The Crosslinguistic Relationship between Ordering Flexibility and Dependency Length Minimization: A Data-Driven Approach

Zoey Liu
Department of Computer Science
Boston College
ying.liu.5@bc.edu

Abstract

This paper asks whether syntactic constructions with more flexible constituent orderings have a weaker tendency for dependency length minimization (DLM). For test cases, I use verb phrases in which the head verb has one direct object noun phrase (NP) dependent and exactly one adpositional phrase (PP) dependent adjacent to each other on the same side (e.g. *Kobe praised [$_{NP}$ his oldest daughter] [$_{PP}$ from the stands]*). Data from multilingual corpora of 36 languages show that when combining all these languages together, there is no consistent relationship between flexibility and DLM. When looking at specific ordering domains, on average there appears to be a weaker preference for shorter dependencies in constructions with more flexibility mostly in preverbal contexts, while no correlation exists between the two in postverbal domains.

1 Introduction

For decades, the literature on syntactic typology has used *ordering flexibility* to classify different language types (Dryer, 1997). The categorization of the flexibility profile of a full language usually relies on the relative order of main constituents such as subject, object and/or indirect object (Siewierska, 1998). The notion of flexibility has mainly been described as a categorical feature when it comes to characterizing syntactic properties of particular languages (Hale, 1982). For instance, English is considered comparatively "fixed" and predominantly SVO (Polinsky, 2012); German is deemed much more flexible than English and is cast as "non-rigid OV" (Hawkins, 1986); Japanese is treated as "rigid OV", with presumably more ordering flexibility than languages like English but less freedom than languages as German (Dryer, 2007).

Ordering flexibility should, however, be considered as a gradient parameter rather than a categorical one, and the characterizations of whether a language is "free" or "rigid" should take into consideration the particular syntactic constructions in question and the specific grammatical roles that these constructions take upon. For instance, when serving as a nominal modifier, adjectives in English regularly appear before head nouns.

(1) **reasonable** argument
(2) **good** point

Also in English, as demonstrated in the following examples, the adverbial phrase *yesterday* can appear quite freely, whether it occurs before or after its syntactic head *raised*, without yielding significant difference in the semantic interpretation of each sentence.

(3) **Yesterday** I raised a good point in the paper.
(4) I raised a good point **yesterday** in the paper.
(5) I raised a good point in the paper **yesterday**.

On the other hand, in Mandarin Chinese, a mixed-type language, when a sentence has the adverbial phrase 昨天, which also means *yesterday*, things are different. As shown in the following examples, while 昨天 has some flexibility when occurring before its syntactic head, 读了, it is rarely placed after the head verb (in written contexts).

(6) **昨天** 我 读了 一本 书
zuotian wo dule yiben shu
yesterday I read-Past one-Classifier book

(7) 我 **昨天** **读了** 一本 书
wo zuotian dule yiben shu
I **yesterday** read-Past one-Classifier book

(8) *我 读了 一本 书 **昨天**
wo dule yiben shu zuotian
I read-Past one-Classifier book **yesterday**
'Yesterday I read a book.'

The evidence for whether a language is more or less flexible traditionally comes from data constructed by grammarians or from corpora of limited

Proceedings of the Society for Computation in Linguistics (SCiL) 2021, pages 264-274.
Held on-line February 14-19, 2021

size (Siewierska, 1998). By contrast, large-scale computational investigations of flexibility and its crosslinguistic relationship with other word order properties have been lacking (Gulordava and Merlo, 2015; Futrell et al., 2015b; Levshina, 2019). Taking advantage of multilingual corpora of 36 languages from the Universal Dependencies project version 2.5 (UD) (Zeman et al., 2019), this paper explores the *gradient* relationship between word order variability and the extent of dependency length minimization (DLM), the latter of which has been claimed to be a universal principle of natural languages (Futrell et al., 2015a).

Results from recent work (Liu, 2020b; Futrell et al., 2020) have shown that the extent of DLM varies across languages. In particular, rigid head-final and mixed-type languages appear to have longer dependencies than predominantly head-initial languages. One proposed explanation is that certain languages have weaker preferences for DLM is due to them having more ordering flexibility (Gildea and Temperley, 2010). The argument goes that when there is more variability in the orders of the constituents within a syntactic construction, the ordering preferences possibly abide less by DLM and be more subject to other factors such as argumenthood status (Culicover and Jackendoff, 2005) and information status (Christianson and Ferreira, 2005).

By contrast, it is also possible that these structures take advantage of this greater ordering freedom by positioning constituents of shorter length closer to their syntactic heads in accordance with DLM. Therefore the relationship between word order freedom and the extent of shorter dependencies across languages is not yet clear. The current study takes up this question.

I use as test cases verb phrases (VP) in which the head verb has a direct object noun phrase (NP) and an adpositional phrase (PP) dependent adjacent to each other on the same side of the verb. The typological distributions of this particular syntactic construction (see §3.1) make it possible to investigate whether in general structures with more flexibility have a weaker preference for shorter dependencies across languages.

Leveraging techniques that have been widely adopted in work on quantitative typology (Cotterell et al., 2019; Elsner et al., 2019), this paper makes contributions towards the realm of ordering flexibility by showing that:

(1) there is on average no clear or consistent relationship between flexibility and DLM across languages;

(2) the preference for DLM is weaker in preverbal than postverbal domains, a contrast that is less constrained by particular language types;

(3) certain preverbal orderings with more variability have a weaker extent of DLM, while no correlation was found between the two postverbally; this contrast is different from the not-yet-verified literature prediction (Hawkins, 1983, 1994) that syntactic orders in the head-final domain are more rigid than those in the head-initial contexts.

2 Background

2.1 Word order flexibility

Recently there have been some attempts to characterize ordering variability of different languages with naturalistic corpora or behavioral data. Futrell et al. (2015b) have quantified word order freedom across 34 languages with multilingual corpora, focusing on the flexibility of a language's full word order profile. Their results corresponded well with previous grammatical descriptions of the word orders of the languages examined. Additionally, their findings offer validations to the long-standing asymmetry in typology that languages which are more flexible have more case marking, but not vice versa (Sapir, 1921; Kiparsky, 1996).

Using diachronic corpora from Latin and Ancient Greek, Gulordava and Merlo (2015) proposed different measures for the flexibility of modifiers (numerals and adjectives) and head noun within NPs. Their results confirmed that the word orders of both languages have become more fixed over time. Levshina (2019) focused on the flexibility of syntactic dependencies using multilingual data. She demonstrated that statistical measures of ordering flexibility can be incorporated into the task of language classifications better than more traditional approaches based on language types.

Using acceptability judgment tasks to measure flexibility, Namboodiripad (2019) compared the ordering variability of the six permutations of subject (S), direct object (O) and verb (V) in English and Malayalam. The results showed that the acceptability rating differences between canonical (SVO in English and SOV in Malayalam) and non-canonical orders are much more significant in English than in

Malayalam. Additionally, the experiments revealed syntactic features of the two languages that previously lack careful descriptions. For instance, based on the acceptability ratings, OVS in English is not considered as that "bad". When the verb does not occur in the sentential-final position in Malayam, SVO and OVS orders are comparably acceptable.

2.2 DLM

The Principle of DLM (Temperley and Gildea, 2018) claims that words that bear syntactic dependency relations with each other prefer to appear close to each other in order to shorten the overall or average dependencies in the language. This principle is closely related to three other theoretical principles that have been widely applied in studies on word orders and psycholinguistic processing: Early Immediate Constituents (Hawkins, 1994), Dependency Locality Theory (Gibson, 1998), and Minimize Domains (Hawkins, 2004). These principles are all motivated by the *efficiency hypothesis* (Hawkins, 2014; Gibson et al., 2019), which suggests that the typological distributions of language structures are shaped by the pressures to ease communication efficiency.

The crosslinguistic tendency for DLM is well-attested. Futrell et al. (2015a) have shown that grammars tend to minimize overall dependency lengths with evidence in 37 languages (see also earlier work by Liu (2008) for evidence from 20 languages and by Gildea and Temperley (2007) for evidence from English and German). Liu (2020b) demonstrated further, with data from 34 languages, that there is a crosslinguistic preference for constituents of shorter length to be adjacent to their syntactic heads in constructions that allow grammatical syntactic alternations.

3 Experiments

3.1 The test case

To address the gradient relationship between ordering variability and DLM, [1] it is necessary to select syntactic structures that are crosslinguistically comparable and also exhibit different extents of word order freedom across languages. To do this, first consider the following examples in English.

(9) Kobe **praised** $\left[_{NP}\right.$ his oldest daughter $\left.\right]$ $\left[_{PP}\right.$ from the stands $\left.\right]$.

[1] Code and data in quarantine at https://github.com/zoeyliu18/Flexibility-DLM

(10) Kobe **praised** $\left[_{PP}\right.$ from the stands $\left.\right]$ $\left[_{NP}\right.$ his oldest daughter $\left.\right]$.

In (9) and (10), the head verb *praised* has a direct object NP dependent and a PP oblique dependent which is prepositional, both occurring after the head verb. The two sentences are truth-conditionally equivalent, in the sense that switching the order of the NP and the PP does not change the grammaticality nor the meaning of the sentence. Between the two dependents, the direct object NP is considered to be more *argument-like* (Merlo and Ferrer, 2006) in comparison to the PP. Previous work has shown extensively for English (Jackendoff, 1977; Pollard and Sag, 1994) and for other languages as well (Tomlin, 1986), that arguments prefer to be adjacent to their syntactic heads than adjuncts, a tendency formulated as the Principle of Argument Closer. By this principle, the structure of (9) will be more preferred to that of (10).

On the other hand, Hawkins (2014) proposed the Principle of Argument Precedence as one of several interacting principles determining the relative order and flexibility of clause-level constituents. This principle states that more argument-like linguistic elements tend to occur before other dependents of the same syntactic head. The motivation lies in conceptual accessiblity (Bock and Warren, 1985): more argument-like phrases are more accessible given sentential context, and thus will occur first. As noted in Hawkins (2014), this principle provides a valid explanation for why languages with a dominant word order of (S)OVX (X representing any oblique phrases) are more prevalent, while languages with a dominant order of (S)XVO are rare: O is considered to be more argument-like than X when they are governed by the same head verb, and therefore prefers ordering precedence.

Now reconsider the English examples above. Both the NP and the PP occur postverbally. When the NP is closer to the head verb, it also precedes the PP. In this case the Principle of Argument Closer and the Principle of Argument Precedence will be cooperating with each other and make the same prediction regarding the order of the two dependents. Accordingly, in a language with postverbal NP and PP headed by the same verb, there should be mostly V-NP-PP orderings, i.e. a strong tendency for the NP to be adjacent to the verb and to be before the PP. In other words, the order of the NP and the PP in such cases will be more fixed.

However, rigid OV languages such as Japanese

have the mirror ordering pattern to that in English. When a head verb has both a direct object NP and a PP dependent on the same side, they tend to occur preverbally and the PP appears as postpositional instead of prepositional. As illustrated below, both sentences are also truth-conditionally equivalent. Again the NP is deemed as more argument-like than the PP. If the more argument-like phrase prefers proximity to the head verb, the structure of (11) will be preferred to that of (12).

(11) [$_{PP}$ 朝食 で] [$_{NP}$ パン を] 食べる
chōshoku de pan o taberu
(I) breakfast for bread -ACC eat

(12) [$_{NP}$ パン を] [$_{PP}$ 朝食 で] 食べる
pan o chōshoku de taberu
(I) bread -ACC breakfast for eat
'I eat bread for breakfast.'

Note that in contrast to English, when the NP occurs closer to the verb in Japanese, the PP precedes the NP instead. While the Principle of Argument Closer predicts a tendency for a PP-NP-V order, the Principle of Argument Precedence will opt for an NP-PP-V structure, thereby favoring (12). In this case the two principles will be competing against each other. Therefore when the language has preverbal NP and PP dependents headed by the same verb, the preference for the NP to be adjacent to the head verb will be weaker. In comparison to when the two dependents occur postverbally as in English, this means that the relative order of preverbal NP and PP will be less fixed and more flexible.

The possibly different ordering flexibility between the NP and the PP across languages laid out here points to a suitable test case to examine the gradient relationship between flexibility and DLM, i.e. whether syntactic constructions with more ordering freedom have a weaker preference for shorter dependencies. In particular, it enables quantitative validations of the potentially contrasting word order freedom and DLM preferences between postverbal and preverbal domains.

3.2 Data and preprocessing

I used multilingual corpora of contemporary languages from UD and extracted VPs in which the head verb has one direct object NP and exactly one PP oblique right next to each other on the same side (e.g. *Kobe praised [$_{NP}$ his oldest daughter] [$_{PP}$ from the stands]*) from every treebank.

For all instances extracted from each treebank, the proportion of when the direct object NP appears closer to the head verb as well as the proportion of when the PP occurs closer were calculated. As the difference in proportion between treebanks for the same language was small, treebanks for the same language were combined. Languages with fewer than a total of 100 extracted instances were not included. After preprocessing, the data set contained 36 languages in total. Among these languages, twenty-nine ended up being Indo-European, except for Arabic and Hebrew (Afro-Asiatic), Indonesian (Austronesian), Wolof (Niger-Congo), Finish and Estonian (Uralic), and Japanese (Japanese).

3.3 Measures

3.3.1 Measures for ordering flexibility

To approximate the ordering flexibility of the NP and the PP within each VP instance, I used entropy, a measure of dispersion that reflects the amount of variability within a distribution. Previous work has applied entropy to measure word order freedom at different levels (Futrell et al., 2015b; Levshina, 2019). Futrell et al. (2015b) adopted conditional entropy to compute word order freedom for a language as a whole. Levshina (2019) used entropy to measure ordering flexibility of particular head-dependent pairs, such as the ordering freedom of determiner and noun.

As shown in Eq 1, X is a binary variable that denotes the relative order of the NP and the PP (e.g. V-NP-PP vs. V-PP-NP; PP-NP-V vs. NP-PP-V); $P(x_i)$ represents the probability of one particular order. For example, consider a language that has 100 instances where the NP and the PP appear postverbally. Within these instances, 80 appear as V-NP-PP while the other 20 appear as V-PP-NP. The probability of the former is then 0.8, and the probability of the latter is 0.2. The ordering freedom for when the NP and the PP are postverbal in this case is then computed as: $-(0.8 * log_2 0.8 + 0.2 * log_2 0.2) = 0.72$.

$$H(X) = -\sum_{i=1}^{2} P(x_i) log_2 P(x_i) \qquad (1)$$

The value of entropy ranges from 0 to 1, with 0 indicating there is no flexibility in the order of the two dependents and 1 corresponding to the maximum amount of flexibility. The higher the entropy value is, the more flexible the ordering is.

95% confidence intervals for ordering flexibility were estimated with bootstrapping (Efron, 1979) for 1,000,000 iterations. The basic procedures for

computation were as follows: (1) given a data set of n VP instances in total, a random sample of n instances was drawn with replacement; (2) within the sample, the ordering flexibility of the NP and the PP was approximated with entropy; (3) step (1) and (2) were repeated for 1,000,000 iterations, which resulted in a resampling distribution of the entropy value from each iteration; (4) the mean and 95% confidence interval of the distribution derived from step (3) were computed.

3.3.2 Measures for DLM

While dependency length has been shown to be one of the most effective predictor in crosslinguistic syntactic alternations (Liu, 2020a), it is not the sole predictor. Therefore to more accurately measure the extent of DLM in the constructions studied here within each language, mixed-effect logistic regression models were adopted. More specifically, within each ordering pattern of every language (see §4), half of the original instances were randomly selected and their structural variants were constructed by exchanging the order of the NP and the PP. The remaining half were kept the way they were.

Factor	1	0	-1
dependency length	shorter phrase closer	NP and PP have same length	longer phrase closer
argument status	NP closer		PP closer
pronominality	pronominal phrase closer	NP and PP have same pronominality	nonpronominal phrase closer

Table 1: Coding for fixed effects in regression models.

The outcome binary variable was the order of the NP and the PP, $Order$, which was coded as 1 for all original instances and 0 for all variants. As presented in Table 1, three automatically measurable factors were included in the models as fixed effects: dependency length, argument status (i.e. whether the NP is closer to the verb) and pronominality, along with an interaction between dependency length and pronominality in cases where the interaction turned out to have a significant role. The head verb of each VP instance was included as a random effect. Regarding dependency length, the length of every phrase was measured as the number of tokens within the phrase based on the treebank annotations; then within each VP instance, the length of the NP and that of the PP were compared to see whether the phrase of shorter length occurs closer to the verb.

Every model was trained to predict the original ordering. Estimates for 95% confidence intervals of the coefficient of each factor were derived from 10,000-fold cross-validations. The coefficient value of dependency length was then taken as an approximate for the extent of DLM. A positive coefficient that is significantly larger than 0 indicates a tendency for the shorter constituent to be closer to the head verb (regardless of whether it is the NP or the PP), and thus corresponds to minimization of overall dependency length.

To further evaluate the differences between the estimated coefficient values of each language and their ordering patterns (Gelman and Stern, 2006), post-hoc pairwise comparison was performed via adding interaction terms between each of the three fixed factors and the particular ordering pattern (e.g. Wolof postverbal vs. German preverbal).

$$Order \sim (dependencylength + argumenthood + pronominality) * ordering\ pattern$$

3.3.3 Measures for correlation between ordering flexibility and DLM

The correlation between ordering flexibility and the extent of DLM was computed with Bayesian linear regression, predicting the extent of DLM as a function of ordering flexibility ($DLM \sim flexibility$). For languages with orders in both preverbal and postverbal domains, given that the extent of DLM is different comparing the two domains for most of these languages (see §4.2), an additional model was fit including the domain as a fixed effect as well as interaction with flexibility ($DLM \sim flexibility * domain$). Similar to the pairwise comparison of DLM described above, the interaction term here is to see whether the role of flexibility differs in the two domains.

A weakly informative prior was adopted for each parameter in the model (Ghosh et al., 2018), which is helpful to set reasonable upper and lower bounds for values within the posterior distributions (Levshina, 2018). The prior followed a Student's t distribution centered around 0 (mean $\mu = 0$), with $\nu = 3$ degrees of freedom and scale $\sigma = 8$. The default Markov Chain Monte Carlo sampling methods implemented in STAN was employed using the R package $brms$ (Bürkner et al., 2017). Each model ran 2 chains with 3000 iterations in each chain, with the first 500 iterations as burn-in samples. 95% confidence intervals for each model

parameter were derived from their respective posterior distributions. A coefficient value significantly smaller/larger than 0 indicates a negative/positive correlation.

4 Results

As this study is taking a data-driven approach, the languages were grouped based on the observed orderings of the NP and the PP from corpora, rather than based on traditional language families. Overall the data set exhibits three ordering patterns:

(1) one for languages with NP and PP only after the head verb (e.g. Wolof; Finnish);

(2) one for languages with NP and PP both after and before the head verb (e.g. Czech; Estonian);

(3) one for languages with NP and PP only before the head verb (e.g. Persian; Hindi).

4.1 Languages with postverbal orderings

Of all the languages, twenty have the NP and the PP dependents both occurring after the head verb. As shown in Figure 1, most of these languages are Germanic or Romance languages. In particular, the four Germanic languages, Danish, English, Norwegian and Swedish all have comparatively less flexibility [2], in contrast to the Romance languages. The PPs appear as prepositional in 19 languages, and postpositional in Finnish. Nevertheless, ordering flexibility does not seem to depend on this difference. The word order freedom in Finnish (0.85) is comparable to languages that are prepositional, such as Galician (0.86) and Greek (0.82).

All of these languages exhibit a significant tendency for DLM. On the other hand, there is no significant correlation between ordering flexibility and DLM for these languages alone, based on Bayesian regression ($\beta = 0.08\ (-0.51, 0..68)$). This means that among these languages, when they have more ordering flexibility, they do not tend to have a weaker extent of shorter dependencies.

4.2 Languages with both postverbal and preverbal orderings

Now consider languages with mixed ordering patterns, that is, languages in which the NP and the PP can appear both after and before the verb, as in Figure 2. These languages mainly fall into three subtypes: Germanic, Romance and Slavic. Overall most of these languages demonstrate great variability in the order of the two dependents in postverbal as well as in preverbal domains. On average, the ordering flexibility of the two dependents when they are postverbal is similar to when they are preverbal. Among these languages, the PPs appear as postpositional in Estonian, and prepositional in the others. However, this does not seem to affect ordering freedom either. The flexibility in Estonian is comparable to that in languages such as Dutch or German in both preverbal and postverbal domains.

On the other hand, in comparison to languages where the two dependents only occur after the verb (Figure 1), it is not obvious whether these languages are more flexible; although the latter have a higher average value of flexibility in both the preverbal and postverbal domains, the average for languages from Figure 1 is influenced by quite a number of languages (e.g. Indonesian, Serbian) with much less flexibility.

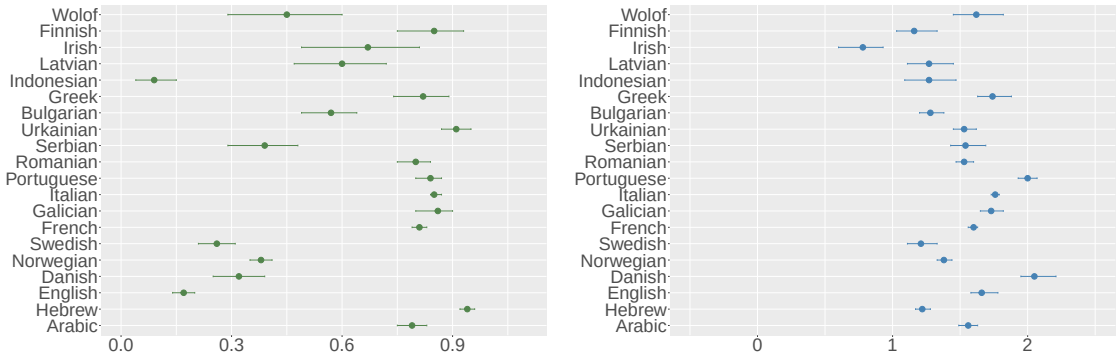
Coupled with results from pairwise comparisons, for most of these languages except Spanish, Croatian and Catalan, the extent of shorter dependencies is significantly different in preverbal and postverbal orderings, with the former being much weaker than the latter (Figure 2).

For preverbal orderings, results from Bayesian regression show a strong negative correlation between ordering flexibility and the extent of DLM ($\beta = -1.85\ (-2.57, -1.11)$). By contrast, the correlation does not hold for postverbal contexts ($\beta = -1.02\ (-2.37, 0.39)$). And when combining the two domains together, there does not appear to be a positive effect of flexibility on DLM ($\beta = -1.01\ (-2.15, 0.19)$) neither. These numbers suggest that the relationship between flexibility and DLM relies on the particular ordering structure in question.

4.3 Languages with preverbal orderings

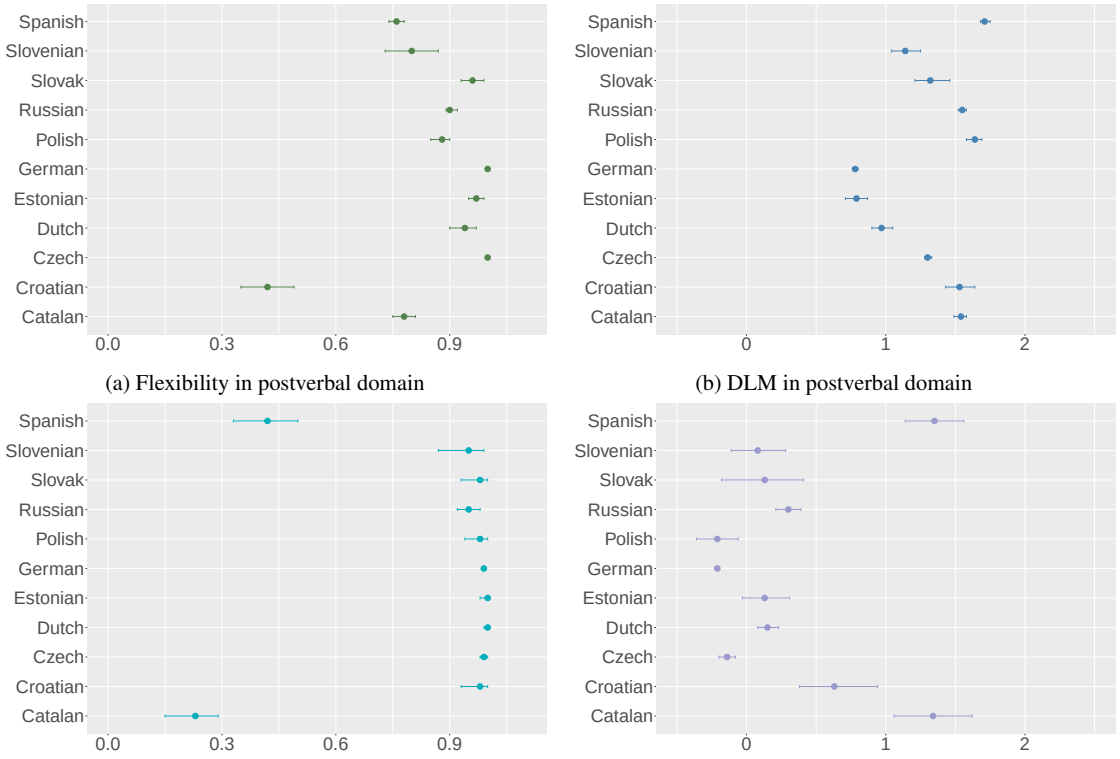
In five languages from the data set, the NP and the PP both occur before the head verb when they are on the same side. As shown in Figure 3, on average the order of the NP and the PP appears to be more flexible than languages with only postverbal orders in Figure 1, and is comparable to languages with mixed orderings in Figure 2. In the three rigid OV languages, including Japanese and the two In-

[2] In this paper, the flexibility of the languages to be analyzed is discussed in the contexts of the particular constructions investigated here.



(a) Flexibility (b) DLM

Figure 1: Ordering flexibility and extent of DLM in languages with postverbal NP and PP. Average ordering flexibility is 0.62 and average DLM is 1.49. Error bars represent 95% confidence intervals.



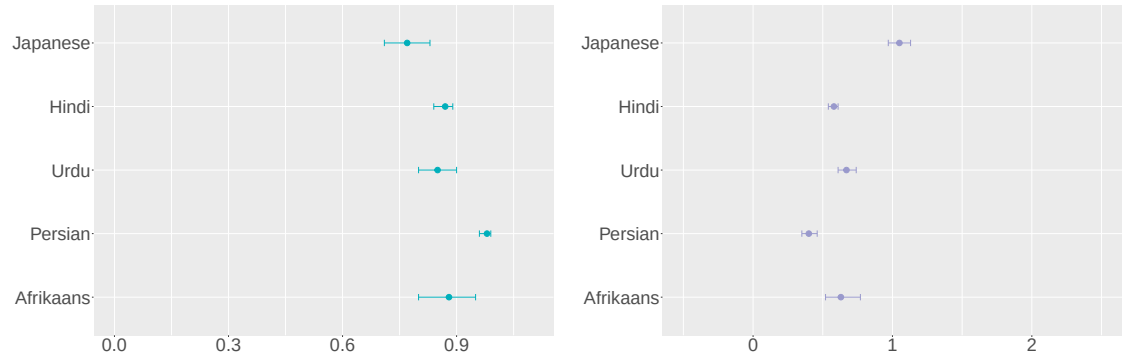
(a) Flexibility in postverbal domain (b) DLM in postverbal domain

(c) Flexibility in preverbal domain (d) DLM in preverbal domain

Figure 2: Ordering flexibility and extent of DLM in languages with mixed orderings. In postverbal domains, average ordering flexibility is 0.86 and average DLM is 1.30. In preverbal domains, average ordering flexibility is 0.85 and average DLM is 0.32. Error bars represent 95% confidence intervals.

dic languages, Urdu and Hindi, the PPs appear as postpositional; whereas in the two non-rigid OV languages, Afrikaans and Persian, the PPs appear as prepositional. Yet there does not seem to be a strong discrepancy in ordering freedom concerning whether the PP is prepositional or postpositional either. The flexibility for the rigid head-final languages and that for Afrikaans is comparable to each other. In contrast, Persian seems to have much more ordering freedom.

On the other hand, the effect of dependency length is robust yet also much weaker for these languages than when the NP and the PP are postverbal in general (Figure 1, Figure 2b), indicating the extent of DLM is stronger in head-initial than head-final domains. While no correlation was



(a) Flexibility (b) DLM

Figure 3: Ordering flexibility and extent of DLM in languages with preverbal NP and PP. Average ordering flexibility is 0.87 and average DLM is 0.67. Error bars represent 95% confidence intervals.

found between flexibility and DLM for this group ($\beta = -2.93$ $(-5.66, 0.15)$), given its relatively small size, statements drawn from these languages alone are not conclusive.

4.4 An overall look at all languages

In order to take all languages into account, a Bayesian linear regression was fit to predict DLM using flexibility, the particular ordering domain and their interaction. Crosslinguistically, the extent of DLM does not appear to grow weaker as ordering flexibility increases ($\beta = -0.24$ $(-0.68, 0.20)$).

Combining all languages, there is a stronger preference for shorter dependencies when the NP and the PP are postverbal than preverbal. The distinction is less dependent on the specific language types and is most obvious in languages with mixed ordering patterns. This lends support to recent findings (Futrell et al., 2020; Liu, 2020b) that the extent of DLM is weaker in head-final contexts.

5 Discussion

Using dependency corpora from 36 languages, this paper has demonstrated that on average there is no apparently consistent relationship between flexibility and DLM. With that being said, certain ordering structures (e.g. preverbal domain in languages with mixed ordering patterns) with more ordering freedom do have a weaker preference for shorter dependencies, while such correlation was not found for any of the cases in the postverbal domain. This draws contrast to the prediction from previous literature (Hawkins, 1983), which is not yet to be empirical validated, that syntactic orderings in the head-final contexts would be more rigid than those in the head-initial. This prediction does not seem to hold here, at least in the structures that have been examined in the current study.

An alternative view of the results presented is that flexibility largely has a significant effect on DLM in head-final domains, and no effect in head-initial ones. This potentially provides an empirical account for the crosslinguistic variation of DLM (Futrell et al., 2020; Liu, 2020b), besides other traditional competing and cooperating principles on syntactic orderings (Hawkins, 2014).

There are a few discrepancies between the current findings and two previous quantitative studies on flexibility, Futrell et al. (2015b) and Levshina (2019). First, although most of the languages with flexible orders examined here have rich case-marking or at least the nominative-accusative case marking (ergative-absolutive for Hindi when the head verb is in past tense) (Dryer and Haspelmath, 2013), there are exceptions such as Hebrew and Portuguese, both of which have high flexibility yet no case-marking system.

Secondly, while Spanish and Catalan have been shown to be fairly flexible based on aggregated entropy measures, the results here indicate that ordering flexibility for these two languages is different depending on whether the constituents are preverbal or postverbal.

Thirdly, both rigid and non-rigid OV languages appear to have quite flexible orderings, a pattern on the contrary to results from Futrell et al. (2015b) and Levshina (2019). These discrepancies speak to the arguments laid out at the beginning of this paper, that ordering flexibility should be studied with the properties of particular syntactic structures in mind. The positions of different constituents within the structures should also be considered to not leave

out interesting quantitative details of the languages.

Although pronominality is not the focus in this study, there is a strong ordering effect of pronominality across languages. The preference for the pronominal phrase to occur first, whether it is the NP or the PP, appears to hold both after and before the head verb. Following the general pattern that pronominal elements tend to be more "given" than nonpronominal items, the preferences demonstrated here support the prediction made by the Principle of Given-before-new (Bock, 1986).

Besides DLM, future work could probe the crosslinguistic relationship between ordering variability and other word order parameters such as topicality. In addition, it would be worthwhile to see whether results from behavioral experiments addressing the same questions yield similar findings. For example, the methodology of acceptability judgment tasks from Namboodiripad (2019) can be extended to syntactic constructions on a larger crosslinguistic scale. Typological studies as such are crucial for painting a more clear picture of the syntactic structures of the languages in question, as they will not only provide evidence to validate previously held grammatical descriptions or assumptions, but also possibly reveal features of the languages that have not been documented or considered theoretically before from the angle of flexibility (Namboodiripad, 2019).

References

J. Kathryn Bock. 1986. Syntactic persistence in language production. *Cognitive Psychology*, 18(3):355 – 387.

J. Kathryn Bock and Richard K. Warren. 1985. Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1):47 – 67.

Paul-Christian Bürkner et al. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28.

Kiel Christianson and Fernanda Ferreira. 2005. Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition*, 98(2):105–135.

Ryan Cotterell, Christo Kirov, Mans Hulden, and Jason Eisner. 2019. On the complexity and typology of inflectional morphological systems. *Transactions of the Association for Computational Linguistics*, 7:327–342.

Peter W. Culicover and Ray Jackendoff. 2005. *Simpler Syntax*. Oxford University Press, Oxford.

Matthew S. Dryer. 1997. On the six-way word order typology. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, 21(1):69–103.

Matthew S. Dryer. 2007. Word order. *Language typology and syntactic description*, 1:61–131.

Matthew S. Dryer and Martin Haspelmath, editors. 2013. *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.

Bradley Efron. 1979. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1–26.

Micha Elsner, Andrea D Sims, Alexander Erdmann, Antonio Hernandez, Evan Jaffe, Lifeng Jin, Martha Booker Johnson, Shuan Karim, David L King, Luana Lamberti Nunes, et al. 2019. Modeling morphological learning, typology, and change: What can the neural sequence-to-sequence framework contribute? *Journal of Language Modelling*, 7(1):53–98.

Richard Futrell, Roger P. Levy, and Edward Gibson. 2020. Dependency locality as an explanatory principle for word order. *Language*, 96(2):371–412.

Richard Futrell, Kyle Mahowald, and Edward Gibson. 2015a. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, 112(33):10336–10341.

Richard Futrell, Kyle Mahowald, and Edward Gibson. 2015b. Quantifying word order freedom in dependency corpora. In *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)*, pages 91–100, Uppsala, Sweden. Uppsala University, Uppsala, Sweden.

Andrew Gelman and Hal Stern. 2006. The difference between "significant" and "not significant" is not itself statistically significant. *The American Statistician*, 60(4):328–331.

Joyee Ghosh, Yingbo Li, Robin Mitra, et al. 2018. On the use of Cauchy prior distributions for Bayesian logistic regression. *Bayesian Analysis*, 13(2):359–383.

Edward Gibson. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68(1):1 – 76.

Edward Gibson, Richard Futrell, Steven Piantadosi, Isabelle Dautriche, Kyle Mahowald, Leon Bergen, and Roger Levy. 2019. How Efficiency Shapes Human Language. *Trends in Cognitive Sciences*, 23(5):389–407.

Daniel Gildea and David Temperley. 2007. Optimizing grammars for minimum dependency length. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 184–191,

Prague, Czech Republic. Association for Computational Linguistics.

Daniel Gildea and David Temperley. 2010. Do Grammars Minimize Dependency Length? *Cognitive Science*, 34(2):286–310.

Kristina Gulordava and Paola Merlo. 2015. Diachronic trends in word order freedom and dependency length in dependency-annotated corpora of Latin and ancient Greek. In *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)*, pages 121–130, Uppsala, Sweden. Uppsala University, Uppsala, Sweden.

Ken Hale. 1982. Preliminary remarks on configurationality. In *Proceedings of the North Eastern Linguistic Society*, volume 12, pages 86–96.

John A. Hawkins. 1983. *Word Order Universals*. New York: Academic Press.

John A. Hawkins. 1986. *A Comparative Typology of English and German: Unifying the contrasts*. London: Croom Helm, and Austin: University of Texas Press.

John A. Hawkins. 1994. *A Performance Theory of Order and Constituency*, volume 73. Cambridge University Press, Cambridge.

John A. Hawkins. 2004. *Efficiency and Complexity in Grammars*. Oxford University Press, Oxford.

John A. Hawkins. 2014. *Cross-linguistic Variation and Efficiency*. Oxford University Press, Oxford.

Ray Jackendoff. 1977. *X-bar Syntax: A Study of Phrase Structure*. Linguistic inquiry monographs 2. Cambridge, Massachusetts: MIT Press.

Paul Kiparsky. 1996. The shift to head-initial VP in Germanic. In Höskuldur Thráinsson, Samuel David Epstein, and Steve Peter, editors, *Studies in Comparative Germanic Syntax*, volume 2, pages 140–179. Kluwer Dordrecht.

Natalia Levshina. 2018. Probabilistic grammar and constructional predictability: Bayesian generalized additive models of help+(to) infinitive in varieties of web-based english. *Glossa: a journal of general linguistics*, 3(1).

Natalia Levshina. 2019. Token-based typology and word order entropy: A study based on Universal Dependencies. *Linguistic Typology*, 23(3):533–572.

Haitao Liu. 2008. Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science*, 9(2):159–191.

Zoey Liu. 2020a. *Data-driven Crosslinguistic Modeling of Constituent Ordering Preferences*. Ph.D. thesis, University of California, Davis.

Zoey Liu. 2020b. Mixed evidence for crosslinguistic dependency length minimization. *STUF-Language Typology and Universals*, 73(4):605–633.

Paola Merlo and Eva Esteve Ferrer. 2006. The Notion of Argument in Prepositional Phrase Attachment. *Computational Linguistics*, 32(3):341–378.

Savithry Namboodiripad. 2019. A gradient approach to flexible constituent order.

Maria Polinsky. 2012. Headedness, again. *Theories of Everything. In Honor of Ed Keenan*.

Carl Pollard and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. University of Chicago Press.

Edward Sapir. 1921. Language: An introduction to the study of speech. *NY: Harcourt, Brace & Co.*

Anna Siewierska, editor. 1998. *Constituent Order in the Languages of Europe*. Walter de Gruyter Co.

David Temperley and Daniel Gildea. 2018. Minimizing Syntactic Dependency Lengths: Typological/Cognitive Universal? *Annual Review of Linguistics*, 4(1):67–80.

Russell S. Tomlin. 1986. *Basic Word Order: Functional Principles*. London: Croom Helm.

Daniel Zeman, Joakim Nivre, Mitchell Abrams, Noëmi Aepli, Željko Agić, Lars Ahrenberg, Gabrielė Aleksandravičiūtė, Lene Antonsen, Katya Aplonova, Maria Jesus Aranzabe, Gashaw Arutie, Masayuki Asahara, Luma Ateyah, Mohammed Attia, Aitziber Atutxa, Liesbeth Augustinus, Elena Badmaeva, Miguel Ballesteros, Esha Banerjee, Sebastian Bank, Verginica Barbu Mititelu, Victoria Basmov, Colin Batchelor, John Bauer, Sandra Bellato, Kepa Bengoetxea, Yevgeni Berzak, Irshad Ahmad Bhat, Riyaz Ahmad Bhat, Erica Biagetti, Eckhard Bick, Agnė Bielinskienė, Rogier Blokland, Victoria Bobicev, Loïc Boizou, Emanuel Borges Völker, Carl Börstell, Cristina Bosco, Gosse Bouma, Sam Bowman, Adriane Boyd, Kristina Brokaitė, Aljoscha Burchardt, Marie Candito, Bernard Caron, Gauthier Caron, Tatiana Cavalcanti, Gülşen Cebiroğlu Eryiğit, Flavio Massimiliano Cecchini, Giuseppe G. A. Celano, Slavomír Čéplö, Savas Cetin, Fabricio Chalub, Jinho Choi, Yongseok Cho, Jayeol Chun, Alessandra T. Cignarella, Silvie Cinková, Aurélie Collomb, Çağrı Çöltekin, Miriam Connor, Marine Courtin, Elizabeth Davidson, Marie-Catherine de Marneffe, Valeria de Paiva, Elvis de Souza, Arantza Diaz de Ilarraza, Carly Dickerson, Bamba Dione, Peter Dirix, Kaja Dobrovoljc, Timothy Dozat, Kira Droganova, Puneet Dwivedi, Hanne Eckhoff, Marhaba Eli, Ali Elkahky, Binyam Ephrem, Olga Erina, Tomaž Erjavec, Aline Etienne, Wograine Evelyn, Richárd Farkas, Hector Fernandez Alcalde, Jennifer Foster, Cláudia Freitas, Kazunori Fujita, Katarína Gajdošová, Daniel Galbraith, Marcos Garcia, Moa Gärdenfors, Sebastian Garza, Kim Gerdes, Filip Ginter, Iakes

Goenaga, Koldo Gojenola, Memduh Gökırmak, Yoav Goldberg, Xavier Gómez Guinovart, Berta González Saavedra, Bernadeta Griciūtė, Matias Grioni, Normunds Grūzītis, Bruno Guillaume, Céline Guillot-Barbance, Nizar Habash, Jan Hajič, Jan Hajič jr., Mika Hämäläinen, Linh Hà Mỹ, Na-Rae Han, Kim Harris, Dag Haug, Johannes Heinecke, Felix Hennig, Barbora Hladká, Jaroslava Hlaváčová, Florinel Hociung, Petter Hohle, Jena Hwang, Takumi Ikeda, Radu Ion, Elena Irimia, Ọlájídé Ishola, Tomáš Jelínek, Anders Johannsen, Fredrik Jørgensen, Markus Juutinen, Hüner Kaşıkara, Andre Kaasen, Nadezhda Kabaeva, Sylvain Kahane, Hiroshi Kanayama, Jenna Kanerva, Boris Katz, Tolga Kayadelen, Jessica Kenney, Václava Kettnerová, Jesse Kirchner, Elena Klementieva, Arne Köhn, Kamil Kopacewicz, Natalia Kotsyba, Jolanta Kovalevskaitė, Simon Krek, Sookyoung Kwak, Veronika Laippala, Lorenzo Lambertino, Lucia Lam, Tatiana Lando, Septina Dian Larasati, Alexei Lavrentiev, John Lee, Phng Lê H`ông, Alessandro Lenci, Saran Lertpradit, Herman Leung, Cheuk Ying Li, Josie Li, Keying Li, KyungTae Lim, Maria Liovina, Yuan Li, Nikola Ljubešić, Olga Loginova, Olga Lyashevskaya, Teresa Lynn, Vivien Macketanz, Aibek Makazhanov, Michael Mandl, Christopher Manning, Ruli Manurung, Cătălina Mărănduc, David Mareček, Katrin Marheinecke, Héctor Martínez Alonso, André Martins, Jan Mašek, Yuji Matsumoto, Ryan McDonald, Sarah McGuinness, Gustavo Mendonça, Niko Miekka, Margarita Misirpashayeva, Anna Missilä, Cătălin Mititelu, Maria Mitrofan, Yusuke Miyao, Simonetta Montemagni, Amir More, Laura Moreno Romero, Keiko Sophie Mori, Tomohiko Morioka, Shinsuke Mori, Shigeki Moro, Bjartur Mortensen, Bohdan Moskalevskyi, Kadri Muischnek, Robert Munro, Yugo Murawaki, Kaili Müürisep, Pinkey Nainwani, Juan Ignacio Navarro Horñiacek, Anna Nedoluzhko, Gunta Nešpore-Bērzkalne, Lng Nguyˆên Thị, Huy`ên Nguyˆên Thị Minh, Yoshihiro Nikaido, Vitaly Nikolaev, Rattima Nitisaroj, Hanna Nurmi, Stina Ojala, Atul Kr. Ojha, Adédayọ̀ Olúòkun, Mai Omura, Petya Osenova, Robert Östling, Lilja Øvrelid, Niko Partanen, Elena Pascual, Marco Passarotti, Agnieszka Patejuk, Guilherme Paulino-Passos, Angelika Peljak-Łapińska, Siyao Peng, Cenel-Augusto Perez, Guy Perrier, Daria Petrova, Slav Petrov, Jason Phelan, Jussi Piitulainen, Tommi A Pirinen, Emily Pitler, Barbara Plank, Thierry Poibeau, Larisa Ponomareva, Martin Popel, Lauma Pretkalniņa, Sophie Prévost, Prokopis Prokopidis, Adam Przepiórkowski, Tiina Puolakainen, Sampo Pyysalo, Peng Qi, Andriela Rääbis, Alexandre Rademaker, Loganathan Ramasamy, Taraka Rama, Carlos Ramisch, Vinit Ravishankar, Livy Real, Siva Reddy, Georg Rehm, Ivan Riabov, Michael Rießler, Erika Rimkutė, Larissa Rinaldi, Laura Rituma, Luisa Rocha, Mykhailo Romanenko, Rudolf Rosa, Davide Rovati, Valentin Roșca, Olga Rudina, Jack Rueter, Shoval Sadde, Benoît Sagot, Shadi Saleh, Alessio Salomoni, Tanja Samardžić, Stephanie Samson, Manuela Sanguinetti, Dage Särg, Baiba Saulīte, Yanin Sawanakunanon, Nathan Schneider, Sebastian Schuster, Djamé Seddah, Wolfgang Seeker, Mojgan Seraji, Mo Shen, Atsuko Shimada, Hiroyuki Shirasu, Muh Shohibussirri, Dmitry Sichinava, Aline Silveira, Natalia Silveira, Maria Simi, Radu Simionescu, Katalin Simkó, Mária Šimková, Kiril Simov, Aaron Smith, Isabela Soares-Bastos, Carolyn Spadine, Antonio Stella, Milan Straka, Jana Strnadová, Alane Suhr, Umut Sulubacak, Shingo Suzuki, Zsolt Szántó, Dima Taji, Yuta Takahashi, Fabio Tamburini, Takaaki Tanaka, Isabelle Tellier, Guillaume Thomas, Liisi Torga, Trond Trosterud, Anna Trukhina, Reut Tsarfaty, Francis Tyers, Sumire Uematsu, Zdeňka Urešová, Larraitz Uria, Hans Uszkoreit, Andrius Utka, Sowmya Vajjala, Daniel van Niekerk, Gertjan van Noord, Viktor Varga, Eric Villemonte de la Clergerie, Veronika Vincze, Lars Wallin, Abigail Walsh, Jing Xian Wang, Jonathan North Washington, Maximilan Wendt, Seyi Williams, Mats Wirén, Christian Wittern, Tsegay Woldemariam, Tak-sum Wong, Alina Wróblewska, Mary Yako, Naoki Yamazaki, Chunxiao Yan, Koichi Yasuoka, Marat M. Yavrumyan, Zhuoran Yu, Zdeněk Žabokrtský, Amir Zeldes, Manying Zhang, and Hanzhi Zhu. 2019. Universal Dependencies 2.5. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.