Phonological Externalization volume 9 (2024), 125-151.
Hisao Tokizaki (ed.), Sapporo University.

*Heavy-Head-Light: Generalizing Greenberg's Universal #25

Hisao Tokizaki and Yasutomo Kuwana
Sapporo University and Asahikawa Medical University

Abstract. Greenberg's (1963) Universal #25 states that if the pronominal object follows the verb, so does the nominal object. In this paper, we argue that this universal is a subcase of a more general universal *Heavy-Head-Light, which states that no language has heavy complements in pre-head position and light complements in post-head position. We examine the word order data of Dryer and Haspelmath (2013) and show that the apparent counterexamples to this generalization are not real ones. We argue that this generalization holds in the world's languages because head-final (left-branching) structure has tighter juncture between head and complement than head-initial (right-branching) structure.*[*]

Keywords: word order, length, heaviness, juncture, branching

1. *DP-V-Prn: Greenberg's Universal #25

Greenberg (1963) states that if the pronominal object (Prn) follows the verb, so does the nominal object. This implicational universal #25 holds in consistent Verb-Object languages, for example in English (1), where the verb consistently precedes its object irrespective of its size (pronoun or full NP). We represent a nominal phrase as DP (determiner phrase) (cf. Abney 1987).[1]

(1) a. John loves her. (V-Prn)
b. John loves that girl. (V-DP)

Here, the pronominal object follows the verb in (1a), as does the nominal object in (1b). This universal vacuously holds in consistent Object-Verb languages, for example in Japanese (2), where both pronominal object and nominal object precede the verb.

(2) a. *Taro-wa sore-o tabeta*. (Prn-V)
Taro-TOP it-ACC ate
'Taro ate it.'

[*] This article is based on our presentation "*Heavy-Head-Light: Generalizing Greenberg's Universal #25" at Association for Linguistic Typology 8, University of California, Berkeley July 23-26, 2009. We would like to thank the participants of the conference.

[1] DP includes nominals without determiners (i.e. articles) such as *sushi* in (2b).

b. *Taro-wa sushi-o tabeta*. (DP-V)
Taro-TOP sushi-ACC ate
'Taro ate sushi.'

That is, there is no language in which a verb is followed by its pronominal object and preceded by its nominal object. In addition to consistent VO languages (V-Prn/DP) and consistent OV languages (Prn/DP-V), there are mixed languages such as Bantu and Romance which show Prn-V-DP order, as shown in (3).

(3) Prn-V-DP: Swahili
a. *ni-li-ki-tafuta*. (Prn-V)
I-PST-it-look
'I looked for it.'
b. *ni-li-tafuta kis*u. (V-DP)
I- PST-look knife
'I looked for a knife.'

However, there is no 'mirror-Bantu/Romance' language (*DP-V-Prn), whose structure we show for illustrative purposes in English in (4).

(4) *DP-V-Prn: unattested
a. # *John that girl loves*. (DP-V)
b. # *John loves her*. (V-Prn)

Greenberg (1963) does not discuss the reason why there are no languages with DP-V-Prn order. In section 3, we argue that it is harder for nominal objects than pronominal objects to be placed before the verb to make head-final constituents.[2]

One might argue that Äiwoo (Austronesian, Oceanic) is a counterexample to *DP-V-Prn: Næss (2010) reports that in this language, nominal objects precede the verb as in (5a) while pronominal objects are postverbal as in (5b).

[2] One might raise a question whether the pronominal elements in the data are clitics or full pronouns. However, in phonological point of view, some 'full' pronouns can be considered clitics (e.g. *Je le vois*). Also, the order of a clitic and its host is relevant here. Thus, we use the term 'pronominal' to include clitic pronouns (e.g. (3a)) as well as full pronouns.

(5) a. *John i-togulo-mu*
John PFV-hit-2MIN.A (PFV=perfective, MIN=minimal number, A=augmented)
'You hit John.'

b. *i-togulo-mu iu*
PFV-hit-2MIN.A 1MIN
'You hit me.'

However, this language has the subject pronoun immediately following the verb and preceding the object pronoun (V-Prn(Subj)-Prn(Obj)). In this sense, (5b) is not a pure counterexample to *DP(Obj)-V-Prn(Obj), which assumes a sister relation between verb and its pronominal object as in (4). It seems that in (5b) the object pronoun *iu*, which does not have much information, is dislocated from the preverbal (clause-initial) position, which is the focus position in this type of sentences with O-Verb (cf. Næss 2017: 7). Then, we can attribute this seeming counterexample to *Heavy-Head-Light to the information structure of this type of sentences in Äiwoo, which also has SVO order in sentences with A-Verbs. We discussed this matter in Tokizaki and Kuwana (2020).

2. *Heavy-Head-Light: Generalizing Greenberg's Universal #25

2.1 *Heavy-V-Light

In this section, we point out some universals similar to Greenberg's #25 (*DP-V-Prn) and propose a generalized constraint *Heavy-Head-Light on word order in the world's languages. Here we define the head of a phrase as the word that determines the syntactic category of that phrase. We also define the head of a word as the affix that determines the syntactic category of that word. First, Newmeyer (2005: 5) points out that no language has nominal objects obligatorily in post-verbal position and sentential objects obligatorily in pre-verbal position.[3] If we represent a clause with a conjunction as CP (complementizer phrase) and nominal objects as DP (determiner phrase) (cf. Radford 1997), we can schematize this generalization as *CP-V-DP.[4]

[3] Newmeyer (2005: 5) cites this universal as a personal communication with Luis Vincente.

[4] CP includes clauses without complementizer (i.e. subordinate conjunctions) such as the subordinate clause in *I think [she is right]*.

(6) DP-V-CP: Persian

a. *Ali ye ketaab xarid*. (DP-V)
Ali a book bought
'Ali bought a book.' (Kahnemuyipour 2009: 10)

b. *An zan mi danat ke an mard sangi partab kard.* (V-CP)
that woman CONT knows COMP that man rock throw did
'The woman knows that the man threw a rock.' (Dryer 1980: 130)

(7) *CP-V-DP: unattested

a. # *The children broke the window*. (V-DP)

b. # *The woman* [*that the man threw a rock*] *knows*. (CP-V)

Sentential objects (CP) are heavier than nominal objects (DP), which are heavier than pronouns (Prn), in the sense that CP is generally longer than DP, which is longer than pronouns.[5] Thus, we can generalize *DP-V-Prn and *CP-V-DP into *Heavy-V-Light, which prohibits a verb from following a heavy (i.e. long) object and preceding a light (i.e. short) object.

2.2 *Heavy-P-Light

*Heavy-V-Light can further be generalized into *Heavy-Head-Light if we consider adpositions (P) and nouns (N) in addition to verbs (V) as the heads of constituents, i.e. PP and NP. First, let us consider the order of an adposition and its object. In addition to P-Prn/DP (preposition) and Prn/DP-P (postposition), there is a mixed type Prn-P-DP in Germanic languages.[6]

[5] One might argue that CP can be lighter than DP (e.g. *Sam saw* [*that they were gone*] (4-word object) vs. *Sam saw* [*the signs of their hasty departure*] (6-word object)). However, the CP in the former sentence does not have so much information as the DP in the latter sentence. In order to express the proposition with the same amount of information, the latter sentence would be *Sam saw* [*their departure*] (2-word object).

[6] The pronominals *there, da-* and *daar-* in (8)-(10) are in fact pronominal adverbs rather than pronouns as the glosses show. Here we analyze these forms as Prn because their meanings are similar to pronouns as shown in the translation.

(8) English

a. therein (< in there) (Prn-P)

b. in that place (P-DP)

(9) German

a. *damit* (< *mit da-*) (Prn-P)
there-with
'with it'

b. *mit Honig* (P-DP)
with honey
'with honey'

(10) Dutch

a. *daarmee* (< *met daar-*) (Prn-P)
there-with
'with it'

b. *met luchtpost* (P-DP)
with airmail
'by airmail'

However, there is no 'mirror-Germanic' language (*DP-P-Prn) as far as we know.

Second, let us assume adverbial subordinators (i.e. subordinate conjunctions, cf. Dryer 2013b) to be a kind of preposition taking a sentential complement. For example, the preposition *for* in English can take a sentential complement (represented here as IP (inflectional phrase, cf. Radford 1997)) as well as a nominal object (*for Mary*; *for* [IP *it was raining*]). Then, subordinate clauses can be represented as P-IP or IP-P. Examination of the data in Dryer (2013a, b) shows that out of 585 languages, 53 languages including Finnish have the DP-P-IP order while only two languages (Buduma (Afro-Asiatic, Chadic, Biu-Mandara) and Gününa Küne (Chon, Puelche)) have the IP-P-DP order.[7] Then, we can formalize another constraint on word order *IP-P-DP.

[7] These are the results of a combination of Order of Adposition and Noun Phrase (#85A) and Order of Adverbial Subordinator and Clause (#94A) in The World Atlas of Language Structures (WALS, https://wals.info). Languages with postpositions and initial subordinator words are taken as DP-P-IP; languages with prepositions and subordinating suffixes (Gününa Küne) or with prepositions and final subordinator words (Buduma) are taken as IP-P-DP.

(11) DP-P-IP: Hungarian

a. *a* *fiúk* *előtt* (DP-P)
the boys before
'before the boys'

b. *Mielőtt* *dolgozni* *kezd,* *iszik* *egy* *kávét.* (P-IP)
before work start drink a coffee
'Before start working, he drinks a cup of coffee.'

Moreover, a closer examination of these two exceptional languages, Buduma and Gününa Küne, shows that the descriptions of adpositions in Dryer (2013a) need to be reconsidered. In addition to prepositions, Buduma has a small number of postpositions (*-ro, -re* 'for, to'; *-ga* 'in'), which are borrowed from Kanuri (Nilo-Saharan, Western Saharan) (Lukas and Nachtigal 1939: 70–71). Importantly, these postpositions, sometimes combined with demonstratives, are also used as adverbial subordinators, as shown in (12) (Lukas and Nachtigal 1939: 79-80).

(12) a. *nacagé* *mare*
spoken this-for
'as he has spoken'

b. *narí* *fú;* *narí* *mare*
PRET-bring town PRET-do this-for
'he brought (it) to town; as he had done (it)'

c. *yamǻ́* *are* *āmáị̀* *ānmá*
PRET-stand up when water PRET-come
'when she got up, the water came'

(13) a. *yāmā* *ga* *yākẹ̀lẹ̀* *bula*
PRET-stand up when PRET-walk around town
'when she got up, she walked around the town'

b. *uli* *dúlima* *mána* *hábahan* *nahaŋgẹ̀* *ga* *nāle*
boy leper word friend's PRET-heard as went away
'as the leper boy heard the word of his friend, he went away'

Thus, Buduma has a consistent head-final order with respect to the adpositions taking a sentential complement (IP/DP-P) and does not provide a counterexample to the constraint *IP-P-DP.

Gününa Küne, the other seeming counterexample to *IP-P-DP, in fact has postpositions instead of prepositions (cf. Dryer 2013a), as shown in the following examples (Casamiquela 1983: 58).[8]

(14) a. *tʃɤku ˈgakutek ka ˈwal-hna*
go up horse-on
'I climb on horseback'

b. *u ˈpat̪-na*
road-by
'by the road'

Thus, Gününa Küne as well as Buduma has the IP/DP-P order. Then, we can say that no language with the IP-P-DP order has been attested in the world.

As we generalized *DP-V-Prn and *CP-V-DP into *Heavy-V-Light in section 2.1, we can generalize *DP-P-Prn and *IP-P-DP into *Heavy-P-Light.

2.3 *Heavy-N-Light

The third subcase of *Heavy-Head-Light is *Heavy-N-Light, which is a generalization of *PP-N-Prn and *CP-N-DP as we will see in this section. First, let us consider *PP-N-Prn. For example, in English, *'s*-genitives are generally preferred for short phrases including pronouns (e.g. *his book*), while *of*-phrases are preferred for longer phrases (e.g. *books of this period*) (Biber et al. 1999) (Prn-N-PP). However, there are only a limited number of 'mirror-English' languages (e.g. Tauya (Madang, Trans-New Guinea) cf. Siewierska 2001: 140). Then, we can formulate another constraint *PP-N-Prn. The problematic languages with genitive-noun order and noun-pronoun (genitive) order (PP-N-Prn) are listed in (15).[9]

[8] The description of Order of Adposition and Noun Phrase in Gününa Küne in WALS (https://wals.info/languoid/lect/wals_code_gku, as of August 17, 2018) is flawed perhaps because of the heading of the section 5.1.6 "Preposición" in Casamiquela (1983: 58). Casamiquela writes "Son en realidad, las preposiciones, verdaderas posposiciones o sufijos. (The prepositions are in fact postpositions or suffixes.)"

[9] Siewierska (2001: 140) also reports that there are seven languages with genitive-noun and noun-pronoun (genitive) or pronoun (genitive)-noun order, which are listed in (i).

(15) Bari (Nilotic, Nilo-Saharan), Dakota (Core Siouan, Siouan), Diola Fogny (North Atlantic, Niger-Kordofanian), Doyayo (Adamawa, Niger-Congo), Hamar (Omotic, Afro-Asiatic), Kobon (Madang, Trans-New Guinea), Tauya (Madang, Trans-New Guinea), Tonkawa (Tonkawa, Tonkawa), Ungarijn (Wororan, Australian), Vanimo (aka, Dumo) (Western Skou, Skou), Valley Yokuts (Yokutsan, Penutian)

Here, we argue that Tauya is not a counterexample to our generalization *PP-N-Prn. Tokizaki and Kuwana (2009) point out that Tauya has two forms of genitive suffix: pi attached to pronouns as in (16a) and na attached to full nouns as in (16b) (MacDonald 1990: 131, 133).

(16) a. *wate* *ne* - *pi* (N-Prn)
house 3S GEN
'his/her house'

b. ʔ*e* *fanu-na* *wate* (PP-N)
DEM man-GEN house
'that man's house'

These two constructions might seem to constitute a counterexample to our *PP-N-Prn, a subcase of *Heavy-N-Light. However, Tokizaki and Kuwana argue that *pi* and *ne* are in different categories: *pi* is a genitive marker while *na* is a kind of postposition. The third person pronoun *ne* suffixed by *-pi* is the D head of the whole Determiner Phrase (DP) as shown in (17a). In (17b) postposition *na* takes DP ʔ*e fanu* 'that man' as its complement; the PP ʔ*e fanu-na* 'that man's' precedes and modifies the head N *wate* 'house' of the whole NP.

(17) a. [$_{DP}$ [$_{N}$ *wate*] [$_{D}$ *ne* - *pi*]] [$_{DP}$ N-D]
house 3S GEN
'his/her house'

b. [$_{NP}$ [$_{PP}$ [$_{DP}$ ʔ*e* *fanu*] *-na*] [$_{N}$ *wate*]] [$_{NP}$ PP-N]
DEM man -GEN house
'that man's house'

(i) Bandjalang (Pama-Nyungan), Dogon (Niger-Kordofanian), Nadeb (Makú), Nasibi [Nasioi (East Bougainville)], Tehit (West Papuan), Yessan Mayo (Sepik), Yimas (Sepik)

It is necessary to carefully examine the data in each of these languages. We will leave that task for future research.

Note that the whole phrase in (17a) is DP while that in (17b) is NP. Thus, (17a) is not a counterexample to *[NP PP-N-Prn]. In other words, Tauya has head-final order in both DP and NP.

The claim that *na* is categorially different from *-pi* is supported by three facts. First, so-called genitive marker *na* in (17b) is homophonous with a relative *na*, as shown in (18) (MacDonald 1990: 289).

(18) [NP [CP [IP *ya -* *ni* *ø* *ø* *-yau -* *e -*] *na*] *fanu*]
1S ERG 3S see 1/2 REL man
'the man I saw'

Then, it is possible to claim that *na* is in fact one word that functions as a head P or C taking DP or IP as its complement in its projection PP or CP.

Second, Tauya has two kinds of expressions for numerals. Numbers from one to four are expressed with a numeral word following the noun. Numbers over four are expressed with a relative clause preceding the noun.

(19) a. [NumP N *awi*]
two
'two N'

b. [NP [*wesaʔa* *awi (fofe)* *te-* *a-* *na*] N]
half two come get 3S REL
'seven N'

These two forms (19a) and (19b) parallel the two forms in (17a) and (17b). In (17a) the modifying word *ne-pi* (D) and in (19a) *awi* (Num(ber)) follow the noun and make their projections DP and Num(ber)P. In (17b) and (19b) *na* functions as a relative modifying the following noun. All of the constituents in (17) and (19), i.e. DP, NumP and NP are head-final.

Third, the pronoun-*pi* N order in (20b) is less likely than the N pronoun-*pi* order in (20a), but is possible in Tauya (MacDonald 1990: 132).

(20) a. *afe* *ten - pi* (N-Prn-*pi*)
mother 2P GEN
'your mother'

b. *ten - pi afe* (Prn-*pi*-N)
2P GEN mother
'your mother'

The partial optionality of word order seems to reflect the ambiguous functions of pronoun-*pi*.

Pronoun-*pi* is more likely to be interpreted as the D head of the DP *afe ten - pi* as in (20a), but it can be interpreted as a modifier of the head noun *afe* in the NP *ten - pi afe* as in (20b). Importantly, the word order is head final in both (20a) and (20b). We can explain the alternative word orders in (20) in terms of the ambiguous grammatical categories of nominal phrases if we assume that N-Prn-*pi* in (20a) is a DP with Prn-*pi* as its head while Prn-*pi*-N in (20b) is an NP with Prn-*pi* as a modifier of the head noun.

Of course we need to investigate the languages with NP-N-Prn order other than Tauya. We plan to investigate the other languages given in (15) and (i) in footnote 9 in future research in order to judge whether they are true counterexamples or not.

Let us turn to *CP-N-DP, the other subcase of *Heavy-N-Light. Our examination of the data in Dryer (2013c, 2013d) shows that out of 755 languages, there are 161 languages with Genitive-N-Relative order and only two languages (Tigré (Afro-Asiatic, Semitic) and Amis (Austronesian, East Formosan)) with Relative-N-Genitive order (*CP-N-DP).[10] Swedish is an example of the DP-N-CP order as shown in (20) (Swedish).

[10] This is the result of combining Order of Genitive and Noun (#86A) and Order of Relative Clause and Noun (#90A). The numbers of languages with each word order combination are shown in (i) (the relevant orders underlined).

(i)

Noun-Genitive / Noun-Relative clause	328
Genitive-Noun / Noun-Relative clause	161
Genitive-Noun / Relative clause-Noun	128
Genitive-Noun / Mixed	50
No dominant order / Noun-Relative clause	44
Genitive-Noun / Internally headed	20
Genitive-Noun / Correlative	6
Noun-Genitive / Mixed	4
Genitive-Noun / Adjoined	4
No dominant order / Internally headed	3

(21) a. *min frus födelsedag* (DP-N)
my wife's birthday
'my wife's birthday'

b. *en vän, som bor i Stockholm* (N-CP)
a friend who lives in Stockholm
'a friend who lives in Stockholm'

However, careful examination of the data shows that Tigré is not a real counterexample to the generalization *CP-N-DP. First, Genitive construction in Tigré has Genitive-Noun order (22a, b) as well as Noun-Genitive (22c) (Raz 1983: 81).

(22) a. *nāy yom məhro*
of today lesson
'today's lesson'

b. *nāy həwān qarbat*
of animal skin
'animal skin'

c. *ʾəb hatte baʿat nāy ladəbaʿ*
in certain cave of bush
'in a certain cave in the bush'

Noun-Genitive / Relative clause-Noun	2
No dominant order / Adjoined	2
Noun-Genitive / Internally headed	1
No dominant order / Mixed	1
Genitive-Noun / Doubly headed	1

The order Noun-Genitive / Mixed (with dotted underline) may be a counterexample to the strong version of *CP-N-DP. The languages of this category are Rukai (Tanan, Austronesian: NRel or RelN), Murrinh-Patha (Australian, Daly: NRel or internally-headed, postnominal relative clauses), Squamish (Salishan, Central Salish: NRel or internally-headed, postnominal relative clauses) and Tukang Besi (Austronesian, Celebic: NRel or internally-headed, postnominal relative clauses). Thus, only Rukai (Tanan) has the problematic order CP-N-DP/CP.

Second, Tigré has Noun-Relative order as in (23) as well as Relative-Noun order as in (24) (Raz 1983: 43).[11]

(23) a. *lakətāb laʾəlu bəka*
the book which it you have
'The book you have'

b. *waʾət ḥatte dəgge* [*dəgge wānin latətbahal*] *nabro ʾalaw*
and in one village village animals which is called living they were
'And they were living in a village called "the village of animals"'

c. *dəgəm worot ʾənās higa kulla laḥəyāyət ʾammer laʾala*
tale one man language all of her the animals knowing who was
'The tale of a man who knew the language of all animals'

(24) a. *laʾəglu ḥazze ʿalko kətāb*
which it looking for I was book
'The book which I was looking for'

b. *ʾəbbəlli ḥəmām lamoṯ̄m bəzḥām kəm tom ʾəb*
because of this illness who died many (pl.) that they are according to
tārik ʾəgəl naʾammər ʾənqaddər
history to [we] know we can
'According to history we know that [those] who died by this illness were many.'

The examples in (23) show that the head noun precedes the relative clause. The data in (22) to (24) show that Tigré has CP/DP-N-DP/CP order. Thus, Tigré is not a counterexample to the generalization *CP-N-DP.

Amis (Austronesian, East Formosan) is the other problematic language for the generalization *CP-N-DP according to Dryer (2013c, 2013d), who describes Amis as a language with Noun-Genitive and Relative clause-Noun order. However, Wu (2006: 94) observes that "the genitive noun modifier can also appear before the modified noun with the optional presence of the linker *a*," as shown in (25b).

---

[11] Citing Palmer (1961), Raz (1983: 43) notes the frequency of N-Rel order in speech as shown in (i).

(i) Most commonly the relative clause precedes the noun it modifies. The general impression is that in literary or formal texts, relative clauses which follow the modified noun are rare. However, in utterances which represent the spoken language the modified-modifier order is met with quite frequently.

(25) a. *tamdaw n-u takaw*
person GEN-CN Kaohsiung (CN: common noun)
'person from Kaohsiung (i.e. currently living there)'
b. *n-u takaw (a) tamdaw*
GEN-CN Kaohsiung LNK person (LNK: linker)
'person of the Kaohsiung team (in contrast with the Taipei team in a sports event)'

Wu notes that when the genitive noun modifier shows up prenominally as in (25b), it gives an emphatic tone to the modifier. Wu also shows a similar example with a prenominal modifier and the linker *a* as in (26).

(26) *takaw a tamdaw*
Kaohsiung LNK person
'person who was born and grew up in Kaohsiung'

Here, a bare noun modifier with a linker modifies the following noun. These examples show that Amis can have Genitive-Noun order as well as Noun-Genitive order (25a) (DP-N-DP).

The linker *a* also appears in what Wu (2006: 95) calls the clausal modifier, because of the factual marker *-ay*, as shown in (27).

(27) a. *kuhting-ay* (*a*) *ayam*
black-FAC LNK bird
'black bird'
b. *Tati'ih k-u-ya ma-ka'en-ay n-i aki a tali.*
bad NOM-CN-that UV-eat-FAC GEN-PPN Aki LNK taro
'That taro that Aki ate was bad.' (UV: undergoer voice, PPN: personal proper noun)

Wu (2006: 96) observes that "adjective-like" clausal modifiers always appear before a preposed head noun as in (27a) while relative clause-like clausal modifiers can appear before or after a preposed head noun as in (28a) and (28b) (CP-N-CP).[12]

[12] Wu (2006: 73) observes that "Like most of the Formosan languages, Amis is a verb-initial, or more precisely, a predicate initial language."

(28) a. *Ya mi-palu-ay ci mayaw-an a ta-tusa-ay a*
that AV-beat-FAC PPN Mayaw-DAT LNK PL-two-FAC LNK
(RC-like modifier)
fa'inayan a ***singsi*** *paka-araw ci sawmah-an.*
man LNK teacher ABLT-see PPN Sawmah-DAT
(Head)
'Those two man teachers who beat Mayaw saw Sawmah.'
(AV: actor voice, ABL: abilitative)

b. *Ya ta-tusa-ay a fa'inayan a* ***singsi*** *(*a)*
that PL-two-FAC LNK man LNK teacher LNK
(Head)
mi-palu-ay ci mayaw-an paka-araw ci sawmah-an.
AV-beat-FAC PPN Mayaw-DAT be.able.to-see PPN Sawmah-DAT
(RC-like modifier)
'Those two man teachers who beat Mayaw saw Sawmah.'

Note that the linker *a* does not appear after the head noun *singsi* in the N-Relative order as shown in (28b). The linker *a* links the modifier on the left of the head to the head as in (25b), (26), (27) and (28a) but not the modifier on the right of the head as shown with (**a*) in (28b). In this sense, the linker *a* is like a postposition in head-final languages. On the other hand, the genitive marker *n* is prepositional in that it takes its object on the right and modifies the head noun on the left as in (25a).

These examples in (25) to (28) show that Amis has CP/DP-N-DP/CP order and is not a pure counterexample to the generalization *CP-N-DP.

2.4 *Heavy-Head-Light

So far we have discussed unattested word orders with three types of head in the world's languages: *Heavy-V-Light, *Heavy-P-Light and *Heavy-N-Light. Thus, we can generalize these three into a constraint *Heavy-Head-Light. The result of the discussion is summarized as in (29).

(29) a. *DP-V-Prn: Greenberg's Universal #25
b. *CP-V-DP
c. *DP-P-Prn (cf. 11 languages, e.g. Tauya)
d. *IP-P-DP (cf. Buduma and Gününa Küne)
e. *DP-N-Prn

f. *CP-N-DP (cf. Tigré and Amis)

We have argued that the exceptional languages shown in the brackets in (29) are not real counterexamples to *Heavy-Head-Light (30).[13]

(30) *Heavy-Head-Light
Heaviness of constituents: CP/IP (clause) > DP (nominal phrase) > Pronoun

If this generalization holds in the world's languages, it must reflect a deep principle governing languages. We will explore the principle in the next section by asking why the generalization holds cross-linguistically.

3. Why *Heavy-Head-Light?

3.1 Problems of functional/performance explanations

Finally, let us consider why this pattern Heavy-Head-Light is not allowed in languages. One might argue that this is due to its marked information structure, which goes against the prevalence of light-heavy order in languages. However, performance theories such as Hawkins's (1994) Early Immediate Constituents (EIC) does not explain why Light-Head-Heavy orders are prevalent in languages, as we have seen in section 2. EIC predicts that languages prefer the word order that is more efficient in parsing. Let us consider the sentences with a dative PP and a complex NP object in (31).

(31) a. I introduced [some friends that John had brought to the party] [to Mary]
1 2 3 4 5 6 7 8 9 10 11
b. I introduced [to Mary] [some friends that John had brought to the party]
1 2 3 4

EIC counts the number of words (or the number of constituents in a more precise definition) between the head and the first word of its daughters. The number is 11 in (31a) and 4 in (31b). Then, EIC correctly chooses (31b) rather than (31a).

Now, let us consider the order of a noun and modifier in Swedish as an example of Light-Head-Heavy. The example (21) is repeated here as (32).

---

[13] This constraint *Heavy-Head-Light is in fact *Heavy complement-Head-Light complement. Thus, the examples such as Heavy subject-Head-Light complement are not counterexamples to *Heavy-Head-Light.

(i) [That they were always so negative] annoys me. [CP-V-pro]

(32) DP-N-CP: Swedish

a. *min frus födelsedag* (DP-N)
my wife's birthday
'my wife's birthday'

b. *en vän, som bor i Stockholm* (N-CP)
a friend who lives in Stockholm
'a friend who lives in Stockholm'

These examples would be (33) in the reverse order CP-N-DP.

(33) a. * *födelsedag min frus* (*N-DP)
birthday my wife's

b. * *som bor i Stockholm en vän* (*CP-N)
who lives in Stockholm a friend

EIC nicely predicts the order of the head noun and relative clause as in (32b) rather than (33b). In (32b), the relative clause immediately follows the head noun; the number of words between the head and the relative clause is 0, the minimum. In (33b) the head noun *vän* is separated by four words from the first word in the relative clause *som*. However, EIC does not explain why (32a) is preferred to the reverse order shown in (33a). The acceptable example (32a) has a word *frus* between the first word of the genitive phrase *min* and the head noun *födelsedag* 'birthday' while the unacceptable example (33a) has no word between the head noun *födelsedag* and the first word of the genitive phrase *min*. According to EIC, (33a) should be preferred to (32a). Thus, EIC does not explain why DP-N-CP is prevalent in languages. Similarly, DP-V-CP (6) and DP-P-IP (11) are also seen in languages, as we saw in section 2. Here we repeat these sentences as (34) and (35).

(34) DP-V-CP: Persian

a. *Ali ye ketaab xarid*. (DP-V)
Ali a book bought
'Ali bought a book.' (Kahnemuyipour 2009: 10)

b. *An zan mi danat ke an mard sangi partab kard*. (V-CP)
that woman cont knows comp that man rock throw did
'The woman knows that the man threw a rock.' (Dryer 1980: 130)

(35) DP-P-IP: Hungarian

a. *a fiúk előtt* (DP-P)
the boys before
'before the boys'

b. *Mielőtt dolgozni kezd, iszik egy kávét.* (P-IP)
before work start drink one koffee
'Before start working, he drinks coffee.'

In (34a), the head of DP (i.e. D *ye*) is separated from the head of VP (i.e. V *xarid*) by a noun *ketaab.* EIC wrongly predicts that the order V-DP in (36), which has no intervening word between V *xarid* and D *ye*, is preferred to the order DP-V in (34a).

(36) * *Ali xarid ye ketaab*. (V-DP)
Ali bought a book

Similarly, as the following example shows, Hungarian does not allow the P-DP order, which should be preferred to the DP-P order in (35a) according to EIC: (37) has no intervening word between P *előtt* and the head of DP *a*, while (35a) has an intervening word *fiúk*. EIC does not explain why Hungarian has DP-P order and P-IP order as in (35b).

(37) * *előtt a fiúk* (P-DP)
before the boys

EIC does not explain the prevalence of these orders, which can be generalized as DP-Head-IP. EIC predicts that complements should appear on the same side of the head even if their lengths are different.

Moreover, EIC does not explain why some languages have disharmonic word orders between Prn-Head and Head-DP as we have seen in the cases of Prn-V-DP (3), Prn-P-DP (8)-(10) and Prn-N-PP (*his book/books of this period*). According to EIC, a pronoun could follow the head with no problem as DP/PP do because a pronoun is a single word. However, there are some languages that places pronouns before the head and DP/PP after the head. EIC does not explain why these languages choose the disharmonic light-head-heavy order rather than the consistent head-light/heavy order if they are equivalent in their EIC score.

Thus, the performance theory of word order using EIC is too strong in ruling out acceptable Light-Head-Heavy orders in languages of the world. We will investigate an alternative theory of possible word orders in the next section.

3.2 Heaviness constraint on compounding

The question to be answered is why Heavy-Head-Light is unacceptable while Light-Head-Heavy is acceptable. In this section, we propose an analysis based on the asymmetry of juncture strength between left-branching and right-branching structures. We define juncture as the strength of connectedness between two constituents.

We have argued that left-branching structure has stronger juncture between its constituents than right-branching structure (Tokizaki 2008, cf. Wagner 2005). In other words, left-branching (i.e. head-final) constituents such as OV and postpositional phrases (DP-P) are phonological compounds while right-branching (i.e. head-initial) constituents such as VO and prepositional phrases (P-DP) are bona fide phrases. We have presented various phenomena as evidence for the junctural asymmetry, which include Sequential Voicing in Japanese (Otsu 1980) and *n*-Insertion in Korean (Han 1994), interfixation in German and Dutch (Krott et al. 2004), agglutinativity in OV languages (Lehmann 1973, Plank 1998, cf. Kayne 1994), junctural asymmetry between prefixes and suffixes (Hyman 2008) and the Left Branch Constraint (Ross 1967). Here we show some examples of these phenomena.

Firstly, Japanese Sequential Voicing occurs across the constituent boundary in left-branching structure but not at the constituent boundary in right-branching structure, as shown in (38) (Otsu 1980).

(38) a. [[*nise tanuki]* *shiru*] > *nise danuki jiru* (**shiru*)
mock badger soup
'mock-badger soup (soup made with mock-badger)'

b. [*nise* [*tanuki shiru*]] > *nise tanuki jiru* (**danuki*)
mock badger soup
'mock badger-soup (soup not made with badger but with pork, etc.)'

In other words, constituents in a left-branching structure are tied closely together while constituents in a right-branching structure are separated. Similarly, Korean *n*-Insertion applies across the constituent boundary in left-branching structure but not in right-branching structure (cf. Han 1994). Note also that the pitch contour in (38a) (LH HHH HL) is different from that in (38b) (LH LHH HL): left-branching (38a) is pronounced with just one pitch fall (HL) at the end as a word while right-branching (38b) is pronounced with two pitch falls as two separate words.

Secondly, Krott et al. (2004) show that 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. (39) shows examples with unmarked interfix (IF) occurring at the constituent boundary, and (40) shows those with marked interfixes occurring in a constituent. The numbers of examples with *-s-* and all interfixes are shown in the parentheses.

(39) a. [[*grond*+*wet*]-*s*-*aartikel*] (-*s*- 25; all 39) [left-branching, unmarked IF]
ground-law-IF-article 'constitution'

b. [*arbeid*-*s*-[*vraag*+*stuk*]] (-*s*- 38; all 60) [right-branching, unmarked IF]
employment-IF-question-issue 'labor issue'

(40) a. [[*scheep*-*s*-*bouw*]+*maatschappij*] (-*s*- 13; all 50) [left-branching, marked IF]
ship-IF-building company

b. [*hoofd*+[*verkeer*-*s*-*weg*]] (-*s*- 3; all 11) [right-branching, marked IF]
main traffic-IF-road

The ratio of the unmarked interfix position (39a) and (39b) to the marked interfix positon (40a) and (40b) is higher in right-branching (-s- 38(39b)÷3(40b)=12.7; all 60(39b)÷11(40b)=5.5) than in left-branching (-s- 25(39a)÷13(40a)=1.9; all 39(39a)÷50(40a)=0.8). Comparing the ratio of right-branching 12.7 and 5.5 with left-branching 1.9 and 0.8, we conclude that interfixes in the unmarked position are more likely to occur in right-branching compounds than in left-branching compounds.

Thirdly, Hyman (2008: 323) argues that suffixes tend to be more tightly bound to their root than prefixes. Similarly, Julien (2002: 226) points out that a suffix bears a close structural relation to the root that it attaches to while the structural relation between a prefix and the root it attaches to is less stable. These observations also support the asymmetry in juncture because [prefix [$_{\text{Root}}$ ...]] is right-branching while [[$_{\text{Root}}$ ...] suffix] is left-branching.

Fourthly, left-branching structure behaves like a word or a compound in that it does not allow extraction of its constituent (Left Branch Condition (Ross 1967)).

(41) a. The boy [[whose guardian's] employer]$_i$ we elected t_i president ratted on us.

b. *The boy [whose guardian's]$_i$ we elected [t_i employer] president ratted on us.

In (41a, b), whose guardian's employer is a left-branching structure whose constituents cannot be extracted as shown in (41b). This contrasts with the fact that extraction from right-branching structure is possible, as shown in (42).

(42) Who$_i$ did you see [a [picture [of t_i]]]?

Thus, we have phonological, morphological and syntactic evidence for the claim that left-branching structure has stronger juncture than right-branching structure.[14]

All these facts support our junctural asymmetry hypothesis: the juncture between constituents is stronger in left-branching structure than in right-branching structure. Note that we assume that morpho-phonological processes such as compound word formation and syntactic constituent structure are essentially the same: the same operation, Merge in the current generative theory, builds up compounds, phrases, clauses and sentences.

If these arguments are on the right track, we can say that a left-branching constituent such as complement-head order is phonologically more compound-like than a right-branching constituent such as head-complement order. We will use an equals symbol (=) for the strong juncture between a head and its complement in complement-head order as in (43a) and a hyphen for the weak juncture between them in head-complement order as in (43b).[15]

(43) a. [$_X$ Complement=Head]
b. [$_{XP}$ Head-Complement]

The compound-like status of complement-head order is shown as the category label X in (43a) and the phrasal status of head-complement order is shown as XP in (43b).[16]

---

[14] One might wonder why Scrambling of object is possible in OV languages. A possible answer is that Scrambling is not movement but base generation (cf. Neeleman 2011).

[15] One might raise questions about OXV languages and OVX languages (X is an oblique phrase). We assume that in OXV languages, X almost equals to O in length (cf. Hawkins 2008: 182) and makes a phonological compound with the following verb (X=V) because these languages have postpositions, which are tied close to their object. Thus, XV in these languages is also a case of complement=head (43a). This phonological compound XV merges with O to make a larger constituent OXV. The asymmetry between the high percentage of OVX languages in OV languages and the very low percentage of XVO in VO languages can also be explained in terms of juncture and *Heavy-Head-Light: O is lighter than X assuming that X is PP. We will argue this point in another paper. Note that the existence of OVX languages are also explained with this account: the OVX is Light=V-heavy if we assume that O is lighter than X, which consists of adposition and its object. OVX conforms to the Light-Head-Heavy pattern.

[16] We argue that all the head-final constituents including phrases and clauses are compound-like because of their strong juncture.

Now we can explain why the Heavy-Head-Light order is impossible in terms of the strong juncture in complement-head order. In order for a language to have a Heavy-Head-Light order, the language must allow a heavy/long phrase in a compound word and prohibit a light/short phrase in a compound, as shown in (44).

(44) a. [$_X$ Heavy=Head]
b. * [$_X$ Light=Head]
c. [$_{XP}$ Head-Light]
d. * [$_{XP}$ Head-Heavy]

However, this combination of word orders is against the length limitation of compounds in a given language. If a language allows a long constituent in a compound as in (44a), it should also allow a short constituent in a compound as in (44b), assuming that it is easier to make a compound with a short constituent than a long one, as shown in (45).

(45) a. sweetheart
b. ? [very sweet] heart
c. * [sweet like candy] heart

Even if the language disallows a light/short constituent as in (44b), it should allow a heavy/long constituent in a phrase as in (44d): there is nothing to prevent a long phrase from occurring at the right of a head because the resulting constituent is a phrase (head-complement) without size limitation and not a compound (complement-head) with the limitation. Thus the combination of the two word orders shown in (44) is impossible.[17] On the other hand, the combination in (46) is possible.

(46) a. [$_X$ Light=Head]
b. * [Heavy=Head]
c. [$_{XP}$ Head-Light]
d. [$_{XP}$ Head-Heavy]

The order Light-Head-Heavy shown in (46) is possible in languages where a complement is incorporated into the head to make a short (phonological) compound if and only if the complement is light (short). If the complement is heavy (long), the resulting

[17] An explanation in terms of language acquisition might also be possible for the unattested Heavy-Head-Light order in (44): children acquire short compounds to begin with and then realize they can make more complex compounds, not vice versa. We consider this as another plausible analysis.

(phonological) compound is too long to be acceptable. In this sense, some languages are sensitive to the length of phrase that can be positioned to the left of the head (cf. AUTHOR1 2010).[18]

Combinations other than (46) are also possible, as shown in (47) and (48).

(47) a. [$_{XP}$ Head-Light]
b. [$_{XP}$ Head-Heavy]

(48) a. [$_{X}$ Light=Head]
b. [$_{X}$ Heavy=Head]

The languages with the combination in (47) do not allow any (phonological) compounds of any size. The languages with (48) allow (phonological) compounds of any size.

To sum up, the unattested order Heavy-Head-Light could be derived by putting the complement to the left of a head and making a (phonological) compound [$_{X}$ complement-head] only if the complement is heavy. However, this is an implausible condition on compounding: compounding is more likely to apply to a short complement than a long one. Thus, we can explain why *Heavy-Head-Light holds in languages.

4. Conclusion

We have argued that Greenberg's (1963) Universal #25 (*DP-V-pron) can be extended to a general constraint on head-complement order *Heavy-Head-Light (*Heavy-X-Light), which prohibits languages from having a heavy/long complement to the left of a head and having a light/short complement to the right of a head. The head in this constraint *Heavy-X-Light can be a verb (V), adposition (P) or noun (N). The relative length or heaviness of constituents can be defined as pronouns (pron) < noun phrases (DP) < clauses (CP). The generality of this constraint seems to show a basic property of languages, which we discussed in terms of the phonological compound character of head-final order. If a language allows a heavy/long complement to the left of a head making a long phonological compound [Heavy=X], the language should allow a light/short complement to the left of the head making a short phonological compound [Light=X] instead of a phrase [X-Light]. Thus, we can derive the constraint *Heavy-X-Light from the asymmetry of juncture

[18] It is interesting to consider phrasal compounds such as *wait-and-see mentality*, *first-in-last-out policy* and *an I-don't-care attitude*. We observe that phrasal compounds are allowed in languages with lefthand word stress (e.g. Germanic) rather than in languages with righthand word stress (e.g. Romance and Slavic (Szymanek 2017)).

strength between left-branching structure and right-branching structure together with the constraint on compound size, which may be different between languages.

We still need to investigate word orders in languages that may be counterexamples to the constraint *Heavy-Head-Light. We hope that this study shows not only an interesting typological universal but also a new approach to the interface between phonology, morphology and syntax.

References

Abney, Steven Paul. 1987. The English noun phrase in its sentential aspect. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman grammar of spoken and written English*. New York: Longman.

Botha, Rudolf. 1983. *Morphological mechanisms: Lexicalist analyses of synthetic compounding*. Oxford: Pergamon Press.

Casamiquela, Rodolfo M. 1983. *Nociones de gramática del gününa küne*. Paris: Centre National de la Recherche Scientifique.

Dryer, Matthew S. 1980. The positional tendencies of sentential noun phrases in universal grammar. *The Canadian Journal of Linguistics* 25(2), 123–195.

Dryer, Matthew S. 2013a. Order of Adposition and Noun Phrase. In Dryer and Haspelmath (eds.)

Dryer, Matthew S. 2013b. Order of Adverbial Subordinator and Clause. In Dryer and Haspelmath (eds.)

Dryer, Matthew S. 2013c. Order of Genitive and Noun. In Dryer and Haspelmath (eds.)

Dryer, Matthew S. 2013d. Order of Relative Clause and Noun. In Dryer and Haspelmath (eds.)

Dryer, Matthew S. and Martin Haspelmath (eds.) 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/)

Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (ed.) *Universals of language*. Cambridge, Mass: MIT Press, 73-113.

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

Hawkins, John. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

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

Julien, Marit. 2002. *Syntactic heads and word formation*. Oxford: Oxford University Press.

Kahnemuyipour, Arsalan. 2009. *The syntax of sentential stress*. Oxford: Oxford University Press.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass: MIT Press.

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

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

Lukas, Johannes, 1939. *Die Sprache der Buduma im Zentralen Sudan*. (Abhandlungen für die Kunde des Morgenlandes, 24/2.) Leipzig: F.A. Brockhaus.

Lukas, Johannes and Gustav Nachtigal (1939) *Die Sprache der Buduma im zentralen Sudan: auf Grund eigener Studien und des Nachlasses von G. Nachtigal*. Leipzig: Deutsche Morgenländische Gesellschaft : Kommissionsverlag F.A. Brockhaus.

MacDonald, Lorna. 1990. *A grammar of Tauya*. Berlin: Mouton de Gruyter.

Næss, Åshild. 2006. Bound nominal elements in Äiwoo (Reefs): A reappraisal of the "multiple noun class systems," *Oceanic Linguistics* 45(2), 269–296.

Næss, Åshild. 2010. A counterexample to the universal number 502 in the Universals Archive, ed by Frans Plank, Thomas Mayer, Tatsiana Mayorava and Elena Filimonova. https://typo.uni-konstanz.de/archive/intro/index.php.

Neeleman, Ad. 1994. Scrambling is a D-structure phenomenon. In *Studies on Scrambling: Movement and Non-Movement Approaches to Free Word-Order Phenomena*, ed. by Norbert Corver and Henk van Riemsdijk, 387–430, Berlin: De Gruyter.

Nettle, Daniel. 1998. *The Fyem language of Northern Nigeria*. Munchen: Lincom Europa.

Newmeyer, Frederick. J. 2005. *Possible and probable languages*. Oxford: Oxford University Press.

Newton, Glenda. 2007. Complementisers and C particles. Handout for FOFC Seminar. University of Cambridge. http://research.ncl.ac.uk/linearization/index.php.

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

Cambridge, MA: Department of Linguistics and Philosophy, Massachusetts Institute of Technology, 207-227.

Palmer, Frank. R. 1961. Relative clauses in Tigré. *Word* 17, 23-33.

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

Radford, Andrew. 1997. *Syntax: A minimalist introduction*. Cambridge: Cambridge University Press.

Raz, Shlomo. 1983. *Tigré Grammar and Texts*. Undena.

Ross, John R. 1967. *Constraints on variables in syntax*. Ph.D. dissertation. Cambridge, MA: Massachusetts Institute of Technology.

Siewierska, Anna. 2001. Order correlations between free and bound possessors: Perspectives from diachronic change. W. Bisang (ed.) *Aspects of typology and universals*. Akademie Verlag, 133-152.

Szymanek, Bogdan. 2017. Compounding in Polish and the absence of phrasal compounding. Carola Trips and Jaklin Kornfilt (eds.) *Further investigations into the nature of phrasal compounding*. Berlin: Language Science Press, 49-79.

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

Tokizaki, Hisao. 2010. Recursive compounds and word-stress locations, a paper presented at Linguistic Interface II. Belfast. Dec, 2010. (http://ling.auf.net/lingbuzz/001467)

Tokizaki, Hisao and Yasutomo Kuwana. 2009. Tauya-go no zokkaku meishi to gojun no fuhensei [Genitive noun in Tauya and universal word order], *Bunka to Gengo* [*Culture and Language*] 70, 31–41. (https://sapporo-u.repo.nii.ac.jp/records/3357)

Tokizaki, Hisao and Yasutomo Kuwana. 2020. Greenberg's Universal 25 and its exception, an abstract for DGfS Annual Meeting 2020. (https://www.zfs.uni-hamburg.de/dgfs2020/programm/abstracts/dgfs2020-ag04-tokizaki.pdf)

Wagner, Michael. 2005. Asymmetries in prosodic domain formation. *MIT Working Papers in Linguistics* 49, Cambridge, MA: Department of Linguistics and Philosophy, Massachusetts Institute of Technology, 329-367.

Wu, Jing-lan Joy. 2006. Verb classification, case marking, and grammatical relations in Amis. State University of New York at Buffalo Ph.D. dissertation.

ISO/Glottolog

Amis amis1246 639-3 ami
Bandjalang band1339 639-3 bdy
Bari bari1284 iso 639-3 bfa
Buduma budu1265 639-3 bdm
Dakota dako1258 639-3 dak
Diola-Fogony jola1263 639-3 dyo
Dogon dogo1299
Doyayo doya1240 639-3 dow
Dutch dutc1256 639-3 nld
English stan1293 639-3 eng
Finnish finn1318 639-3 fin
German stan1295 639-3 deu
Gününa Küne puel1244 639-3 pue
Hamer(Hamar) hame1242 639-3 amp
Japanese nucl1643 639-3 jpn
Kanuri cent2050 639-3 knc
Kobon kobo1249 639-3 kpw
Korian kore1280 639-3 kor
Murrinh-Paths murr1258 639-3 mwf
Nadëb nade1244 639-3 mbj
Nasibi (Nasioi) naas1242 639-3 nas
Persian west2369 639-3 pes
Rukai (Tanan) ruka1240 639-3 dru
Squamish squa1248 639-3 squ
Swahili swah1253 639-3 swh
Swedish swed1254 639-3 swe
Tauya tauy1241 639-3 tya
Tehit tehi1237 639-3 kps
Tigré tigr1270 639-3 tig
Tonkawa tonk1249 639-3 tqw
Tukang Besi tuka1247
Ungarinjin ngar1284 639-3
Vanimo (Dumo) vani1248 639-3 vam
Valley Yokuts (Yokuts) yoku1256 639-3 yok
Yessan-Mayo yess1239 639-3 yss
Yimas yima1243 639-3 yee

Abbreviations

1/2	first, second singular
1S	first singular
2P	second plural
A	augmented
ABL	abilitative
ACC	accusative
AV	actor voice
CN	common noun
COMP	complementizer
CON	conditional
CONT	continuous
DAT	dative
DEM	demonstrative
ERG	ergative
FAC	factual marker
IF	interfix
GEN	genitive
LNK	linker
MIN	minimal number
NOM	nominative
OBJ	object marker
PFV	perfective
PL	plural
PPN	personal proper noun

PRET	preterite	TOP	topic
PST	past	UV	undergoer voice
REL	relative		