Mixed Evidence for Crosslinguistic Dependency Length Minimization

Zoey Liu
yiliu@ucdavis.edu
Department of Linguistics
University of California, Davis

Abstract

We investigate whether and to what extent the principle of Dependency Length Minimization (DLM) predicts crosslinguistic syntactic ordering preferences. More specifically, we ask: (i) is there a typological tendency for shorter constituents to appear closer to their syntactic heads in constructions with flexible constituent orderings? (ii) how does the extent of DLM in these constructions vary for languages with different structural characteristics? Our study uses prepositional and postpositional phrase (PP) typology as a testbed. Leveraging multilingual corpora for 34 languages, we focus on sentences with verb phrases that have exactly two PP dependents on the same side of the head verb, the ordering of which under certain conditions contains flexibility. Overall we show a pronounced preference for shorter PPs to be closer to the head verb, establishing the first large-scale quantitative evidence that DLM exists in crosslinguistic syntactic alternations. Furthermore, we present evidence that while the efficacy of DLM depends on the specific ordering structures of different language types, across languages there appears to be a much stronger preference for DLM when the two PPs appear postverbally, compared to no or a much weaker tendency for shorter dependencies when the two PPs occur preverbally. This contrast is the most visible in mixed-type languages with head-initial PPs that can appear both after or before the head verb. Within the limited number of rigid OV languages in our dataset, which have head-final PPs before the head verb, we observe no robust tendency for DLM, in contrast to the patterns in languages with head-initial PPs after the head verb. This contradicts previous findings of a longer-before-short preference in preverbal orders of head-final languages.

keywords: PP order typology; multilingual corpora; language processing; computational measures

1 Introduction

Languages differ appreciably yet share considerable similarities across a wide range of structures. The characterization and quantification of crosslinguistic regularities and differences cannot only further our fundamental understanding of languages, but also have practical implications in the development of language technologies. The availability of multilingual corpora annotated with syntactic information allows for investigation of issues in linguistic typology using computational methods (Merlo 2015, Wang and Eisner 2017, Östling 2015), shedding new light on the underlying mechanisms behind structural patterns in natural languages. In addition, typological information can help improve the performance of multilingual natural language processing systems (Bender 2009, Ponti et al. 2019) and facilitate extending current computational techniques and applications from resource-rich languages to low-resource or understudied languages (O'Horan et al. 2016).

Previous experiments have claimed a universal principle of human languages, namely Dependency Length Minimization (DLM) (Hawkins 1994, Gibson 1998, 2000, Wasow and Arnold 2003, Yamashita 2002, Temperley 2007). Corpus studies have evidenced that grammars tend to minimize the overall or average distance between syntactic heads and their dependents, shown across 20 languages in Liu (2008) and 37 languages in Futrell et al. (2015a). Other work has tested the predictions of DLM in choosing among multiple possible orders for distinct syntactic structures in specific languages. They have shown that when grammatical alternatives exist for the construction, there is a preference for constituents of shorter length to occur closer to their syntactic heads, yielding shorter overall dependency length in the sentence (Jaeger and Norcliffe 2009, Temperley and Gildea 2018, Dyer 2017).

As fruitful as previous findings are, two questions regarding DLM have not been addressed. First, are shorter dependencies preferred in syntactic constructions with flexible constituent orderings across languages? In other words, it is unclear whether there is a typological preference for DLM in sentences with syntactic alternations, when language users presumably have a choice towards the relative ordering of the constituents. Previous results are not directly comparable as they examined mostly different structures and the scope of the studies is limited to one or a few languages (Wasow and Arnold 2003, Yamashita 2002, Gulordava et al. 2015, Rajkumar et al. 2016).

Second, how does the efficacy of DLM in ordering preferences vary among languages with different syntactic characteristics? Very few studies in the literature have tried to tackle this issue (Futrell et al. 2015a, Gildea and Temperley 2010, Liu 2010, Ros 2018, Liu 2019). In the corpus experiment by Gildea and Temperley (2010), the grammar of German appeared to minimize dependency length to a lesser extent than that of English. Results from Futrell et al. (2015a) suggested that head-final

languages such as Japanese and Turkish have longer dependencies than head-initial languages like Italian and Indonesian, though without explicit statistical tests. Liu (2008) demonstrated that among the 20 languages examined, Mandarin Chinese has the longest mean dependency length. A close examination of double prepositional phrase orderings in English and Mandarin Chinese (Liu 2019) also found that Mandarin Chinese has a weaker tendency for shorter dependencies. Looking at transitive and ditransitive constructions in Basque, Polish and Spanish, Ros (2018) showed that Basque, an OV language with a mix of verb-medial and verb-final orders, has weaker DLM effects compared to the other two languages.

Our work empirically probes the aforementioned questions, using prepositional and postpositional phrase (PP) typology as a testbed. We focus specifically on sentences with verb phrases (VP) that have exactly two PP dependents appearing on the same side of the head verb, the ordering of which under some context permits flexibility (e.g. *She danced with the band at the dinner party* vs. *She danced at the dinner party with the band*). These cases allow language users to presumably have a choice regarding the order of the two PPs . Going beyond previous studies on syntactic ordering preferences, including experiments on PP order which have mainly looked at English (Hawkins 1999, Wiechmann and Lohmann 2013), we leverage multilingual corpora data for 34 languages from the Universal Dependencies project (Zeman et al. 2019). We address:

(1) whether there is a crosslinguistic tendency for the shorter PP to appear closer to the head verb;

(2) how the efficacy of DLM varies in different PP ordering patterns across languages with different word order features.

The remainder of the paper is structured as follows: in Section 2 we review related work on the effects of dependency length in syntactic alternations; Sections 3 and 4 give detailed descriptions of our methods and results; we discuss the implications of our study in Section 5 and conclude with Section 6.

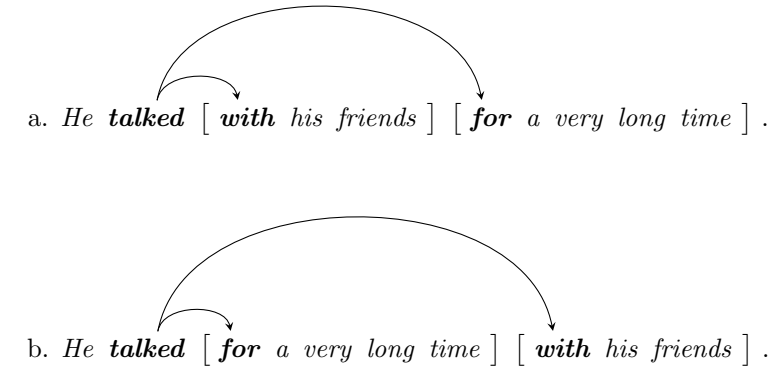
2 Related Work

Besides DLM, the role of dependency length in constituent orderings has been addressed in different linguistic principles in the literature, including Early Immediate Constituents (Hawkins 1994), Minimize Domains (Hawkins 2004) and Dependency Locality Theory (Gibson 1998). While DLM was proposed based on a dependency grammar framework, most other principles were built upon a phrase structure framework. Though they have not used the term *dependency length* directly, the notion can be easily derived. These principles all suggest the same core idea concerning syntactic ordering preferences, that shorter dependencies are preferred in

order to ease processing and efficient communication (Jaeger and Tily 2011, Gibson 2000, Hawkins 2014).

As an illustration of how DLM applies to constituent orderings, let us consider the following two sentences in English:

(1)



Here both (1a) and (1b) have two PPs, *with his friends* and *for a very long time*, shown within square brackets. Switching the order of the two PPs results in alternatives that are both grammatical and convey the same meaning. We opt for function head over content head here for ease of demonstration for DLM and treat the prepositions in both PPs, *with* and *for*, to be the heads of their respective constituents.[1] As indicated by the syntactic dependency arcs, we consider the two prepositions as dependents of the verb *talked*, the head of the VP in each sentence. The dependency length between each PP dependent and its head verb is the linear distance between *talked* and the preposition. Comparing (1a) and (1b), the dependency distance between *talked* and its closest PP is the same; whereas the distance between *talked* and the farther PP is shorter in (1a). Thus the overall dependency length in (1a) is shorter than that in (1b), by placing the shorter PP, *with his friends*, closer to its head. Based on predictions by DLM, the structure of (1a) will be preferred.

Various attempts have been made to investigate the role of dependency length in predicting constituent ordering preferences. Though different metrics for measuring dependency length have been applied before, these metrics tend to be highly correlated (Gildea and Jaeger 2015, Wasow and Arnold 2003). Most studies have focused on syntactic alternations in English (Gildea and Temperley 2007, Rajkumar et al. 2016), including but not limited to postverbal PP orders (Hawkins 1999, Wiechmann and Lohmann 2013), verb particle constructions (Lohse et al. 2004), heavy noun phrase shift (Wasow 1997, Arnold et al. 2000) and the dative alternation (Wasow and Arnold 2003, Bresnan et al. 2007). Other work has examined one or a small number of other languages. For example, online experiments from Yamashita and

[1] If we treat the nominal head rather than the preposition as the head of the PP, the overall dependency length is still shorter when the short PP occurs closer to the head verb.

Chang (2001) showed that speakers tend to place long arguments ahead of short ones before the head verb in both transitive and ditransitive constructions in Japanese. Kizach (2012) demonstrated that constituent length affects the ordering of four syntactic structures in Russian: postverbal PPs, the double object construction, adversity impersonals and the order of subject, verb and object. In a corpus analysis, Rasekh-Mahand et al. (2016) modeled the effects of constituent length for relative clause extraposition in Persian. Studies on adjective-noun word order alternations in five Romance languages found that adjective placement is influenced not only by the dependency distance between the adjective and the noun, but also by the dependency lengths among the dependency subtrees of the noun phrase (Gulordava et al. 2015).

3 Data and methods

3.1 Data and preprocessing

We used multilingual corpora from the Universal Dependencies (UD) project version 2.5 (Zeman et al. 2019), which provides consistent and gold standard annotations across 157 treebanks for 90 languages. Each treebank is a syntactically parsed and annotated textual corpus, and certain languages have more than one treebank available.

Here we focus on contemporary languages and their treebank data. We searched in every treebank for sentences with VPs containing exactly two PP dependents that both occur on the same side of the head verb.[2] Abiding by the UD annotation scheme, we selected verbs with a part of speech tag as VERB, which denotes only main verb or content verb. PPs with no lexical head, in other words, PPs that only consist of a preposition or a postposition were not considered. The dependency relation between the lexical head of each PP and the head verb in the instances that we extracted was always *obl* (oblique). Following The World Atlas of Language Structures (WALS) (Dryer and Haspelmath 2013), we coded the word order features as well as the language family and genus of each language. In cases where there are multiple treebanks for the same language, we first calculated the effects of dependency length for each treebank individually (cf. Section 3.2). As we found the effect differences between the treebanks of the same language to be very small, we combined those treebanks to one for analysis. Languages with fewer than a total of 50 sentences that fit our criteria (i.e. containing a VP with two PPs attached on the same side of the verb) were not included in our analysis, leaving us with a dataset for 34 languages. These languages end up being mostly Indo-European, with

[2]Codes, data and figures for all results in the article are available at https://bitbucket.org/zoeyliu18/crosslinguistic-dlm/src/master/.

the exceptions of Arabic and Hebrew (Afro-Asiatic), Indonesian (Austronesian), Japanese (Japanese), and simplified (Mandarin) Chinese (Sino-Tibetan).[3]

3.2 Measures for dependency length effects

To approximate the effect of dependency length on PP orderings, we adopted a similar procedure to the one described by Hawkins (1999). We measured the lengths of both the PP closer to the verb and the PP farther from the verb as the number of word tokens in each PP. Each token has its own index in the UD annotation format, with both its POS and dependency relation encoded. If the preference for DLM holds for PP ordering typology, we will expect the proportion for when the shorter PP is closer to be significantly higher than that for when the longer PP is closer.

For hypothesis testing of the effect for dependency length and estimates of confidence intervals, we employed bootstrapping. The basic procedure for computation is as follows: (1) From a dataset of n VP instances for a certain language that we have extracted after preprocessing, we drew a sample of n VP instances with replacement, where some of the cases from the original dataset can appear multiple times; (2) Given this sample, we calculated the proportion for three different cases: when shorter PP occurs closer, when longer PP is closer, and when the two PPs are of equal length; (3) We repeated step (1) and (2) for 1,000,000 iterations to derive an empirical resampling distribution for each of the three different case scenarios described in (2); (4) For each resampling distribution, we calculated the mean and 95% confidence interval.

4 Results

Overall our dataset shows three different PP ordering patterns:

- languages with head-initial PPs after the head verb (e.g. Indonesian);
- languages with mixed PP orders where head-initial PPs appear after or before the head verb (e.g. German);
- languages with head-final PPs before the head verb (e.g. Japanese).

As we are taking a data-driven approach, we group the languages based on their PP ordering patterns first, while we take into consideration the grammatical and syntactic features of different language types as well within our analysis.

[3]Originally there are two treebanks for Chinese in UD 2.5 that contain enough instances which fit our search criteria. One appears in the simplified orthography and the other appears in the traditional orthography. Here we focus only on the simplified version.

4.1 Languages with head-initial PPs after head verb

For twelve languages in our dataset, when a head verb has two PP dependents occurring on the same side, they appear as head-initial PPs after the head verb. These languages are all head-initial, or have a dominant VO order. If the hypothesis for DLM holds, the PP that is the first in the sequence of the two PPs, which is closer to the head verb, will be preferred to be shorter. As presented in Figure 1, green, grey and yellow bars represent the proportion of cases when shorter PPs are closer, when longer PPs are closer, and when the two PPs have equal length, respectively. Among the twelve languages, eleven show a strong preference for DLM except for Latvian. On average, the number of sentences with the shorter PP closer to the verb is 3.9 times that of sentences with the longer PP closer to the verb. None of them displays a preference for the longer PP to occur first.

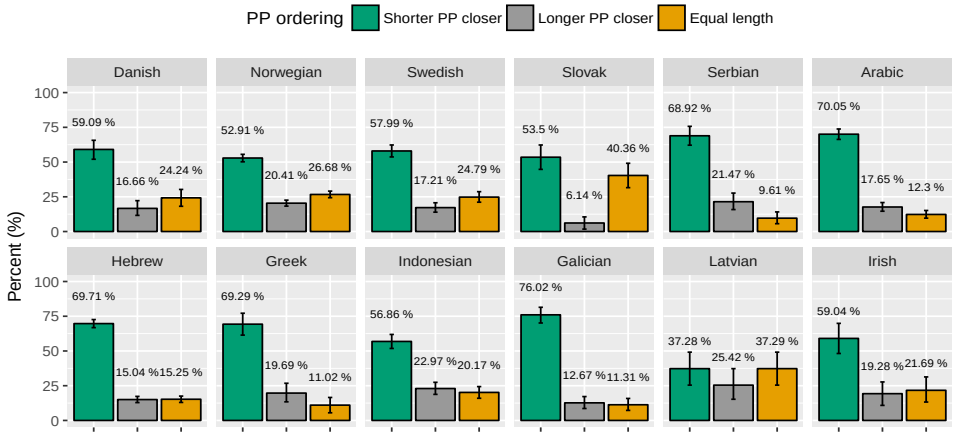
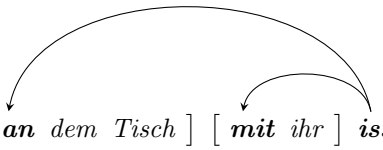


Figure 1: DLM in languages with head-initial PPs after the head verb. Error bars represent 95% confidence interval.

4.2 Languages with head-initial PPs after or before head verb

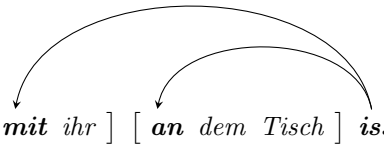
Sixteen languages in our dataset, shown in Figure 2, exhibit mixed PP ordering patterns. These languages fall into three different subtypes (genera): Germanic, Slavic and Romance, most of which are mixed-type languages. For these languages, we observe the same PP orders as those in Figure 1: when a head verb has two PP dependents on the same side, they occur as head-initial PPs after the head verb. In addition, they also have cases which place two head-initial PPs before the head verb. As an illustration of the latter ordering pattern, consider the following sentences in German.

(2) a. *dass er* [***an*** *dem Tisch*] [***mit*** *ihr*] ***isst***



that he [at the table] [with her] eats

b. *dass er* [***mit*** *ihr*] [***an*** *dem Tisch*] ***isst***



that he [with her] [at the table] eats

'that he eats with her on the table.'

If dependency length exerts a positive effect, the structure of (2a) will be preferred to that of (2b) since in (2a) the shorter PP is closer to the head verb. In this case the shorter PP is actually the second PP in the sequence of the two PPs. However, when a head verb has a head-initial PP dependent, putting the PP after the head verb will result in shorter dependency length than when ordering it before the verb. Particularly in the examples studied here, where the head verb has two head-initial PP dependents on the same side, if the shorter PP is closer, placing both PPs postverbally rather than preverbally leads to optimal overall dependency length. In addition, when two head-initial PPs appear before the head verb, the cost of the longest dependency between the verb and the first PP, which is farther away from the head verb, is incurred no matter what the order of the PPs is. In this case it does not seem to matter as much whether the shorter PP is closer. Thus when two head-initial PPs are preverbal, we expect there to be much weaker effect for dependency length at most, or no preference for DLM, compared to instances when the PPs are after the head verb.

As shown in Figure 2a, when the PPs occur after the head verb, dependency length also appears to be a strong predictor of PP ordering, which is comparable to the results in Figure 1. There is a dominant preference for the shorter PP to appear closer, or occur first, and none of these languages shows a tendency to put the longer PP closer to the head verb. The average ratio between the number of VPs with the shorter PP closer to the verb and the number of VPs with the longer PP closer is 3.0. On the other hand, when the PPs occur before the head verb as in Figure 2b, the preference for DLM almost disappears across the sixteen languages. Though Dutch exhibits a tendency to put the shorter PP closer to the head verb when the two PPs occur preverbally, the extent is less pronounced compared to when the two PPs appear postverbally in the language (1.4 vs. 2.2). Most of the Romance languages here in particular, except Portuguese (e.g. (3)) and Romanian, actually demonstrate a robust short-before-long ordering preference, where the shorter PPs tend to appear first, and farther away from the head verb.

(3) a. [***Desde*** *23 de Maio*] [***em*** *Portugal*] *António* ***viveu*** *em Lisboa*

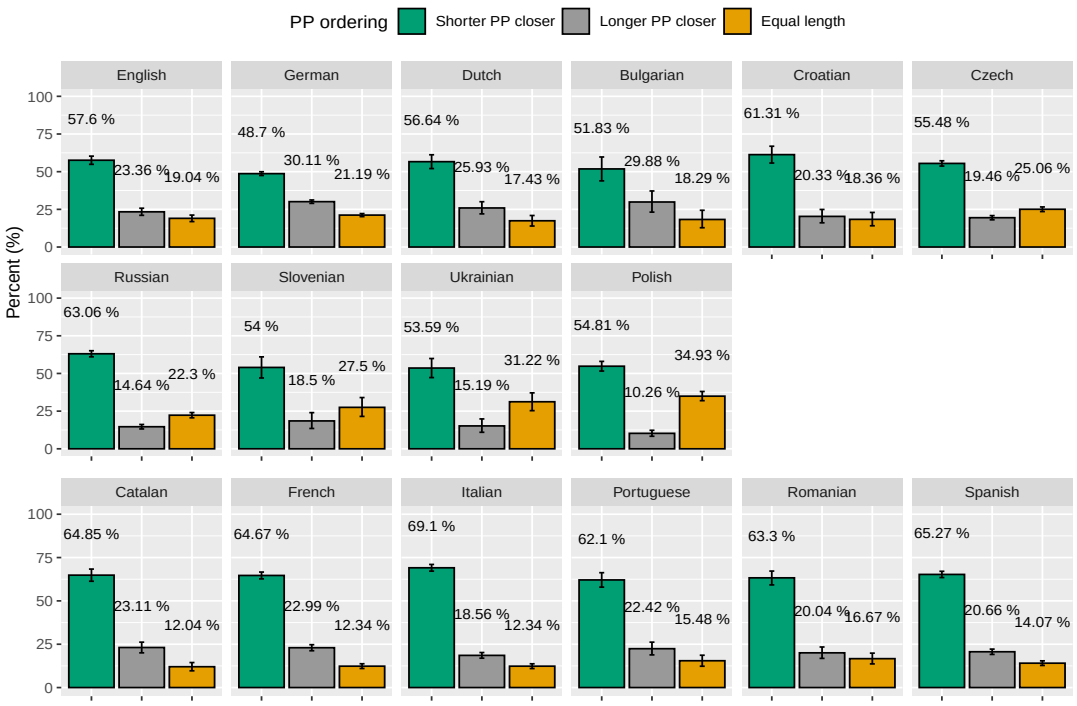
[Since 23 of May] [in Portugal] António lived in Lisbon

b. [***Em*** *Portugal*] [***desde*** *23 de Maio*] *António* ***viveu*** *em Lisboa*

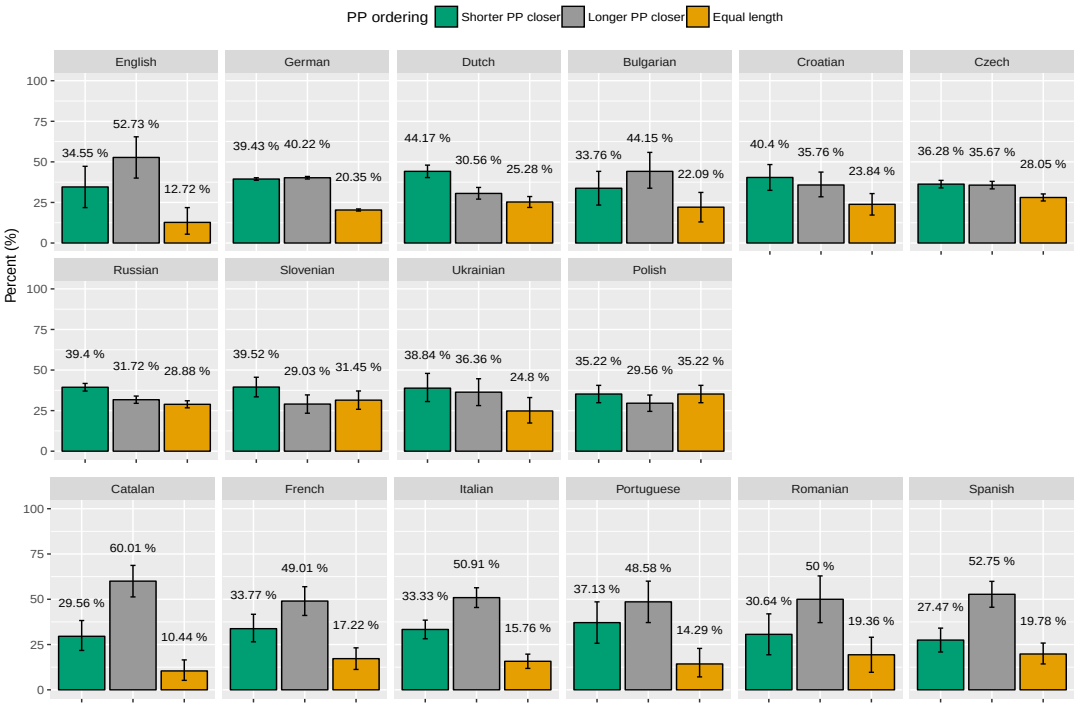
[In Portugal] [since 23 of May] António lived in Lisbon

'Since May 23, in Portugal, António lived in Lisbon.'

It is worth noting both the different and similar syntactic characteristics among the three language subtypes. Within the three Germanic languages, English differs from German and Dutch in that it is dominantly (S)VO. Though English has relatively rigid word orders, under certain pragmatic or social context, the object is preposed before the head verb, forming an OSV order (Prince 1988). On the other hand, German and Dutch are both classified as non-rigid OV, allowing a mixture of verb-medial and verb-final clausal structures. For these two languages,



(a) DLM in mixed-type languages when PPs appear after the verb.



(b) DLM in mixed-type languages when PPs appear before the verb.

Figure 2: DLM in mixed-type languages where PPs can appear after or before the verb. Error bars represent 95% confidence interval.

in main clauses with one or more auxiliary verbs as well as in subordinate clauses, the finite and lexical verbs regularly come in the clause-final position after the object. On the other hand, when the main clauses do not have an auxiliary verb, the finite and lexical verbs move to the front of the object (Hawkins 2019). Despite the relatively rigid requirement to place the verb in the final position given the aforementioned conditions, both German and Dutch do allow flexibility in the ordering of subject, direct and indirect object (MacWhinney et al. 1984, Kempen and Harbusch 2004). Both the Slavic and Romance languages have dominantly VO constructions. Different from German and Dutch, they do not exhibit differences for verb positions or verb movements that depend on clause types and whether the clause has auxiliary verbs. Nevertheless, it has been attested that Slavic and Romance languages also show variations of the order for core arguments of the head verb (Siewierska 1998, Liu 2010).

In spite of the similar and distinct syntactic properties of the Germanic, Slavic and Romance languages, combining the observations in both Figure 2a and Figure 2b, it is obvious that the contrast regarding the tendency for DLM between postverbal and preverbal PP orderings is shared across the three language subtypes: there is a strong preference for shorter dependencies in postverbal domains, whereas the preference is much weaker or disappears in preverbal structures. Though this pattern has been theoretically predicted based on a phrase structural framework by Hawkins (1994), to the best of our knowledge, our results here are the first empirical demonstration of this clean and sharp distinction.

A puzzle remains here, which is why there are instances with two head-initial PPs before the head verb in the first place. A recent study by Futrell et al. (2015b) has quantified the degree of word order freedom across a variety of languages. Their results showed that non-English Germanic, Slavic and Romance languages have more flexible orderings compared to other language genera. Thus one reason might be due to the relative flexibility of certain types of PPs in relation to the head verb, especially those that denote temporal expressions or locations. For instance, the temporal PP *in 1990* in German, which means 'in 1990', can occur both preverbally and postverbally. The PP *en casa* in Spanish, which describes the location 'at home', is also able to appear either before or after the head verb. Moreover, it is possible that the freedom of these PPs is less dependent on the rigidity of a language's full word order profile. This explains why we also see the same ordering pattern in English, though it has a much smaller number of total occurrences (55). These types of PPs are likely to appear as the one that is farther away from the head verb in the sequence of two PPs, and might tend to be relatively short on average as well. In a qualitative examination of all the 55 instances in English which have two head-initial PPs before the head verb, among the 29 cases which have a short-before-long orderings, 12 have a temporal PP occurring first (41.4%).

Besides the sixteen languages seen above, three other typologically mixed languages, Afrikaans, Persian and Chinese, present only instances where two head-initial PPs occur before the head verb. Among these three, Afrikaans and Persian are non-rigid OV languages and they are both prepositional. Afrikaans shows a significant preference for the shorter PP to appear closer to the head verb, though the tendency for DLM is weaker than in cases which have head-initial PPs after the head verb. The average ratio between the number of VPs with shorter PPs closer and the number of VPs with longer PPs closer is 2.2, compared to 3.9 for languages in Figure 1 and 3.0 for languages in Figure 2a. In contrast to German and Dutch

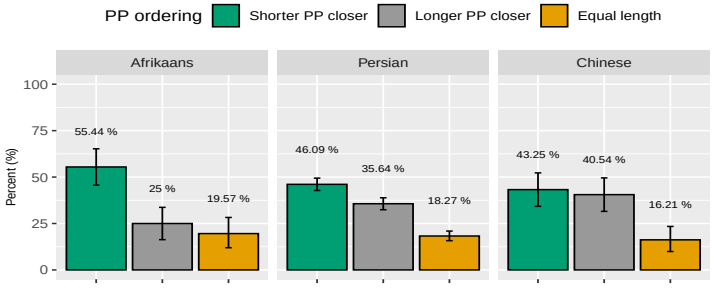
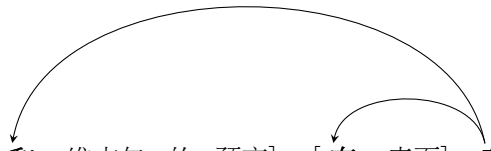


Figure 3: DLM in Afrikaans, Persian and Chinese. Error bars represent 95% confidence interval.

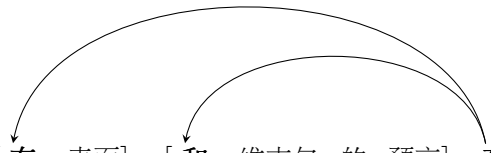
in Figure 2, it seems surprising that there are no VP instances with head-initial PPs placed after the head verb in Afrikaans. This might result from the specific genre of the Afrikaans corpus, which is mainly government texts. Additionally, the number of Afrikaans tokens in our dataset is small (92). On the other hand, Persian only exhibits a weak effect for dependency length, while in Chinese we observe no preference for DLM at all.

One unique characteristic about Chinese that should be pointed out here is that in contrast to the other languages in our study, Chinese has profound inconsistencies in the headedness of its various structures (Dryer 1991). Although the Chinese VP instances that we extracted all have two head-initial PPs before the head verb, as illustrated below in (4), the adposition system of Chinese has been argued to be more complicated (Li and Thompson 1974). While previous work has quantitatively classified Chinese as an SVO language (Sun and Givón 1985, Mei 1980, Liu et al. 2009), it has been suggested that Chinese has both prepositions and postpositions that form head-initial (e.g. (5a)) or head-final (e.g. (5b)) PPs respectively (Hawkins 1983, Wang and Sun 2015). This is different from the languages we have discussed so far, which are all prepositional (languages in Figure 1, Figure 2, Figure 3).

(4) 

zhexie yanlun [he weijier de yuyan] [zai biaomian] you suo churu

These comments [with Virgil's prophecy] [on the surface] have differences.



zhexie yanlun [zai biaomian] [he weijier de yuyan] you suo churu

These comments [on the surface] [with Virgil's prophecy] have differences.

'These comments have differences on the surface with Virgil's prophecy.'

(5) a. 在 美国
zai meiguo
in America
'in America'

b. 书包 里
shubao li
bag in
'in the bag'

c. 从 法庭 上
cong fating shang
from court on
'from court'

Nevertheless, cases such as (5c) also exist in Chinese, in which a phrase takes both a preposition and a postposition. Cai (2013) noted that the preposition and postposition in such contexts together form a circumposition, whereas Zhang (2013) proposed that what appears in the post-nominal position should be treated as a bound localizer rather than a postposition.

In our Chinese data, when examples similar to (5c) occur, the part of speech of the post-nominal form is annotated as an adposition, and it is treated as a clausal

modifier of its head noun. To test if this particular annotation scheme affects DLM, we removed sentences with cases as (5c). After the removal we are left with 78 sentences in total, with 42.31% having the shorter PP closer to the head verb and 35.9% having the longer PP closer to its head.

4.3 Languages with head-final PPs before head verb

In our dataset, three languages turned out to be rigid OV, which are Japanese and two Indic languages, Hindi and Urdu. For these languages, when a head verb has two PP dependents, they appear as head-final PPs before the head verb. If DLM holds, when the shorter PP occurs closer to the head verb, it is also the second PP in the sequence of the two PPs. This is similar as instances with head-initial PPs before the head verb, and contrasts cases with head-initial PPs after the head verb. As illustrated in the examples below, the ordering structure of (6a) will be preferred to that of (6b), because by putting the shorter PP, 自分 で, second and closer to the head verb, the overall dependency distance is shorter in (6a).

(6) a. 彼 は [赤い ペン **で**] [自分 **で**] **書いた**

kare wa [akai pen de] [jibun de] kaita

He [a red pen with] [himself by] wrote.

b. 彼 は [自分 **で**] [赤い ペン **で**] **書いた**

kare wa [jibun de] [akai pen de] kaita

He [himself by] [a red pen with] wrote.

'He wrote by himself with a red pen.'

As presented in Figure 4, the preference for DLM in the three rigid OV languages is much weaker overall. There appears to be only a mild effect of dependency length for Japanese. The ratio between the number of VPs with the shorter PP closer to the verb and the number of VPs with the longer PP closer is 1.5. Though this is substantially lower than the 3.9 and 3.0 average ratio, respectively, for the languages in Figure 1 and those in Figure 2a. The contrast between Hindi and Urdu

to languages with postverbal head-initial PPs is even more pronounced, with there being no observed tendency for shorter dependencies at all.

The observations presented here are against findings for consistently head-final languages in previous studies (e.g. Japanese (Yamashita 2002); Korean (Choi 2007). Their experiments have examined the relative order of subject, direct and indirect object in relation to the head verb, and their results show robust preferences for DLM. On the other hand, we use PP obliques as a test case, the order of which contains more word order flexibility and is less governed by specific grammatical constraints, compared to the order of the three core arguments as mentioned above. In our view, this offers more insights into probing the direct role of dependency length as a predictor for constituent orders.

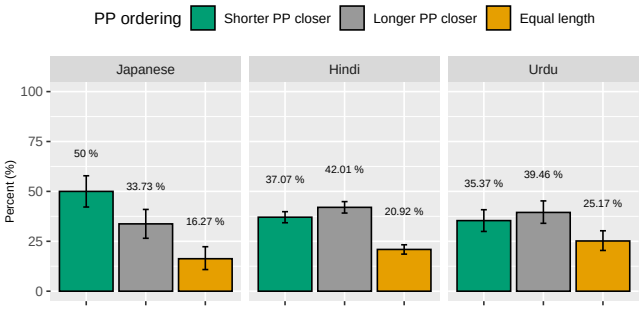


Figure 4: DLM in languages with head-final PPs before the head verb. Error bars represent 95% confidence interval.

We understand that the sample size of rigid OV languages in our dataset is quite small in contrast to the dominantly VO and mixed-type languages previously seen. Accordingly, we are cautious about drawing any conclusive statements here regarding the different preference for DLM between the different language (sub)types. More work is needed in the future to test whether the tendency for shorter dependencies exists and is consistent in other rigid head-final languages. Nevertheless, this does not necessarily undermine our findings here, as we believe the reasons for the small sample size presented in Figure 4 are not arbitrary.

First, unlike prepositions in head-initial languages, postpositions in head-final languages are relatively less "productive" (Tsunoda et al. 1995, Hawkins 2014). That is, rigid OV languages like Japanese have smaller inventories of postposition types compared to inventories of preposition types in dominantly head-initial languages such as English. This is likely due to the fact that the postpositions might have got lost during language change, or they have developed into suffixes rather than staying as freestanding grammatical elements. As a sanity check, we extracted all PP obliques that are dependents of verbs from the English treebank, then counted the different types of prepositions that are the functional heads of these PPs. We did the same for the different types of postpositions in the treebanks for Japanese,

Hindi and Urdu as well. As shown in Table 1, the three rigid OV languages have smaller inventories of postposition types in comparison to that of prepositions in English.

Table 1: Count of preposition or postposition in English, Japanese, Hindi and Urdu.

	English	**Japanese**	**Hindi**	**Urdu**
preposition	158			
postposition		68	123	112

These numbers lead to our second point, which is that what is expressed as a PP oblique in English is possibly realized in a different syntactic structure in these head-final languages. For example, dative verbs in English such as *give* and *send* can have two alternative argument structures, the double object structure, as in *I gave her a book*, as well as the prepositional dative structure, as in *I gave a book to her*. In the latter the indirect object *her* is marked by the preposition *to* and the two together form a PP oblique. However, in Japanese, as shown in (7), the indirect object for the corresponding verb of *give* is always realized as an NP via the attachment of the case marker に . The case markers are different from postpositions or prepositions as observed in PPs, which potentially results in fewer instances where the head verb has two PP dependents on the same side.

(7) 私 は 彼女 に 本 を 与えました
watashi wa kanojo ni hon o ataemashita
I -NOM she -DAT book -ACC gave
'I gave her a book.'

5 Discussion

5.1 Crosslinguistic effect and efficacy of DLM

Now we turn back to our initial questions. First, is there a preference for DLM in syntactic constructions with flexible constituent orderings across languages? We show empirically that the answer is yes. Dependency length is an effective typological determinant in PP order typology. Overall there is a strong crosslinguistic tendency for shorter PPs to appear closer to the head verb.

The second question we raised at the beginning is how the extent of DLM in ordering preferences varies for languages with different syntactic characteristics. Overall, we show that the efficacy of DLM depends on the specific ordering structures in different language types, yielding two general patterns. First, the preference for shorter dependencies is the most pronounced when head-initial PPs appear after the verb. By contrast, when they occur before the head verb, there is no or a

much weaker tendency for DLM. This contrast is the most visible in languages with mixed PP orders, shared across different language genera for most of the mixed-type languages in our dataset. Second, acknowledging their limited sample size in our dataset, we also found no or a much weaker preference for DLM in rigid head-final languages, which have head-final PPs before the head verb. Taking the two general patterns into consideration, our results indicate that across languages, preverbal PP orderings have no pronounced tendency for DLM in comparison to postverbal PP orders. This observation appears to be less dependent on the structural characteristics of specific language types and seems to be regardless of whether the PP is prepositional or postpositional.

Based on previous work, it appears that two aspects could explain the weaker tendency for DLM in certain language types: word order freedom and rich case marking. These two properties have been shown to be strongly correlated across a number of languages (Sapir 1921, Futrell et al. 2015b). Gildea and Temperley (2010) attributed the reason for why German has longer dependency length than English to the fact that German has comparatively more word order freedom and more verb-final structures. This might lead to more inconsistent head-dependent orderings and longer preverbal dependencies. Both Ros (2018) and Futrell et al. (2015a) suggested that head-final languages exhibit longer dependencies due to their rich case marking systems, which lead to ordering flexibility.

However, word order freedom and rich case marking are not always correlated across languages. When a language has a rich case marking system, that does not necessarily mean it has flexible word orders (e.g. Icelandic (Kiparsky 1996)). On the other hand, languages such as Lao (Enfield 2009) and Riau Indonesian (Gil 2017) have relatively free word orders, but they do not have rich case marking. Accordingly, how to precisely tease apart the separate roles of the two aspects in constituent ordering preference and how they are related to our study need further quantitative validations.

As analyzed in Section 4.2, the relative flexibility of certain types of PPs could explain why there are instances with two head-initial PP dependents before the head verb in the first place. However, it still remains unclear whether the structure would exploit this freedom and put the shorter PP closer to the head verb. Although it is possible that PPs which can only appear either preverbally or postverbally have more fixed positions in general, and might prefer ordering proximity to the head verb, regardless of whether that leads to overall shorter dependencies. In addition, most of the mixed-type and the rigid OV languages in our dataset have rich case marking. Both types also have longer dependencies in preverbal PP orderings. Nevertheless, when a language or a structure has more case markers but no pronounced tendency for DLM, it does not empirically verify that the long dependencies are caused by the case markers.

5.2 Speakers' vs. listeners' perspective

One possible reason for some of the variants in our results may lie in processing differences between syntactic comprehension and production, that is, whether the ordering patterns are mainly for the benefits of the listener or the speaker. In addition, languages with distinct word order features favor different factors or adopt language-specific strategies in order to facilitate either comprehension or production ease (Ueno and Polinsky 2009).

From the listeners' perspective, putting shorter constituents closer to their heads shortens the overall dependency lengths of the sentence, and will potentially ease comprehension and online parsing. This has been widely documented in comprehension tasks for different structures in a number of languages (Gibson 1998, 2000, Gibson and Wu 2013, Levy et al. 2013, Liu 2008), where longer dependency lengths lead to more processing difficulty, measured by reading times. Similarly here, the PP orderings might place the shorter PP closer to the head verb for the benefits of syntactic comprehension.

However, several studies examining the comprehension of verb-final structures indicate that the comprehension process for these structures is actually faster when the opposite patterns to DLM are observed (Levy 2013). For instance, it has been shown that in certain verb-final constructions in German, adding preverbal dependents between the arguments and the clause-final verb makes the sentence easier to process, at the processing of the verb at least (Konieczny 2000, Konieczny and Döring 2003). Vasishth and Lewis (2006) demonstrated that interposing dependents between arguments of the clausal-final verb facilitates processing in Hindi. Evidence from these studies suggests that adding more dependents before the verb makes the verb more predictable, and faster to read. The connection between these studies and ours may not be seemingly obvious, as all our VP instances have exactly two PP dependents. Nevertheless, it is possible that the absence of DLM for certain PP orderings in our study can be explained along a similar line. When the two PPs both appear before the head verb, the head verb might have more predictability if the PPs are ordered as short-before-long; or the longer PP might be more predictive of the following head verb.

From the speakers' perspective, however, as syntactic production is an incremental process (Ferreira and Swets 2002), the speakers might not abide by DLM to order their utterances. They might opt for constituent orderings that will facilitate utterance planning and ease production efforts (Arnold et al. 2000). In this case speakers are likely to produce short constituents first which are formulated faster (De Smedt 1994, MacDonald 2013), more predictable given the preceding context or easier to access conceptually (Lohmann and Takada 2014), and choose to postpone the longer ones in order to have more planning time for the whole sentence

they intend to speak. This is shown consistently in our study across languages when they have head-initial PPs after the head verb. In cases where the two PPs both occur before the same head verb, regardless of the headedness of the PP, there might also be a tendency for the shorter PP to appear before the longer one. Nevertheless, this short-before-long preference in constituent utterances has been countered by production studies of head-final languages such as Japanese (Yamashita and Chang 2001) and Korean (Choi 2007).

On the other hand, in particular from a production-based point of view, if a speaker orders the two PP dependents based on DLM, that indicates the speaker has planned ahead what to say for both phrases, then figures out which one is shorter and puts it closer to the head verb (Chang 2009). Though this is likely when both PPs are short, the scenario becomes less plausible when one or both PPs are relatively long. This advance planning seems even harder and against the incremental nature of language production when two PPs appear preverbally rather than postverbally, since the former requires planning of not just the two PPs, but also what the head verb is and where it occurs. By contrast, previous production studies of head-final languages have demonstrated that there is not always advanced selection of verb during processing (Momma et al. 2016). Thus whether and under what context constituent orders abide more by "short closer" or "short first" deserve more thorough exploration.

5.3 Other competing and cooperating factors

Crosslinguistic constituent ordering preference is multifactorial by nature (Hawkins 2014); it is a function of competing and cooperating factors, that multiple constraints are at play in order to derive the ideal ordering structure. In our study, the fact that the variation of the extent for DLM exists and that there are instances where dependency length has no effect (e.g. PPs with equal length) indicates that there are other constraints which work conjointly with or pull in different directions against dependency length to decide what the PP order will be. Accordingly, effects from other factors simultaneously lead to the different extent of DLM across languages.

Previous studies focusing on postverbal PP orders in English (Hawkins 1999, Wiechmann and Lohmann 2013) in particular have examined the roles of semantic closeness, discourse status and the traditional ordering rule of adverbial phrases, Manner before Place before Time. Though they have only found mild or no effect for the latter two factors, both experiments have confirmed the significant predictive power of semantic closeness. The earliest observation for semantic closeness stems from Behaghel (1932), who suggests that constituents that are semantically related should occur closer together syntactically. The preference for semantic closeness has also been proposed in the Minimize Domains principle by Hawkins (2004), which

states that minimizing lexical or semantic dependency is necessary as well in order to facilitate processing for the human parser. The strong effect of semantic closeness appears to also hold for preverbal PP orderings in Mandarin Chinese (Liu 2019), where semantic closeness is approximated as the argumenthood status of the PP. These studies together have shown that in cases where dependency length has weak or no effect, semantic closeness plays comparatively a stronger role. Thus it is possible that there is a strong preference for the semantically closer PP to occur closer to the head verb across the languages here, regardless of or less dependent on the dependency distance between the PP and the head verb.

In addition, the preverbal PP ordering patterns here corroborate results from previous experiments on English that have looked at adverb placement and preverbal modifying adjuncts (Gildea and Temperley 2010, Temperley 2007, Rasekh-Mahand et al. 2016). They have evidenced observations against DLM consistently. One of the reasons they have argued is that languages such as Chinese and Japanese tend to be more "topic-prominent", and they would prefer to put the topic phrase first before the head verb. Further studies are necessary to establish whether that was the case, which would in turn explain the weaker tendency for DLM in Chinese and Japanese in our dataset.

6 Conclusion and Future Work

With an investigation of PP ordering preferences for 34 languages, we present large-scale evidence for the first time that DLM exists in syntactic structures with flexible constituent orders crosslinguistically. Further more, we conclude that the efficacy of DLM depends on the specific ordering structures. Our results show a strong preference for DLM in postverbal PP orders, while no or a much weaker tendency for shorter dependencies in preverbal PP orders across languages.

One follow-up study could be quantifying the occurrence flexibility of the PPs for languages that have mixed PP orderings. As we discussed in previous sections, it is likely that the PP that is farther away from the head verb can appear more flexibly in either preverbal or postverbal domains. A fruitful measure as an approximation for the word order variability is entropy (Futrell et al. 2015b, Levshina 2019). Additionally, the appearance of a PP with more word order flexibility is less predictable given the same context, which can be quantified with probabilistic esimates from computational language models. Other challenges reside in identifying additional factors that also influence constituent orders as well as in quantitatively approximating their effects and interactions with dependency length. In particular, future research needs to carry out related investigations from a crosslinguistic perspective, instead of making strong claims about languages with studies that are

only English-centric.

What is more, the observations drawn from our study, combined with previous findings, can be generalized to other syntactic structures as well. For example, though it is not quite straightforward to clearly tease apart DLM and the principle of "short first", constructions such as the alternations of verb and adverb might provide more insights into the matter, especially since similar to PPs, adverbs can appear in both preverbal and postverbal domains as well, and tend to be shorter than other types of dependents of the same head verb in general. On another note, given the indication that the semantically closer, or more argument-like PP occurs closer to the head verb, a generalized question one might ask is whether more argument-like phrases prefer proximity to their syntactic heads across languages. Constructions such as those in which a head verb has a direct object and a PP oblique occurring on the same side could be potential test cases (Hawkins 2014) to establish typological quantitative evidence for this question, and to compare the predictive power of DLM and argumenthood status.

References

Arnold, Jennifer E., Anthony Losongco, Thomas Wasow, and Ryan Ginstrom. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76(1). 28–55.

Behaghel, Otto. 1932. *Deutsche Syntax: eine geschichtliche Darstellung*. Heidelberg: Carl Winters Universitätsbuchhandlung.

Bender, Emily M. 2009. Linguistically Naıve != Language Independent: Why NLP Needs Linguistic Typology. In *Proceedings of the EACL 2009 workshop on the interaction between linguistics and computational linguistics: virtuous, vicious or vacuous?*, 26–32. Athens, Greece: Association for Computational Linguistics.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina, and R Harald Baayen. 2007. Predicting the dative alternation. In *Cognitive foundations of interpretation*, 69–94. KNAW.

Cai, Lei. 2013. The Semantic Functions of Prepositions and Postpositions in Chinese Spatial Circumpositions – A Perspective from Language Typology. In Pengyuan Liu and Qi Su (eds.), *Chinese lexical semantics*, 248–257. Berlin, Heidelberg: Springer Berlin Heidelberg.

Chang, Franklin. 2009. Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language* 61(3). 374–397.

Choi, Hye-Won. 2007. Length and Order: A Corpus Study of Korean Dative-Accusative Construction. *Discourse and Cognition* 14(3). 207–227.

De Smedt, KJMJ. 1994. Parallelism in Incremental Sentence Generation. *Parallel natural language processing*. 421–447.

Dryer, Matthew S. 1991. SVO languages and the OV: VO typology. *Journal of Linguistics* 27(2). 443–482.

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

Dyer, William Edward. 2017. *Minimizing Integration Cost: a General Theory of Constituent Order*. University of California, Davis dissertation.

Enfield, Nicholas J. 2009. 'Case relations' in Lao, a radically isolating language. In *The Oxford handbook of case*, 808–819. Oxford University Press.

Ferreira, Fernanda and Benjamin Swets. 2002. How Incremental is Language Production? Evidence from the Production of Utterances Requiring the

Computation of Arithmetic Sums. *Journal of Memory and Language* 46(1). 57–84.

Futrell, Richard, 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.

———2015b. Quantifying word order freedom in dependency corpora. In *Proceedings of the third international conference on dependency linguistics (depling 2015)*, 91–100. Uppsala, Sweden: Uppsala University, Uppsala, Sweden.

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

———2000. The Dependency Locality Theory: A Distance-based Theory of Linguistic Complexity. In Alec Marantz, Yasushi Miyashita, and Wayne O'Neil (eds.), *Image, language, brain*, 95–126. Cambridge, Mass.: MIT Press.

Gibson, Edward and H.-H. Iris Wu. 2013. Processing Chinese relative clauses in context. *Language and Cognitive Processes* 28(1-2). 125–155.

Gil, David. 2017. Chapter 19 - Isolating-Monocategorial-Associational Language. In Henri Cohen and Claire Lefebvre (eds.), *Handbook of categorization in cognitive science (second edition)*, Second Edition, 471–510. San Diego: Elsevier.

Gildea, Daniel and T. Florian Jaeger. 2015. Human languages order information efficiently. *arXiv preprint arXiv:1510.02823*.

Gildea, Daniel and David Temperley. 2007. Optimizing Grammars for Minimum Dependency Length. In *Proceedings of the 45th annual meeting of the association of computational linguistics*, 184–191. Prague, Czech Republic: Association for Computational Linguistics.

———2010. Do grammars minimize dependency length? *Cognitive Science* 34(2). 286–310.

Gulordava, Kristina, Paola Merlo, and Benoit Crabbé. 2015. Dependency length minimisation effects in short spans: a large-scale analysis of adjective placement in complex noun phrases. In *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 2: short papers)*, 477–482. Beijing, China: Association for Computational Linguistics.

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

Hawkins, John A. 1994. *A performance theory of order and constituency*. Vol. 73. Cambridge University Press.

———1999. The relative order of prepositional phrases in English: Going beyond Manner–Place–Time. *Language variation and change* 11(3). 231–266.

———2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

———2014. *Cross-linguistic Variation and Efficiency*. Oxford: Oxford University Press.

———2019. Word-external properties in a typology of Modern English: A comparison with German. *English Language & Linguistics* 23(3). 701–727.

Jaeger, T. Florian and Elisabeth J. Norcliffe. 2009. The Cross-linguistic Study of Sentence Production. *Language and Linguistics Compass* 3(4). 866–887.

Jaeger, T. Florian and Harry Tily. 2011. On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science* 2(3). 323–335.

Kempen, Gerard and Karin Harbusch. 2004. A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In Thomas Pechmann and Christopher Habel (eds.), *Multidisciplinary approaches to language production*, vol. 157, 173–182. Mouton de Gruyter.

Kiparsky, Paul. 1996. The shift to head-initial VP in Germanic. In Höskuldur Thráinsson, Samuel David Epstein, and Steve Peter (eds.), *Studies in comparative germanic syntax*, vol. 2, 140–179. Kluwer Dordrecht.

Kizach, Johannes. 2012. Evidence for weight effects in Russian. *Russian linguistics* 36(3). 251–270.

Konieczny, Lars. 2000. Locality and Parsing Complexity. *Journal of Psycholinguistic Research* 29(6). 627–645.

Konieczny, Lars and Philipp Döring. 2003. Anticipation of clause-final heads: Evidence from eye-tracking and SRNs. In *Proceedings of the iccs/ascs joint international conference on cognitive science*, 13–17.

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

Levy, Roger. 2013. Memory and Surprisal in Human Sentence Comprehension. In Roger P. G. van Gompel (ed.), *Sentence processing*, 78–114. Hove: Psychology Press.

Levy, Roger, Evelina Fedorenko, and Edward Gibson. 2013. The syntactic complexity of Russian relative clauses. *Journal of Memory and Language* 69(4). 461–495.

Li, Charles N and Sandra A. Thompson. 1974. Co-verbs in Mandarin Chinese: Verbs or prepositions? *Journal of Chinese Linguistics* 2(3). 257–278.

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

———2010. Dependency direction as a means of word-order typology: a method based on dependency treebanks. *Lingua* 120(6). 1567–1578.

Liu, Haitao, Yiyi Zhao, and Wenwen Li. 2009. Chinese Syntactic and Typological Properties Based on Dependency Syntactic Treebanks. *Poznań Studies in Contemporary Linguistics* 45(4). 509–523.

Liu, Zoey. 2019. A Comparative Corpus Analysis of PP Ordering in English and Chinese. In *Proceedings of the first workshop on quantitative syntax (quasy, syntaxfest 2019)*, 33–45. Paris, France: Association for Computational Linguistics. `https://doi.org/10.18653/v1/W19-7905`. `https://www.aclweb.org/anthology/W19-7905`.

Lohmann, Arne and Tayo Takada. 2014. Order in NP conjuncts in spoken English and Japanese. *Lingua* 152. 48–64.

Lohse, Barbara, John A. Hawkins, and Thomas Wasow. 2004. Domain Minimization in English Verb-Particle Constructions. *Language* 80(2). 238–261.

MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in psychology* 4. 226.

MacWhinney, Brian, Elizabeth Bates, and Reinhold Kliegl. 1984. Cue validity and sentence interpretation in English, German, and Italian. *Journal of verbal learning and verbal behavior* 23(2). 127–150.

Mei, Kuang. 1980. Is modern Chinese really a SOV language? *Cahiers de Linguistique-Asie Orientale* 7(1). 23–45.

Merlo, Paola. 2015. Predicting word order universals. *Journal of Language Modelling* 3(2). 317–344.

Momma, Shota, L Robert Slevc, and Colin Phillips. 2016. The timing of verb selection in Japanese sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 42(5). 813.

O'Horan, Helen, Yevgeni Berzak, Ivan Vulić, Roi Reichart, and Anna Korhonen. 2016. Survey on the Use of Typological Information in Natural Language Processing. In *Proceedings of COLING 2016, the 26th international con-*

ference on computational linguistics: technical papers, 1297–1308. Osaka, Japan: The COLING 2016 Organizing Committee.

Östling, Robert. 2015. Word Order Typology through Multilingual Word Alignment. In *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 2: short papers)*, 205–211. Beijing, China: Association for Computational Linguistics.

Ponti, Edoardo Maria, Helen O'horan, Yevgeni Berzak, Ivan Vulić, Roi Reichart, Thierry Poibeau, Ekaterina Shutova, and Anna Korhonen. 2019. Modeling Language Variation and Universals: A Survey on Typological Linguistics for Natural Language Processing. *Computational Linguistics* 45(3). 559–601.

Prince, Ellen F. 1988. On pragmatic change: The borrowing of discourse functions. *Journal of pragmatics* 12(5). 505–518.

Rajkumar, Rajakrishnan, Marten van Schijndel, Michael White, and William Schuler. 2016. Investigating locality effects and surprisal in written English syntactic choice phenomena. *Cognition* 155. 204–232.

Rasekh-Mahand, Mohammad, Mojtaba Alizadeh-Sahraie, and Raheleh Izadifar. 2016. A corpus-based analysis of relative clause extraposition in Persian. *Ampersand* 3. 21–31.

Ros, Idoia. 2018. *Minimizing dependencies across languages and speakers. Evidence from Basque, Polish and Spanish and native and non-native bilinguals.* University of the Basque Country dissertation.

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

Siewierska, Anna (ed.). 1998. *Constituent order in the languages of Europe.* Walter de Gruyter.

Sun, Chao-Fen and Talmy Givón. 1985. On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implications. *Language.* 329–351.

Temperley, David. 2007. Minimization of dependency length in written English. *Cognition* 105(2). 300–333.

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

Tsunoda, Tasaku, Sumie Ueda, and Yoshiaki Itoh. 1995. Adpositions in word-order typology. *Linguistics* 33(4). 741–762.

Ueno, Mieko and Maria Polinsky. 2009. Does headedness affect processing? A new look at the VO–OV contrast. *Journal of Linguistics* 45(3). 675–710.

Vasishth, Shravan and Richard L. Lewis. 2006. Argument-Head Distance and Processing Complexity: Explaining Both Locality and Antilocality Effects. *Language* 82(4). 767–794.

Wang, Dingquan and Jason Eisner. 2017. Fine-Grained Prediction of Syntactic Typology: Discovering Latent Structure with Supervised Learning. *Transactions of the Association for Computational Linguistics (TACL)* 5. 147–161.

Wang, William S-Y and Chaofen Sun. 2015. *The Oxford handbook of Chinese linguistics*. Oxford Handbooks.

Wasow, Thomas. 1997. End-Weight from the Speaker's Perspective. *Journal of Psycholinguistic research* (3). 347–361.

Wasow, Thomas and Jennifer Arnold. 2003. Post-verbal constituent ordering in English. *Topics in English Linguistics* 43. 119–154.

Wiechmann, Daniel and Arne Lohmann. 2013. Domain minimization and beyond: Modeling prepositional phrase ordering. *Language Variation and Change* 25(1). 65–88.

Yamashita, Hiroko. 2002. Scrambled sentences in Japanese: Linguistic properties and motivations for production. *Text – Interdisciplinary Journal for the Study of Discourse* 22(4). 597–634.

Yamashita, Hiroko and Franklin Chang. 2001. "Long before short" preference in the production of a head-final language. *Cognition* 81(2). B45–B55.

Zeman, Daniel, 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, Loıc 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, Slavomır 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, Katarına 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 Grūzıtis, Bruno Guillaume, Céline Guillot-Barbance, Nizar Habash, Jan Hajič, Jan Hajič jr., Mika Hämäläinen, Linh Hà M, 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, lájıdé Ishola, Tomáš Jelınek, 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 Martınez 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 Nguy˜ên Th, Huy`ên Nguy˜ên Th Minh, Yoshihiro Nikaido, Vitaly Nikolaev, Rattima Nitisaroj, Hanna Nurmi, Stina Ojala,

Atul Kr. Ojha, Adéday 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, Benoıt Sagot, Shadi Saleh, Alessio Salomoni, Tanja Samardžić, Stephanie Samson, Manuela Sanguinetti, Dage Särg, Baiba Saulıte, 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, Taksum 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. `http://hdl.handle.net/11234/1-3105`.

Zhang, Niina Ning. 2013. Adpositions. In Rint Sybesma, Wolfgang Behr Yueguo Gu, Zev Handel, and C.-T. James Huang James Myers (eds.), *Encyclopedia of Chinese Language and Linguistics*, 116–122. Brill.

7 Appendix: Descriptive statistics of total VP instances

The number of total VP instances examined in each language under different case scenarios is presented below.

Table 2: Descriptive statistics of total VP instances for languages with head-initial PPs after the head verb.

Language	**Total VP instances**
Danish	198
Norwegian	1323
Swedish	488
Arabic	561
Hebrew	964
Greek	127
Indonesian	357
Galician	221
Latvian	59
Irish	83
Serbian	177
Slovak	114

Table 3: Descriptive statistics of total VP instances for Afrikaans, Persian and Chinese.

Language	**Total VP instances**
Afrikaans	92
Persian	870
Chinese	111

Table 4: Descriptive statistics of total VP instances for languages with head-final PPs before the verb.

Language	**Total VP instances**
Japanese	166
Hindi	1152
Urdu	294

Table 5: Descriptive statistics of total VP instances when head-initial PPs appear after the verb, in languages where head-initial PPs can also appear before the verb.

Language	**Total VP instances**
English	1250
German	6400
Dutch	459
Bulgarian	164
Ukrainian	237
Slovenian	200
Russian	2179
Czech	3093
Croatian	305
Polish	936
French	2236
Spanish	2672
Portuguese	504
Romanian	564
Italian	2188
Catalan	714

Table 6: Descriptive statistics of total VP instances when head-initial PPs appear before the verb, in languages where head-initial PPs can also appear after the verb.

Language	**Total VP instances**
English	55
German	14892
Dutch	625
Bulgarian	77
Ukrainian	121
Slovenian	248
Russian	1655
Czech	1640
Croatian	151
Polish	318
French	151
Spanish	182
Portuguese	70
Romanian	62
Italian	330
Catalan	115