used in main clauses in which there is no auxiliary verb, while OV order is used in clauses with an auxiliary verb and in subordinate clauses introduced by a subordinator.

Combining Maddieson's and Dryer's classification of languages by syllable structure (#12) and word order (#83) in *WALS Online* (http://wals.info/index) gives us the results shown in Table 1 below. The average complexity of syllable structure in each word order is calculated with simple = 1, moderately complex = 2 and complex = 3. For example, the average syllable complexity of languages with OV order is 2.25 = (18 x 1 + 93 x 2 + 60 x 3) ÷ 171.

Table 1: Syllable comlexities and object-verb order: number of languages

| **Total 359 languages** | Order of Object and Verb | | |
|---|---|---|---|
| Syllable Structure | OV (171) | VO (165) | No domin. order (23) |
| Simple (44) | 18 | 23 | 3 |
| Moderately complex (198) | 93 | 95 | 10 |
| Complex (117) | 60 | 47 | 10 |
| Complexity Average | 2.25 | 2.15 | 2.30 |

These results do not seem to show the expected correlation between the object-verb order and the syllable structure, that we have seen in (1) (i.e. OV → simple syllable) and (2) (i.e. VO → complex syllable) above. Even worse, the 23 languages with simple syllable structure and VO orders outnumber the 18 languages with simple syllable structure and OV order. The 60 languages with complex syllable structure and OV order outnumber the 47 languages with complex syllable structure and VO order. These data are in fact the opposite of what we expected, given the previous studies we have seen above. It may be that the results can be improved by refining our quantitative approach.

First, Dryer (1992, 2009) argues that typological work should not be based on the number of languages, but on the number of genera. Genera are groups of languages whose similarity is such that their genetic relatedness is uncontroversial (Dryer 1992: 84). Dryer argues that counting genera rather than languages controls for the most severe genetic bias. Counting the numbers of genera instead of languages slightly improves the results, as shown in Table 2.

Table 2: Syllable complexities and object-verb order: number of genera

| **Total 272 genera** | Order of Object and Verb | | |
|---|---|---|---|
| Syllable Structure | OV (132) | VO (117) | No domin. order (23) |
| Simple (36) | 17 | 16 | 3 |
| Moderately complex (140) | 67 | 63 | 10 |
| Complex (96) | 48 | 38 | 10 |
| Complexity Average | 2.23 | 2.19 | 2.30 |

The 17 genera with simple syllable structure and OV order outnumber the 16 genera with simple syllable structure and VO order. However, the 48 genera with complex syllable structure and OV order still outnumber the 38 genera with complex syllable structure and VO order.

Second, Dryer (1992, 2009) argues that genera should also be divided into six macro areas. He emphasizes that it is dangerous to use data from raw totals of languages without examining their distribution over areas. Dividing genera into macro areas gives Table 3.

Table 3: Syllable complexities and object-verb order: number of genera in six macro areas

| **a. Africa (58)** | Order of Object and Verb | | |
|---|---|---|---|
| Syllable Structure | OV (16) | VO (37) | No domin. order (5) |
| Simple (10) | 2 | 7 | 1 |
| Moderately complex (37) | 12 | 22 | 3 |
| Complex (11) | 2 | 8 | 1 |

| **b. Eurasia (45)** | Order of Object and Verb | | |
|---|---|---|---|
| Syllable Structure | OV (33) | VO (10) | No domin. order (2) |

| | | | |
|---|---|---|---|
| Simple (0) | | | |
| Moderately complex (15) | 13 | 2 | |
| Complex (30) | 20 | 8 | 2 |

| **c. South East Asia (39)** | Order of Object and Verb | | |
|---|---|---|---|
| Syllable Structure | OV (9) | VO (30) | No domin. order (0) |
| Simple (3) | | 3 | |
| Moderately complex (25) | 6 | 19 | |
| Complex (11) | 3 | 8 | |

| **d. Australia (45)** | Order of Object and Verb | | |
|---|---|---|---|
| Syllable Structure | OV (33) | VO (8) | No domin. order (4) |
| Simple (8) | 7 | 1 | |
| Moderately complex (25) | 17 | 6 | 2 |
| Complex (12) | 9 | 1 | 2 |

| **e. North America (46)** | Order of Object and Verb | | |
|---|---|---|---|
| Syllable Structure | OV (16) | VO (22) | No domin. order (8) |
| Simple (2) | | 2 | |
| Moderately complex (21) | 10 | 8 | 3 |
| Complex (23) | 6 | 12 | 5 |

| **f. South America (40)** | Order of Object and Verb | | |
|---|---|---|---|
| Syllable Structure | OV (25) | VO (11) | No domin. order (4) |
| Simple (13) | 8 | 3 | 2 |
| Moderately complex (19) | 10 | 7 | 2 |
| Complex (8) | 7 | 1 | |

Table 3 shows that there are more OV genera than VO genera with simple syllable structure in (d) Australia (7:1) and (f) South America (8:3). However, these areas also have more OV genera than VO genera with complex syllable structure, i.e. (d) Australia (9:1) and (f) South America (7:1). In the other areas, (a) Africa, (b) Eurasia, (c) South East Asia and (e) North America, the number of OV genera with simple syllable structure is not more than that of VO genera with simple syllable structure. These results show that the data in *WALS* do not give straightforward support for the hypothesis that OV languages have simple syllable structure.

However, in the next section we argue that OV languages do have simple syllable structure if we consider the geographical gradation of the variety of word-final consonants, the fine classification of syllable complexity and head-complement orders, the coda inventory and the simplification of syllable structure within words and between words.

## 3. Reconsidering syllable structure in OV languages

### 3.1 Geographical gradation of coda inventory

First, as we saw in Section 2, Maddieson (2005) in *WALS* defines CV as "simple" syllable structure, (C)CVC [Onset CC limited] as "moderately complex" and others such as CCVC [CC free], CCCV... and ...VCC as "complex." However, this three-way distinction of syllable structure is not fine enough to enable us to see possible correlations with other features such as word orders. For example, syllable complexity should be defined on the basis of the number and variety of coda consonants. Hashimoto (1978) argues that both coda and tone are simpler in north Asia than in south Asia, as shown in Table 4.

Table 4: Number of tones and codas in Asian languages (cf. Hashimoto 1978)

| | Manchu | Gansu | Beijing | Nanshang | Guanzhou | Thai |
|---|---|---|---|---|---|---|
| # tones | 0 | 3 | 4 | 6 | 8 (9) | 8 |
| coda | n/ŋ | n/ŋ | n/ŋ | n/ŋ/t/k | m/n/ŋ/p/t/k | m/n/ŋ/p/t/k |
| direction | North | ← | | | → | South |

Southern languages have a wider variety of coda consonants than northern languages. Thai, a VO language, has the most complex syllable, and Manchu, an OV language, has the simplest syllable among these languages. However, both of them are classified as "moderately complex" in *WALS*. Japanese, another OV language, is classified as having a "moderately complex" syllable. However, its syllable is (C)V(n), which is quite close to the "simple" syllable structure (C)V.

Interestingly, this geographical gradation of the coda inventory correlates with the variety of head-complement orders in these languages. The northern language Manchu has consistent head-final order in words and constituents of a variety of sizes: Stem-Suffix, Genitive-Noun, Adjective-Noun, Noun Phrase-Postposition, Object-Verb, Clause-Adverbial Subordinator (complement underscored). We define head as a non-branching category and complement as a (potentially) branching category. Head-complement order, shown with (–), increases as we move south, and as the coda inventory and number of tones increase, as shown in Table 5.[3]

Table 5: Number of tones, coda variety and complement-head orders (+) (Stem-Suffix, Genitive-Noun, Adjective-Noun, Noun Phrase-Postposition, Object-Verb, Clause-Adverbial Subordinator)

| Language | #tones | coda | St-Suf | G-N | A-N | N-P | O-V | Cl-Sb |
|---|---|---|---|---|---|---|---|---|
| Manchu | 0 | n/ŋ | + | | + | + | + | + |
| Gansu | 3 | n/ŋ | | | | | | |
| Beijing | 4 | n/ŋ | | + | + | – | – | |
| Nanshang | 6 | n/ŋ/t/k | | | | | | |
| Guanzhou | 8 (9) | m/n/ŋ/p/t/k | | + | +– | – | – | +– |
| Thai | 8 | m/n/ŋ/p/t/k | | – | – | – | – | – |

This table shows that the distinction between OV/VO languages is not sufficient to explain the correlation between head-complement orders and syllable complexity. We will try to show such fine correlation in the next section.[4]

3.2 Number of segments and degree of head-complement order

In order to decide the degree of head-complement/ complement-head order of a language, we checked the languages reported in Gil (1986) with the six head-complement orders in Table 5, which correspond to the features in *WALS* shown in (4).

(4)
a. Prefixing vs. suffixing in inflectional morphology (#26)
b. Order of object and verb (#83)
c. Order of adposition and noun phrases (#85)
d. Order of genitive and noun (#86)
e. Order of adjective and noun (#87)
f. Order of adverbial subordinator and clause (#93)

We assign 1 to each feature when it is a head-complement order (i.e. Prefix-Stem, Verb-Object, Preposition-Noun Phrase, Noun-Genitive, Noun-Adjective, Adverbial Subordinator-Clause) and –1 when it is a complement-head order (i.e. Stem-Suffix, Object-Verb, Noun Phrase-Postposition, Genitive-Noun, Adjective-Noun, Clause-Adverbial Subordinator). Then, the total score of a consistent head-initial language such as Bantoid and Mixtecan is 6; that of a consistent head-final language such as Turkic and Semitic is –6; that of a mixed language such as Baltic or Athapascan is 0 (–1x3 plus 1x3).

For syllable complexity, we used the number of segments in 170 languages listed in Gil (1986), which is based on the *Stanford Phonology Archive* and the *UCLA Phonological Segment Inventory Database*. Following Dryer (1992, 2005), we counted the number of genera rather than languages.

We grouped the genera according to the number of segments in a syllable, and calculated the average value of the head-complement orders. The result is shown in Table 6.

Table 6: The score of head-complement values sorted by number of segments in a syllable

| Number of Segments | Average of HC score | Number of genera |
|---|---|---|
| 2 | *1.33* | 6 |
| 3 | -1.45 | 29 |
| 4 | -0.84 | 57 |
| 5 | *0.92* | 39 |
| 6 | -0.07 | 14 |
| 7 | 1.20 | 5 |
| 8 | 2.50 | 2 |
| 9 | *2.33* | 3 |

Although the data are insufficient in some cases, Table 6 shows a tendency: as the number of segments increases, the value of head-complement orders increases. Except for the languages with two, five and nine segments in a syllable, which have the head-complement scores of 1.33, 0.92 and 2.33 respectively (italicized), the HC score gradually increases from -1.45 to 2.50. This result at least shows that we can expect a fine correlation between syllable complexity and head-complement orders including OV/VO order.

### 3.3 Limited coda inventory in OV languages

The coda inventory is more limited in OV languages than in VO languages. A list of OV languages with possible coda consonants is shown in (5).[5]

(5)
a. Japanese: n
b. Kanuri (Saharan): n, m, l, ɹ
c. Avar (Avar-Andic-Tsezic): n, m, w, j
d. Tamil (Southern Dravian): n, ɲ, ŋ, m, l, ɭ, ɾ, r, j
e. Moghol (Mongolic): n, m, r, d
f. Rutul (Lezgic): d, l, s, x
g. Lezgian (Lezgic): m, b, k, l, z, r
h. Chukchi (Chukotko-Kamchatkan): n, l, w, j, t, k
i. Korean: n, ŋ, m, l, p, t, k
j. Kurdish (Central) (Iranian): w, n, m, r, k, t, v ,š, ž

This list shows that there is a general order of consonants appearing in the coda position in OV languages. VanDam (2004) argues that languages tend to simultaneously prefer a manner hierarchy (nasal > liquid > obstruent > glide) and a place hierarchy (alveolar > velar > retroflex, tap). This tendency seems to be generally true in languages in (5). We could argue that, at most, OV languages tend to have nasals, liquids, and some voiceless obstruents as a coda. In this sense, syllable structure in OV languages is simpler than in VO languages, which may have a full variety of obstruents and glides.[6]

Note that Kurdish (Central) in (j) has a rich variety of coda consonants. However, this language has head-complement orders in other constituents than OV order: Stem-Suffix, Noun-Genitive, Noun-Adjective, Preposition-Noun Phrase, Adverbial Subordinator-Clause (complements underscored). Thus, Kurdish (Central) is more of a head-complement language than a complement-head language, even though it has OV order: its value of head-complement order is 2 (=(–1)x2+1x4). This example again shows that we need to check word orders other than OV/VO in order to see the correlation between word orders and syllable structure, as we saw in Section 3.2.

A question to ask is whether the coda inventory in VO languages is not as limited as in OV languages. As we will argue, our analysis predicts that syllable structure in OV languages is simple while that in VO languages may be either complex or simple. In fact, we find a number of VO languages or genera with no coda, such as Igbo (Igboid: Niger-Congo).

However, these languages/genera are not counterexamples to our analysis. We will return to this point in Section 4.

### 3.4 Limited consonant clusters within words in OV languages

Now let us consider the consonant clusters between words in languages. OV languages of "(moderately) complex" syllable structure may have phonological changes such as epenthesis and deletion, which simplify syllable structure. We propose (6) as working hypotheses.

(6) a. Consonant clusters are reduced in OV (head-final) languages.
    b. Consonant clusters are not reduced in VO (head-initial) languages.

For example, consonant clusters may be avoided by epenthesis of vowels, deletion of consonants and coalescence, as schematized in (7).

(7) Consonant clusters can be reduced by
    a. Epenthesis (CC → CVC)
    b. Deletion (CC → C)
    c. Coalescense (CC → C)

These phonological changes are found in such languages as Hindi and Basque, which are classified as "complex" syllable structure in *WALS*, but should be called "moderately complex."

First, let us look at the case of epenthesis, which can be found in a number of OV languages, as shown in (8) (cf. Lee and Ramsey (2000) for Korean).[7]

(8) a. Nambiqara: *w'aklsú* → *w'akəlisú* 'alligator'
    b. Persian: *drožki* (Russian) → *doroške* 'droshky'
    c. Basque: *libru* (Latin) → *liburu* 'book'
    d. Kannada: *magal* (Old) → *magalu* (New) 'daughter'
    e. Japanese: *drink* (English) → *dorinku*
    f. Korean: *text* (English) → *teyksuthu*

In these examples, consonant clusters are reduced by epenthesis of a vowel.

Second, we have cases of deletion within words, as shown in the Basque examples in (9) (Hualde and de Urbina 2003: 63).

(9) a. *gloria* (Latin) → *loria* ‘glory’
b. *ecclesia* (Latin) → *eliza* ‘church’

A consonant is deleted in the word-initial position in (9a) and in the medial position in (9b).

Third, Korean used to have consonant clusters at the onset position in the era of Middle Korean. These clusters CC changed into reinforced consonants in Modern Korean.[8] The examples in (10) show the process of coalescence.

(10) a. *stʌr* → *ttal* “daughter”
b. *pskur* → *skur* → *kkul* “honey”

Here *tt* and *kk* show reinforced consonants (cf. Lee 1975: 152).

On the other hand, VO languages seem to have few examples of deletion of consonant clusters. Although it is difficult to show that this is universally the case, there are examples showing that VO languages may delete vowels to make consonant clusters. For example, consider English names of Japanese companies in (11), where vowels are deleted to make consonant clusters or codas.

(11) a. Matsuda → Mazda
b. Yasukawa → Yaskawa
c. Noritsu → Noritz

These examples show that VO languages such as English do not need to simplify consonant clusters. Note that we are not claiming that every VO language has the means of making consonant clusters and codas illustrated here. As we will discuss in Section 4, our analysis predicts that the syllable structure in OV languages should be simple while that in VO languages can be either complex or simple.

3.5 Limited consonant clusters between words in OV languages

Finally, we would like to point out that consonant clusters between words (as well as those within words) are also limited in OV languages. For example, Korean, which has a number of nouns ending in a coda consonant, in fact has particles attached to them to show their cases. Korean has two forms of particles, which are phonologically conditioned, as shown in (12).

(12) a. nominative: *-i/ka*
b. accusative: *-ul/lul*
c. instrumental: *-ulo/lo*
d. comitative: *-kwa/wa*
e. vocative: *-a/ya*
f. topic: *-un/nun*

In (12), the first form of each pair attaches to a word ending with a consonant and the second form to a word ending with a vowel.[9] Thus, particles and the words they attach to do not make a consonant cluster even if the words end in a consonant.

Note also that these particles end in a vowel *i/a/o* or a consonant *l/n*. Thus, constituents consisting of a noun phrase and a particle end in a vowel or *l/n*.

These features of Korean morpho-phonology make Korean more like a syllable-timed or moraic language with the form CVCV.... Then, Korean is not a real counterexample to the universal tendency for head-final languages to have simple syllable structure.

Similar examples can be found in Moghol (Mongolic), which has eight types of case suffixes (Weiers 2003: 254).

(13) a. genitive: *-i/-ai*
b. accusative: *-i/-'i*
c. dative: *-du/-do/-tu* [cf. *du* (preposition)]
d. ablative: *-sa/-sah, -asa/-asah* [cf. *sah* (preposition)]
e. instrumental: *-ar*
f. comitative: *-la/-lah*
g. vocative: *-å*

In the ablative (13d), consonant stems normally require the presence of an extra vowel segment, which avoids making a consonant cluster with the preceding stem. These case suffixes end in vowels or *h/r*; constituents consisting of a noun phrase and a particle also end in a vowel or *h/r*.

Nivkh also has epenthesis in the case of third person singular pronouns (Shiraishi 2006: 39, 41) as shown in (14) and (15).

(14) a. *ŋ-ɨmɨk* 'my mother'
my mother

b. *cʰ-ɨtɨk* ‘your father’
your father

(15) a. *ŋi̱- ˈzŋaj* ‘my picture’
my picture

b. *cʰi̱- ˈzŋaj* ‘your picture’
your picture

Pronominal clitics attach to a vowel-initial host in (14) and to a consonant-initial host in (15) where the vowel *i* is inserted.

In this section we argued that OV languages do have simple syllable structure if we consider the geographical gradation of the variety of word-final consonants, the fine classification of syllable complexity and head-complement orders, the coda inventory and the simplification of syllable structure within words and between words. We used examples from a range of languages to illustrate these points.

**4. Why do OV languages have simple syllable structure?**

We have argued that OV languages tend to have simple syllable structure with fewer consonant clusters between words and within words. In this section, we consider why word orders correlate with syllable structure. Tokizaki (2008) argues that left-branching structure has stronger juncture between its constituents than right-branching structure. The juncture between B and C in left-branching (16a) is stronger than the juncture between A and B in right-branching (16b).

(16) a. [[A B] C]
b. [A [B C]]

In this sense, the juncture is asymmetrical between left-branching and right-branching structure. Tokizaki (2008) shows phonological and morpho-syntactic evidence for this junctural asymmetry. Let us review some of the arguments about Japanese and Korean presented there and discuss some new data from Dutch and German. First, consider *Rendaku* (sequential voicing) in Japanese, which applies to the first consonant in a word preceded by another word ending with a vowel. For example, the first consonant in the second word in (17a) and (17b) is voiced when it is a part of a compound.

(17) a. *nise tanuki* → *nise danuki*
mock badger ‘mock-badger’
b. *tanuki shiru* → *tanuki jiru*
badger soup ‘badger-soup’

The voicing rule also applies to three-word compounds if they have left-branching structure as in (18a), but it is blocked if they have right-branching structure as in (18b) (Otsu (1980)).

(18) a. [[*nise tanuki*] *shiru*] → *nise danuki jiru*
mock badger soup ‘mock-badger soup’
b. [*nise* [*tanuki shiru*]] → *nise tanuki jiru*
mock badger soup ‘mock badger-soup’

Let us assume that *Rendaku* is the process that assimilates a word-initial consonant to the preceding vowel with respect to the feature [+voice]. Then *Rendaku* is blocked when there is a left bracket between a word-final vowel and a word-initial consonant as in (18b). Thus Japanese *Rendaku* is a case of left/right-branching asymmetry with respect to blocking phonological change.

Another case of left/right-branching asymmetry is *n*-Insertion in Korean. In Standard Korean, *n* is inserted before a stem beginning in *i* or *y* when it is preceded by another stem or prefix which ends in a consonant. For example, *sæk* ‘color’ and *yuli* ‘glass’ may make *sæŋ nyuli* ‘colored glass’. This rule can apply in compounds with left-branching structure while it cannot in compounds with right-branching structure (Han (1994)).

(19) a. [[*on chən*] *yok*] → *on chən nyok*
hot spring bathe
‘bathing in a hot spring’
b. [[*mæŋ caŋ*] *yəm*] → *mæŋ jaŋ nyəm*
cecum bowel fire
‘appendicitis’
(20) a. [*kyəŋ* [*yaŋ sik*]] → *kyəŋ yaŋ sik*/**kyəŋ nyaŋ sik*
light Western food
‘a light Western meal’
b. [*myəŋ* [*yən ki*]] → *myəŋ yən gi*/**myəŋ nyən gi*
fame play skill
‘excellent performance’

A left bracket in a compound blocks *n*-Insertion as in (19), and a right bracket does not, as in (20).

The left/right-branching asymmetry is also seen in languages other than Japanese and Korean. According to Krott et al. (2004), interfixation in Dutch three-word compounds shows the left/right-branching asymmetry. In Dutch, the occurrence of interfix including *-s-* in tri-constituent compounds matches the major constituent boundary better in right-branching compounds than in left-branching compounds. In (21) and (22), the numbers of compounds with *-s-* and all interfixes are shown in parentheses after the examples.

(21) a. [*arbeid-s*-[*vraag stuk*]] (-*s*- 38; all 60)
employment+question-issue
b. [*hoofd* [*verkeer-s-weg*]] (-*s*- 3; all 11)
main+traffic-road
(22) a. [[*grond wet*]-*s-aartikel*] (-*s*- 25; all 39)
ground-law+article, constitution
b. [[*scheep-s-bouw*] *maatschappij*] (-*s*- 13; all 50)
ship-building+company

The ratio of the unmarked interfix position (21a) and (22a) to the marked interfix position (21b) and (22b) is higher in right-branching (21) (-*s*- 38÷3=12.7; all 60÷11=5.5) than in left-branching (22) (-*s*- 25÷13=1.9; all 39÷50=0.8). That is, interfixes occur at the constituent break more often in right-branching compounds than in left-branching compounds. This result is expected if we assume that the juncture between constituents in right-branching is weaker than that in left-branching structure.

Moreover, Wagner (2005) shows that there is a phrasing asymmetry between OV and VO orders: OV is pronounced as a prosodic phrase while VO is pronounced as two prosodic phrases. In (23), parentheses show prosodic phrases.

(23) a. (*Sie hát*) (*einen Tángo getanzt*)
she has a-Acc tango danced
'She has danced a tango.'
b. (*Sie tánzte*) (*einen Tángo*)
she danced a-Acc tango
'She danced a tango.'

The OV in (23a) [[*einen Tángo*] *getanzt*] is left-branching and is included in a prosodic phrase. The VO in (23b) [*tánzte* [*einen Tángo*]] is right-branching and is divided into different prosodic phrases.

These arguments support the idea of left/right-branching asymmetry. Now let us see how the asymmetry sheds light on the relation between word orders and syllable structure in languages. Let us consider how simple syllable structure allows an object to move to the left of the verb to make left-branching structure. For example, a verb phrase tends to have right-branching structure in a head-initial language (24a), and left-branching structure in a head-final language (24b).

(24) a. $[_{VP}\ V\ [_{NP}\ ..\ N\ ..]]$
b. $[_{VP}\ [_{NP}\ ..\ N\ ..]\ V]$

However, if we assume the left/right-branching asymmetry discussed above, head-final languages in fact have compound-like verb ‘phrases’.

(25) $[_{V}\ \llbracket ..\ N\ .. \rrbracket\ V]$

The object and the verb in (25), separated only by a weak bracket (represented by ⟧), are more closely connected to each other than the object and the verb in (24a), which are separated by a strong boundary. Simple syllable structure such as CV fits nicely into the stronger juncture in (25) without making a consonant cluster, as in (26).

(26) $[_{V}\ \llbracket ..\ C\underline{V \rrbracket\ C}V]$

Then VO languages are allowed to have complex syllable structure because strong boundaries separate the coda of the verb and the onset of the object as shown in (27).

(27) $[_{VP}\ ..\ CCCV\underline{CC}\ [_{NP}\ \underline{CCC}VCC\ ..\ ]]$

Thus, left/right-branching asymmetry gives us an interesting way to explain a correlation between syntax and phonology.[10]

## 5. Conclusion

We have seen that data in *WALS* do not show a clear correspondence between OV languages and simple syllable structure. However, we have argued that this is partly due to the crude distinction between syllable complexity in *WALS*. We have pointed out that we should take into account the geographical gradience of coda variety, coda inventory, phonological simplification and particles attached to nouns, and complement-head orders other than OV/VO. These points limit consonant clusters within words and between words in OV languages. Thus, the correlation between OV order and simple syllable structure is more realistic than it seems. This correlation is predicted by the notion that left-branching structure has stronger juncture than right-branching structure.

Needless to say, we need to investigate the points just mentioned more carefully and thoroughly. We hope that this research is a step toward a typology of syllable complexity and its relation to other components of grammar.


## Acknowledgments

We would like to thank Theo Vennemann for invaluable comments and suggestions. We are also grateful to Bingfu Lu for his comments on Chinese dialects. This work is supported by Grant-in-Aid for Scientific Research (A20242010, C18520388) and Sapporo University.


## References


Donegan, Patricia J. and David Stampe 1983 Rhythm and the holistic organization of language structure. In: John F. Richardson, Mitchell Marks and Amy Chukerman (eds.), *Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax*, Chicago: Chicago Linguistics Society, 337–353.

Dryer, Matthew S. 1992 The Greenbergian word order correlations. *Language* 68: 81138.

Dryer, Matthew S. 2005 Order of object and verb. In: Haspelmath et al. (eds.), 338–339.

Dryer, Matthew S. 2009 Problems testing typological correlations with the online WALS. *Linguistic Typology* 13: 121–135.

Gil, David 1986 A prosodic typology of language. *Folia Linguistica* 20: 165–231.

Han, Eunjoo 1994 Prosodic structure in compounds. Doctoral dissertation, Stanford University.

Hashimoto, Mantaro 1978 *Gengo ruikei chiri-ron* (Typological and geographical linguistics). Tokyo: Kobundo. Also in *Hashimoto Mantaro Chosaku-shu* vol. 1, Tokyo: Uchiyama-shoten, 29–190.

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

Hualde, José Ignacio and Jon Ortiz de Urbina 2003 *A grammar of Basque*. Berlin: Mouton de Gruyter.

Kamei, Takashi, Rokuro Kohno and Eiichi Chino (eds.) 1988-2001 *Gengogaku Daijiten* (The Dictionary of Linguistics). Tokyo: Sanseido.

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

Lee, Iksop and S. Robert Ramsey 2000 *The Korean Language*. Albany: State University of New York Press. The Japanese edition is published as *Kankokugo Gaisetsu*. Tokyo: Taishukan, 2004.

Lee, Ki-Moon 1975 *Kankokugo-no Rekishi* (History of Korean Language). Supervised by Shichiro Murayama and translated by Yukio Fujimoto. Tokyo: Taishukan.

Lehmann, Winfred P. 1973 A structural principle of language and its implications. *Language* 49: 47–66.

Maddieson, Ian 2005 Syllable structure. In: Haspelmath et al. (eds.), 54–55.

Maddieson, Ian 2010 Correlating syllable complexity with other measures of phonological complexity. *On-in Kenkyu* (Phonological Studies) 13, 105–116.

Mehler, Jacques, Núria Sebastián-Gallés and Marina Nespor 2004

Biological foundations of language acquisition: evidence from bilingalism. In: Michael S. Gazzaniga, Emilio Bizzi and Ira B. Black (eds.) *The cognitive neurosciences III*, Cambridge, MA: Bradford, MIT Press, 825–836.

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.

Plank, Frans 1998 The co-variation of phonology with morphology and syntax: A hopeful history. *Linguistic Typology* 2: 195–230.

Plank, Frans 2009 *WALS* values evaluated. *Linguistic Typology* 13, 41–75.

Price, David P. 1976 Southern Nambiquara phonology. *International Journal of Anthropological Linguistics* 42, 338–348.

Rastorgueva, Vera. S. 1964 *A short sketch of the grammar of Persian,* (translated by Steven P. Hill ; edited by Herbert H. Paper.) Bloomington: Indiana University.

Shiraishi, Hidetoshi 2006 *Topics in Nivkh Phonology*. Groningen Dissertations in Linguistics 61. University of Groningen.

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

VanDam, Mark 2004 Word final coda typology. *Journal of Universal Language* 5: 119–148.

Wagner, Michael 2005 Asymmetries in prosodic domain formation. *MIT Working Papers in Linguistics* 49, 329–367.

Weiers, Michael 2003 Moghol. In: Juha Janhunen (ed.) *The Mongolic languages*, London: Routledge, 248–264.

---

1 http://typo.uni-konstanz.de/archive/intro/index.php (Accessed on August 25, 2009)

2 Maddieson (2009) admits the crudity of this three-way distinction of syllable complexity, and proposes a refinement of syllable typology by scoring the complexity of onset, nucleus and coda, as shown in (i)-(iii).

(i) Contribution of Onset:

0 = Maximal onset is single C
1 = Maximal onset is C + liquid, glide (or nasal)
2 = Maximal onset is CC where C2 may be an obstruent
3 = Maximal onset is CCC or longer

(ii) Contribution of Nucleus:
1 = Nucleus is only simple (monomoraic) V
2 = Nucleus may be long vowel or diphthong

(iii) Contribution of Coda:
0 = No codas allowed
1 = Maximal coda is single C
2 = Maximal coda is CC
3 = Maximal coda is CCC

The refined syllable typology has eight steps on a scale (1-8). Maddieson claims that distribution of languages across categories is approximately normal with N = 605 languages. According to this typology, 'simple' languages (maximal syllable CV) = 1, Japanese = 3 (maximal syllable CjVVC) and Dutch/English = 8. We expect to be able to see the correlation between syllable structure and word orders if we use this typological data on syllable complexity; however, these data are not available at present.

[3] We consider interesting the geographical gradation of coda inventory and head-complement orders because it might show us a relation between linguistics and anthropology. However, this topic is far beyond the scope of this paper and we leave the matter open.

[4] It is an open question whether a similar geographical gradation of coda inventory can be found in languages other than Chinese dialects. Unfortunately, we do not have sufficient data about coda inventory in the world's languages to identify such cases. We leave the problem for future research.

[5] The coda data for Kanuri, Korean, Tamil and Chukuchi in list (5) are from VanDam (2004). We also checked the other languages by analyzing the data in Kamei et al. (1988-2001).

[6] One might argue that our selection of languages in this section and the next is arbitrary. We admit that we have not checked all languages in a principled manner. However, the point of our argument is to show that there are at least a number of OV languages whose syllable structure is simpler than previously reported, and that this is an area worthy of future investigation.

[7] We selected the languages in (8) from the OV languages with a description of syllable simplification in Kamei et al. (1988-2001). The examples in (8) are taken from: Price (1976) (8a), Rastorgueva (1964) (8b), Hualde and Urbina (2003) (8c), Kamei et al. (1988-2001) (8d), and Lee and Ramsey (2000) (8f). Price (1976: 346) reports that in Nambiquara "a nondistinctive vowel occurs between all combinations of consonants that involves a change in the oral place of articulation."

[8] We would like to thank John Whitman for discussion on Korean phonology.

[9] We need to consider the reason why *-kwa* instead of *-wa* is used after a word ending with a consonant to make a consonant cluster. Another remaining problem is why the genitive case marker *-uy* does not have another form with an onset consonant.

[10] Mehler et al. (2004) report experimental work showing the correlation between head-complement order and rhythm, i.e. head-complement = stress-timed vs. complement-head = mora-timed. Although it is based on data from only fourteen languages, the result seems to apply to other languages as well.