A comparative corpus analysis of PP ordering in English and Chinese

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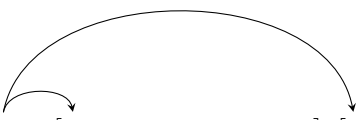University of California, Davis
`yiliu@ucdavis.edu`

Abstract

We present a comparative analysis of PP ordering in English and (Mandarin) Chinese, two languages with distinct typological word order characteristics. Previous work on PP orderings have mainly focused on English using data of relatively small size. Here we leverage corpora of much larger scale with straightforward annotations. We use the Penn Treebank for English, which includes three corpora that cover both written and spoken domains, and the Chinese Penn Treebank for Chinese. We explore the individual effect of dependency length, the argument status of the PP (argument or adjunct) and the traditional adverbial ordering rule, Manner before Place before Time. In addition, we evaluate the predictive power of dependency length and argument status with weights estimated from logistic regression models. We show that while dependency length plays a strong role across genre for English, it only exerts a mild effect in Chinese. On the other hand, the argument status of the PP has a pronounced role in both languages, that is, there exists a strong tendency for the argument-like PP to appear closer to the head verb than the adjunct-like PP. Our work contributes empirically to the long-standing proposal in linguistic typology that crosslinguistic word ordering preference is driven by cooperating and competing principles.
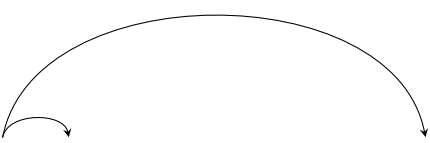
1 Introduction

Recent research has presented typological evidence that the overall or average dependency lengths between syntactic heads and their dependents tend to be minimized by their grammars as a whole (Futrell et al., 2015). Other experiments looking at specific syntactic constructions of individual languages that have alternative constituent orderings have also shown that speakers opt for constituents of shorter length to appear closer to their syntactic heads and thus shorten overall dependency distance in the sentence (Jaeger and Norcliffe, 2009). It has been argued as well as demonstrated in psycholinguistic and corpus studies that the preference for shorter dependencies is driven by processing efficiency (Gibson, 1998; Levy, 2013) and ease of communication (Hawkins, 2014; Gibson et al., In Press). As an illustration of how dependency length minimization (DLM) applies to constituent orderings, consider the following sentences in English:

(1) a. *Dylan presented* $\begin{bmatrix} \textbf{\textit{on}} & \textit{something} & \textit{linguistic} \end{bmatrix}$ $\begin{bmatrix} \textbf{\textit{to}} & \textit{her} & \textit{colleagues} & \textit{and} & \textit{friends} \end{bmatrix}$.

b. *Dylan presented* $\begin{bmatrix} \textbf{\textit{to}} & \textit{her} & \textit{colleagues} & \textit{and} & \textit{friends} \end{bmatrix}$ $\begin{bmatrix} \textbf{\textit{on}} & \textit{something} & \textit{linguistic} \end{bmatrix}$.

Both (1a) and (1b) have two PPs, shown within square brackets: *on something linguistic* and *to her colleagues and friends*. Switching the order of the two PPs does not change the grammaticality nor the semantic meaning of the sentence. As indicated by the syntactic dependency arcs, we consider the prepositions in both PPs to be the heads of their respective constituents, and to be dependents of the verb *presented*, which is the head of the VP in each sentence. The length of the dependency that attaches each PP to its corresponding VP is then the linear distance between the head of the dependency relation (the verb *presented*) and the preposition, which serves as the dependent. In both (1a) and (1b), the dependency length between *presented* and its closest PP is the same; however, the distance between *presented* and the farther PP is shorter in (1a), where the PP of shorter length is placed closer to the verb. From this example, we can see that in cases where the VP has two PP dependents occurring on the same side of the head verb, DLM predicts that there is a preference for placing the shorter PP closer to its head.

The effect of dependency length on syntactic preferences has been examined in various ways (Gibson, 2000; Gildea and Temperley, 2007; Gildea and Temperley, 2010; Temperley, 2007). Although strong evidence for DLM has been found, it is clear that it is not the only motivation in determining preferred word orders. Other competing and/or cooperating factors must also be at play that govern ordering preferences. The interaction between DLM and other principles and constraints in different contexts is currently under investigation (Gulordava et al., 2015; Wiechmann and Lohmann, 2013).

This study makes a contribution to the aforementioned research direction. We present a comparative analysis of PP orderings in English and Chinese, two languages with distinct typological properties. We focus in particular on VP instances with exactly two PP dependents appearing on the same side of the head verb, the ordering of which permits flexibility. Previous work on PP orderings has mainly focused on English with relatively small amounts of data (Hawkins, 1999; Wiechmann and Lohmann, 2013). Here we resort to corpora of much larger scale with straightforward annotations. For English, we use the Penn Treebank (PTB) (Marcus et al., 1993), which includes syntactic structures for approximately one million words of text from each of: the Wall Street Journal (WSJ), the Brown corpus (Kučera and Francis, 1967) and transcriptions of spontaneous spoken conversations from the Switchboard corpus (Godfrey et al., 1992). For Chinese, we exploit the Penn Chinese Treebank (CTB) (Xue et al., 2005), which has a total of 500K words. We probe to what extent dependency length, the argument status of the PP (argument or adjunct), and the traditional adverbial ordering rule, Manner before Place before Time, explain the observed PP ordering patterns. We explore how the effects of the three factors and their interactions differ across genres for English, and between the two languages.
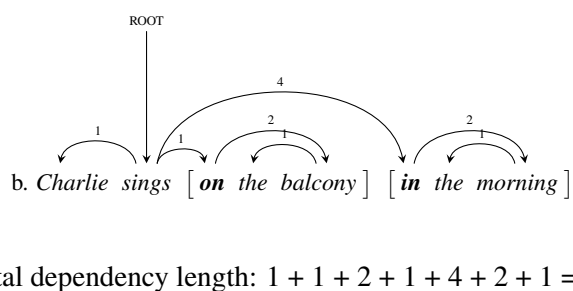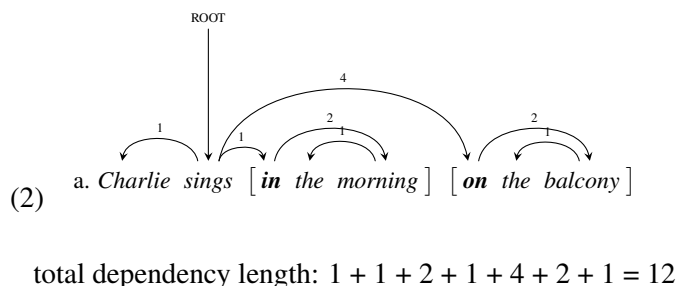
2 Related Work

2.1 Dependency length

Preceding DLM, the preferences for shorter syntactic dependencies have been formulated in various principles, including Early Immediate Constituent (Hawkins, 1994), Minimize Domains (Hawkins, 2004) and Dependency Locality Theory (Gibson, 2000). These principles all suggest the same idea that if grammatical alternatives exist for the syntactic constructions, there is a tendency to put shorter constituents closer to the syntactic heads and to avoid longer dependencies. Empirical support for the significant effects of dependency length in constituent ordering preferences has been found in various studies. Most work has focused on one specific or few syntactic structures in English, ranging from heavy NP shift (Wasow, 1997a; Arnold et al., 2000), dative alternation (Wasow and Arnold, 2003; Bresnan et al., 2007), verb particle constructions (Lohse et al., 2004), to postverbal PP orderings (Hawkins, 1999; Wiechmann and Lohmann, 2013) and so on. Some studies have extended their investigations to constructions in a small number of languages other than English, including Japanese (Yamashita and Chang, 2001; Yamashita, 2002), Korean (Choi, 2007), Russian (Kizach, 2012), Persian (Rasekh-Mahand et al., 2016) and certain Romance languages (Gulordava and Merlo, 2015).

As powerful as its effects are, dependency length itself will not suffice for predicting syntactic orderings across languages. First of all, dependency length is not able to indicate which ordering structure might be preferred when switching the order of constituents does not change the overall dependency length. As seen in the following examples, both (2a) and (2b) have two PPs, which are of equal length,

occurring after the head verb *sings*. Here we calculate dependency length as the distance from the head to its dependents, including the head verb. Changing the order of the two PPs does not appear to affect the total dependency lengths in (2a) and (2b) [1].

(2)
a. *Charlie sings* [***in** the morning*] [***on** the balcony*]

total dependency length: 1 + 1 + 2 + 1 + 4 + 2 + 1 = 12

b. *Charlie sings* [***on** the balcony*] [***in** the morning*]

total dependency length: 1 + 1 + 2 + 1 + 4 + 2 + 1 = 12

Additionally, previous studies have shown that longer dependencies are preferred in certain syntactic constructions in corpora (e.g. preverbal adjuncts ordering in English (Rajkumar et al., 2016; Temperley, 2007)). Other psycholinguisitc experiments presented that subordinate clausal structures with longer dependencies are easier to process in both rigid OV (e.g. Hindi (Vasishth and Lewis, 2006)) or non-rigid OV languages (e.g. German (Konieczny, 2000; Konieczny and Döring, 2003)).

What's more, the efficacy of DLM appears to vary crosslinguistically. Comparing German and English, Gildea and Temperley (2007) showed that German tends to have longer dependencies and minimizes dependency lengths to a lesser extent. They argued that the prevalent OV structures in German, where the verbs are in the final position, enlarge the dependency distance between the verb and its preverbal dependents. One other possible explanation that they discussed was that German has relatively free word order, which means that the constituent orderings in German may be driven more by considerations other than DLM. Looking at 37 languages, Futrell et al. (2015) suggested that head-final languages such as Japanese have longer dependencies compared to head-initial languages like English and Arabic. They conjectured that rich case marking systems in head-final languages allow more word order freedom, which lead to longer dependencies. Regardless of the proposed explanations, the fact that dependency length is minimized to different extents, and that it is not always minimized in certain cases indicate there are other cooperating and competing biases, cognitive or structural, that are effective and interact with DLM (Hawkins, 2014; MacWhinney et al., 2014)

2.2 Argument status

The role of argument status in constituent orderings is hardly new. Arguments prefer to be adjacent to their syntactic heads compared to adjuncts, which has been shown extensively in English (Culicover et al., 2005; Jackendoff, 1977; Pollard and Sag, 1994) as well as in other languages (Tomlin, 1986; Dyer, 2017).

Previous literature has distinct ways of deciding whether a constituent is an argument or an adjunct when investigating its effects on syntactic ordering preferences. For instance, in an examination of heavy NP shift, Wasow (1997b) found different shifting patterns when the verb and the PP are collocations than

---

[1] Following the measure from Hudson (1995), which approximates dependency length as the number of intervening tokens between the head and its dependents, the total dependency length is also the same for both sentences, except that the value will be 0 + 0 + 1 + 0 + 3 + 1 + 0 = 5.

when they are not. When the verb and the PP are collocational, in other words, when the PP is considered to be an argument of the verb, (e.g *take into account*), there is a greater tendency to shift the NP and place the PP immediately after the verb. On the other hand, when the verb and the PP are not collocational (e.g., *take to the store*), the proportion of examples where the NP is shifted is much smaller. Using 394 relevant sentences in English, Hawkins (1999) noted the significant roles of syntactic dependency and the argument status of the PP, namely that the PP which is a complement of the head verb tends to appear closer to the verb. Wiechmann and Lohmann (2013)found similar results with 1,256 sentences from both the written and spoken sections of the International Corpus of English. Both Hawkins (1999) and Wiechmann and Lohmann (2013) used entailment tests to define the argument status of the PP in relation to the verb. For instance, the PP *on his family* in the sentence *He counts on his family* is an argument of the verb *counts*, since the sentence does not entail *He counts*. By contrast, the PP *in the park* in the sentence *He played in the park* is an adjunct of the verb *played* because the sentence does entail *He played*. With the same entailment tests, Lohse et al. (2004) showed that the length of the object NP as well as the argument status of the particle in relation to the verb influence the orders of verb particle constructions in English.

2.3 Manner Place Time (MPT)

Proposed in Quirk et al. (1985), the traditional ordering rule for PPs and adverbials in postverbal position in English appears to follow Manner before Place before Time (MPT), as in *Zoey danced [$_{manner}$ elegantly] [$_{place}$ on the dance floor] [$_{time}$ at night]*. In contrast, this rule applies in the opposite direction when the PPs and adverbials occur in preverbal positions. That is, the ordering of preverbal PPs and adverbials follows Time before Place before Manner (TPM) (Hawkins, 1999). While Hawkins (1999) found that MPT plays no significant role in PP ordering in English, Wiechmann and Lohmann (2013) showed that it has a statistically significant yet weak effect.

3 Experiments[2]

3.1 Data

We searched for sentences in PTB and CTB with verb phrases containing exactly two PPs attached to the same side of the same head verb, where the ordering of the PPs allows certain flexibility.

Corpus	Total
WSJ	3596
Brown	3033
Switchboard	1187
Ctb	250

Table 1: Total VP instances for each corpus

3.2 Measures of each factor

3.2.1 Dependency length

Though different metrics have been applied to approximate dependency length in the literature, these measures have been demonstrated to be closely correlated (Gildea and Jaeger, 2015). To estimate the effect of dependency length on PP ordering, we followed the simple procedure as Hawkins (1999). We measured the lengths of the PP closer to the verb and of the PP farther from the verb as the number of tokens in each PP. We approximated phrase length using the number of tokens according to the treebank tokenization. We then calculated the proportion of cases where the shorter PP occurs closer to the head verb, the longer PP appears closer and when the two PPs are of equal length, for each corpus separately.

---

[2]Codes available at `https://zoeyliu18@bitbucket.org/zoeyliu18/pp-order-in-english-vs-chinese.git`

3.2.2 Argument status

To decide the argument status of a PP constituent, we borrowed the coding scheme from Merlo and Ferrer (2006), which carefully distinguishes PP arguments and adjuncts given their annotated grammatical function and semantic tag from the treebanks, shown in Table 2. As described in their paper, the motivation to include untagged PPs as arguments is due to that in the corpora, NPs (direct object & indirect object) and sentential constituents that are clearly arguments of the verb are left untagged (Marcus et al., 1994; Bies et al., 1995). The difference between argument and adjunct is gradient and not a binary distinction. Rather than looking at each PP as strictly an argument or an adjunct, we interpret the notion as an approximation for how argument-like and adjunct-like each PP is relative to the head verb. To analyze the effect of argument status, we only examined VP instances that have one argument-like PP and one adjunct-like PP (WSJ: $n = 1371$, Brown: $n = 1048$, Switchboard: $n = 470$, CTB: $n = 68$). We then computed the proportion of cases when the argument-like PP occurs closer. Statistical significance of the effects for both dependency length and argument status in each language were evaluated with Monte Carlo permutation test for 1,000,000 iterations.

Argument-like PPs	
-CLR	dative object if dative shift not possible (e.g., donate); phrasal verbs; predication adjuncts
-EXT	marks adverbial phrases that describe the spatial extent of an activity
-PUT	locative complement of *put*
-DTV	dative object if dative shift possible (e.g., *give*
-BNF	benefactive (dative object of *for*
-PRD	non VP predicates
untagged PPs	
Adjunct-like PP	
-DIR	direction and trajectory
-LOC	location
-MNR	manner
-PRP	purpose and reason
-TMP	temporal phrases

Table 2: Grammatical functions and semantic tags of PP constituents in the treebanks

3.2.3 Manner Place Time

In the treebanks, certain PPs have function tags that denote manner (PP-MNR), place (PP-LOC) or time (PP-TMP). We restricted our analysis to sentences that have both PPs annotated with these function tags. For English, we calculated whether the ordering of the two PPs follow MPT. For Chinese, we computed whether the ordering of the two PPs folow TPM.

3.3 Logistic Regression Models

We further compare the predictive power of dependency length and argument status in PP ordering with logistic regression modeling, which has been widely applied to model structural preferences (Bresnan and Ford, 2010; Gulordava et al., 2015; Levy and Jaeger, 2007; Morgan and Levy, 2015; Wasow et al., 2011). We did not include the rule of MPT in the model as the number of cases where it applies is quite small (see Section 4.5). Following similar methods in Rajkumar et al. (2016), we trained the logistic

regression models to predict the original observations in the corpora. For each model, we evaluated its prediction accuracy with Monte Carlo permutation test for 10,000 iterations.

We randomly selected half of the original instances extracted from the corpora and left them the way they were. For the other half, we constructed their structural variants simply by switching the order of the two PPs. Hence for the dataset of each corpus, half of the sentences are the originals while the other half are the constructed variants. The outcome binary variable is the ordering of the two PPs, represented as *Order*. We code *Order* as 1 for all original sentences, and 0 for all variants. Dependency length and argument status are included as the predictors in the model. For dependency length, we code it as 1 when the shorter PP is closer to the head verb, -1 when the longer PP is closer, and 0 when the two PPs have the same length. For argument status, we code it as 1 when the argument-like PP appears closer, -1 when the adjunct-like PP occurs closer, and 0 when the argument status of the two PPs is the same. A summary of our coding for the predictors is presented in Table 3.

| Factor	1	-1	0
dependency length	short PP closer	long PP closer	equal length
argument status	argument-like PP closer	adjunct-like PP closer	same argument status

Table 3: Coding for Predictors in Logistic Regression Models

4 Results & Analysis

4.1 PP Ordering in Chinese

Previous work has demonstrated empirically that Chinese has a dominant SVO order (Sun and Givón, 1985; Liu et al., 2009; Mei, 1980). Nevertheless, compared to English, which has more consistent head-dependent orderings, the headedness of different structures in Chinese is profoundly inconsistent. The adposition system in Chinese has been argued and shown to have both prepositions and postpositions (Hawkins, 1994).

The VP instances that fit our search criteria in CTB (i.e. cases with exactly two PP dependents attached to the same side of the head verb) appear as (5), where two head-initial PPs are placed before the head verb. Different from the PP orderings in English (see Section 2.1), where both the VP and the PP are head-initial, here we observed inconsistent headedness between the VP and the PP. Though based on predictions by DLM, the structure of (5a) will be more preferred to that of (5b), as the shorter PP is closer to the head verb in (5a). Nevertheless, when the head verb has head-initial PP dependents, to derive optimal overall dependency lengths, the PPs should occur after, rather than before the head verb like Chinese. In the cases below, the longest dependency length between the first PP and the head verb is already incurred regardless of the orderings of the two PPs, so it may not matter as much whether the shorter PPs are closer to the head verb or not. Accordingly, we expect there to be much weaker or even no effect for dependency length in PP orderings in Chinese.

(5) a. 他们 将 [在 生产 电子产品 方面] [和 中国] 进行 互利 合作
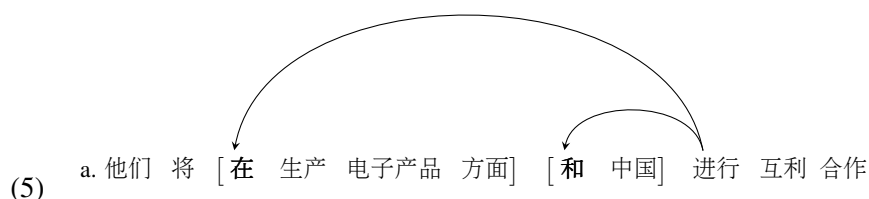
They will [in the aspects of electronic device production] [with China] conduct mutually beneficial collaboration.

b. 他们 将 [和 中国] [在 生产 电子产品 方面] 进行 互利 合作
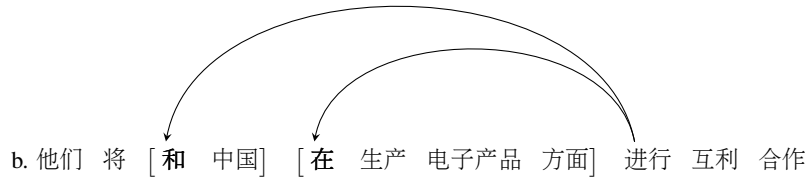
They will ⌈ with China ⌉ ⌈ in the aspects of electronic device production ⌉ conduct mutually beneficial collaboration.

They will collaborate in a mutually beneficial fashion with China in the production of electronic devices.

4.2 Effect of dependency length

As shown in Figure 1[3], the order predicted by DLM is strongly preferred in English. The number of sentences that have the shorter PP closer to the verb is 1.8 to 3.5 times larger than the number of sentences that have the longer PP closer to the verb. However, in roughly 20% of all sentences, DLM makes no prediction, since the two PPs have the same number of tokens. Although these numbers suggest that the preference for DLM is not as strong in spoken data as it is in written text, the preference for shorter dependencies is substantial across all three domains.

On the contrary, Chinese shows only a mild tendency for DLM. The number of cases when the shorter PP appears closer is not significantly much larger than that of instances when the longer PP is closer. This aligns with what we expected originally, that when inconsistent headedness exists between the VP and the PP, as in Chinese, dependency length does not seem to play a strong role.
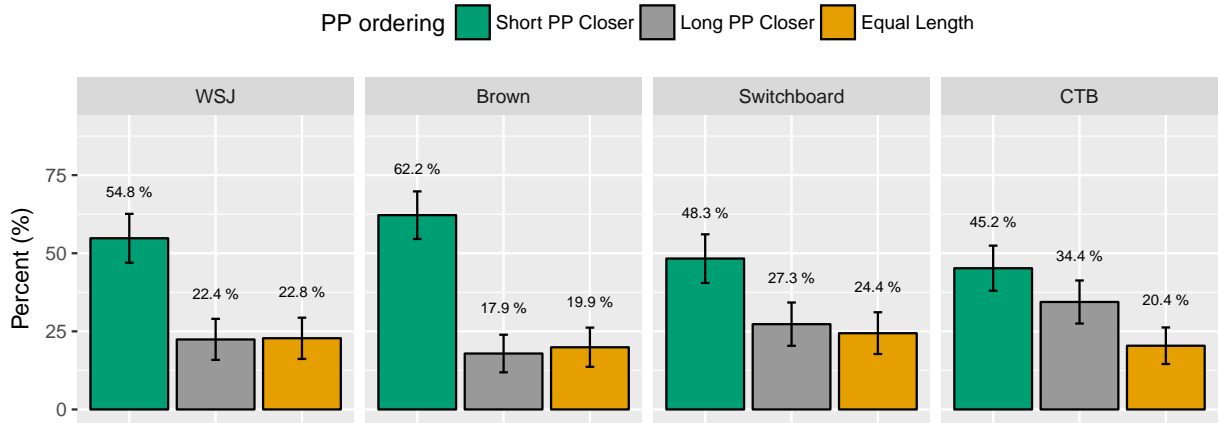

Figure 1: Effect of dependency length

To acquire a better understanding of why the efficacy of DLM is weaker in spoken genre than in written texts for English, we took a closer examination at the PP lengths of the extracted instances from the three corpora in PTB. We conjectured two possible reasons. First, compared to written texts, the average PP length for spoken genre might be much shorter. Second, spoken data might have more cases where the length difference between the two PPs is relatively small. Both indicate it might be less necessary to put the shorter PP closer to the head verb in Switchboard, leading to overall weaker preference for DLM. To test our conjectures, we computed the average PP lengths as well as the number of cases where the lengths of the two PPs differ by only 1-2 words. Nevertheless, as shown in Table 4, the average PP length in Switchboard is comparable to that in Brown, and only mildly shorter than that of WSJ (by 0.3 word). The proportion of cases where the two PPs have small length difference in Switchboard is similar to Brown, while slightly higher than WSJ (by 1.2%). This suggests that there are other potential constraints

---

[3]We estimated effects of dependency length after removing punctuation, as well as repetition in Switchboard for comparison, which did not appear to affect the results much. Thus we included punctuation for our calculation.

possibly competing with dependency length and working in different directions. They play stronger roles in the spoken than written domains in English and have overruled the impact of dependency length.

Corpus	Average PP length	% with small PP length difference
WSJ	5.4 ± 0.6	34.7 ± 6.7
Brown	4.7 ± 0.6	42.8 ± 6.9
Switchboard	4.0 ± 0.5	49.5 ± 6.9

Table 4: Comparisons of PP lengths

4.3 Effect of argument status

Consistent across domains for English and for Chinese, there appears to be a strong preference for argument-like PP to be close to the head verb. The number of instances where an argument-like PP is more adjacent is 1.5 to 2.7 times larger than when the adjunct-like PP occurs closer. By comparison, argument status has stronger effect in Switchboard and CTB than WSJ and Brown.
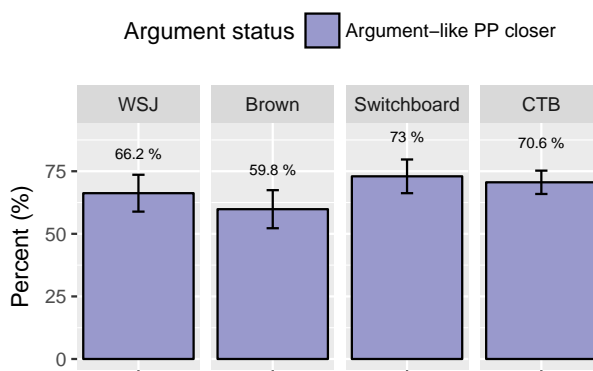

Figure 2: Effect of argument status of the PP

Now it is natural to ask how argument status interacts with dependency length pertaining the order of the two PPs. We estimated and compared the effects of argument status in sentences when the shorter PP appears closer versus when the longer PP is closer. In particular, in cases where shorter PPs are closer, it might matter less whether these shorter PPs are argument-like or not, since dependency length is already exerting a positive effect. Comparatively, in instances where longer PPs are closer, it is possible that most of the longer PPs are arguments of the verb, and tend to be more adjacent. Though results from Figure 3 do not align exactly with our initial thoughts, we observe some interesting patterns. In WSJ, when the longer PPs are close, the number that those longer PPs are arguments of the head verbs is significantly much larger. The preference for argument-like PP to be adjacent even when it is the longer PP suggests that when dependency length and argument status have the opposite effect, there will be strong competition between the two factors. This indicates that in WSJ, dependency length and argument status might have comparable predictive power in deciding what the PP ordering will be. In Switchboard, on the other hand, the argument-like PP is more adjacent to the head verb regardless of whether it is the shorter or the longer PP. The consistently pronounced effect for argument status here suggests that it might bear a stronger role than dependency length. When the two factors are pulling in different directions, the PP ordering might abide more by predictions of argument status than of DLM. However, in Brown, the number of the argument-like PPs being close is not much higher than by chance in spite of its length. This suggests there might be more cooperation rather than competition between the two constraints.

4.4 Cooperation and competition between dependency length and argument status

To further compare the cooperation and competition between the dependency length and argument status, we turn to evaluate and quantify the predictive power of the two factors with logistic regression mod-
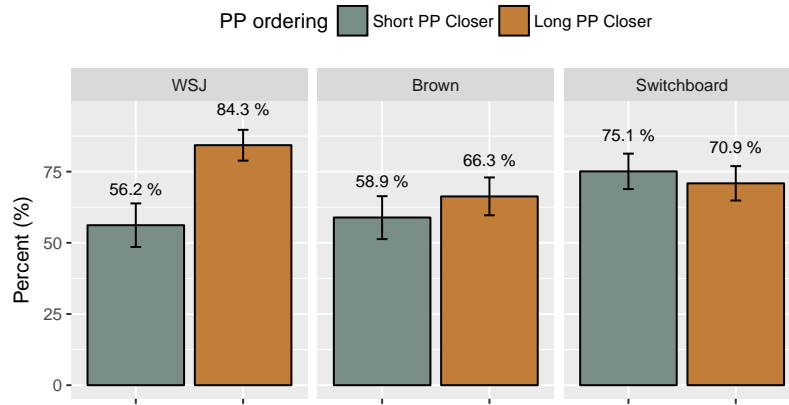
Figure 3: Effect of argument status when short vs. long PP is closer

els. We examined cases where at least one of the two constraints has an effect. Results from Figure 4 demonstrate that dependency length and argument status cooperate as well as compete with each other to different extents in the treebanks. The relative strengths of the two factors vary across domains in English and across the two languages. The most strongly preferred order is when the PP that is both shorter and argument-like to be adjacent to the head verb. On the other hand, competition between the two factors arise when they pull in the opposite directions (i.e. when the shorter PP is an adjunct or when the longer PP is an argument). The comparable predictive power for the two constraints in WSJ speaks to what we suggested earlier (see Section 4.3), that there is strong competition when dependency length and argument status are working against each other. In Brown, dependency length appears to be more predictive than argument status, indicating that the shorter PP is still more likely to be closer even when it is not an argument. In other words, the PP orderings in Brown will align more with predictions by DLM. In both Switchboard and CTB, argument status has a more pronounced role. This contrast suggests that in these two corpora, the orders tend to put the argument-like PP adjacent to the head verb, even if it is the longer PP between the two PPs.
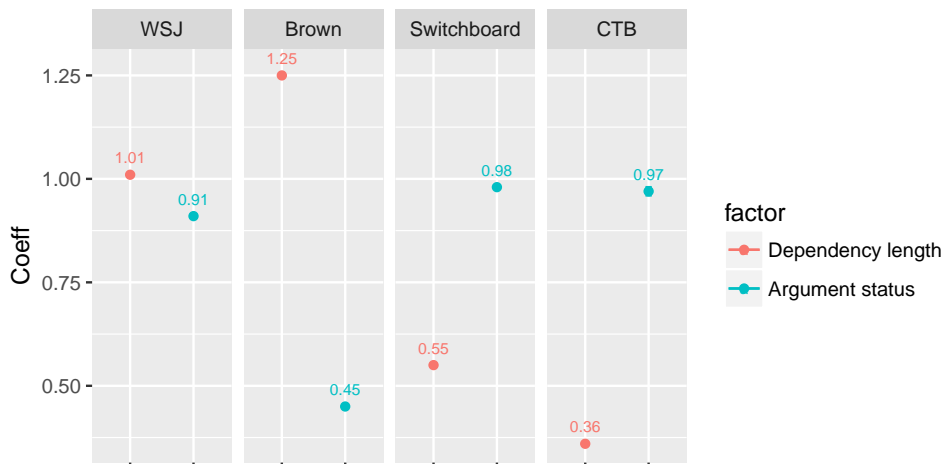

Figure 4: Coefficient estimates for dependency length and argument status in each corpus. Dots show point estimates with bars indicating standard errors.

4.5 Effect of Manner Place Time

Though there were not enough cases that TPM applies in Chinese, we found a significant role for MPT in English. This differs from Hawkins (1999) and Wiechmann and Lohmann (2013), which have shown no or weak effect for MPT, respectively. It is possible that their results are due to the use of smaller language

samples. In our dataset, MPT applies to about 6% of all instances. Within this set, it correctly accounts for the order of 89.3% of sentences in WSJ, 100% in Brown, and 100% in Switchboard. However, because it applies so infrequently, its overall impact is much smaller than that of dependency length and argument status.

5   Discussion & Conclusion

We analyzed the effects of dependency length, argument status and MPT in PP orderings for both English and Chinese. Consistent with previous studies, dependency length serves as a strong predictor for PP ordering across domains in English. Nevertheless, it only exerts a mild effect in Chinese. This relates to previous studies, which have shown that whether the preference for DLM exists and its efficacy are dependent on the headedness of the specific structures for languages with different typological characteristics (Lohmann and Takada, 2014; Faghiri and Samvelian, 2014; Faghiri et al., 2014). The argument status of the PP also has a pronounced effect on the orderings. It appears to play a comparable or even stronger role when compared to dependency length using logistic regression modeling. Overall, our results provide direct and quantitative evidence that dependency length and argument status are competing and cooperating motivations in PP ordering preferences across English and Chinese.

As effective as dependency length and argument status are, it is clear that around 30% of the data in English and around 40% of the data in Chinese remain unexplained based on model prediction accuracy presented in Table 5. Other constraints and their interactions with dependency length and argument status await to be discovered. One other factor that has been addressed previously on PP orderings is pragmatic information status (Hawkins, 1999; Wiechmann and Lohmann, 2013). Though Hawkins (1999) found no significant role for pragmatic information, it seems to have a mild effect based on results from Wiechmann and Lohmann (2013).

Corpus	Accuracy (%)
WSJ	67.3 ± 0.03
Brown	73.4 ± 0.03
Switchboard	66.4 ± 0.04
CTB	63.3 ± 0.10

Table 5: Model prediction accuracy with dependency length and argument status

Finally, previous experiments have presented contrary evidence regarding whether shorter dependencies will facilitate processing in Chinese relative clauses. Different from the head-initial relative clause structure in English, the head noun of relative clauses in Chinese comes in the final position. This results in longer dependencies in subject-extracted (SR) than object-extracted relative clauses (OR), whether the relative clause is modifying the subject or the object of the sentence. Certain studies have found that ORs are easier to process than SRs (Gibson and Wu, 2013; Hsiao and Gibson, 2003), providing support for predictions by DLM. On the other hand, findings from others have shown significantly shorter reading times for SRs for both adults (Hsiao and MacDonald, 2013; Vasishth et al., 2013; Chen et al., 2012; Chen et al., 2010) and children of different ages (Hu et al., 2016). As argued in Jäger (2015), expectation-based accounts are able to offer more thorough explanations. It is possible that SRs are processed faster due to its higher conditional probabilities given the preceding context in the sentence. Following this line of thought, it is likely that one's probabilistic knowledge of the grammars for a language as well as the overall structural distributions of the language affect constituent ordering preferences. Extensions of these predictions to word order variations across languages will lead to a more fruitful research direction.

Acknowledgements

References

Jennifer E Arnold, 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.

Ann Bies, Mark Ferguson, Karen Katz, Robert MacIntyre, Victoria Tredinnick, Grace Kim, Mary Ann Marcinkiewicz, and Britta Schasberger. 1995. Bracketing guidelines for treebank ii style penn treebank project. *University of Pennsylvania*, 97:100.

Joan Bresnan and Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in american and australian varieties of english. *Language*, 86(1):168–213.

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

Zhong Chen, Qiang Li, Kuei-Lan Kuo, and Shravan Vasishth. 2010. Processing chinese relative clauses: Evidence for the universal subject preference. *Unpublished manuscript*.

Zhong Chen, Kyle Grove, and John Hale. 2012. Structural expectations in chinese relative clause comprehension. In *Proceedings of the 29th West Coast Conference on Formal Linguistics*, pages 29–37. Cascadilla Proceedings Project Somerville, MA.

Hye-Won Choi. 2007. Length and order: A corpus study of korean dative-accusative construction. *Discourse and Cognition*, 14(3):207–227.

Peter W Culicover, Ray S Jackendoff, Ray Jackendoff, et al. 2005. *Simpler syntax*. Oxford linguistics. Oxford University Press.

William Edward Dyer. 2017. *Minimizing Integration Cost: A General Theory of Constituent Order*. Ph.D. thesis, University of California, Davis.

Pegah Faghiri and Pollet Samvelian. 2014. Constituent ordering in persian and the weight factor.

Pegah Faghiri, Pollet Samvelian, and Barbara Hemforth. 2014. Accessibility and word order: The case of ditransitive constructions in persian. In *The 21th International Conference on Head-Driven Phrase Structure Grammar*, pages pages–217.

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

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

Edward Gibson, Richard Futrell, Steven Piantadosi, Isabelle Dautriche, Kyle Mahowald, Leon Bergen, and Roger Levy. In Press. How efficiency shapes human language. *Trends in Cognitive Sciences*.

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

Edward Gibson. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. *Image, language, brain*, pages 95–126.

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

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. Association for Computational Linguistics.

Daniel Gildea and David Temperley. 2010. Do grammars minimize dependency length? *Cognitive Science*, 34(2):286–310.

John J. Godfrey, Edward C. Holliman, and Jane McDaniel. 1992. Switchboard: Telephone speech corpus for research and development. In *1992 IEEE International Conference on Acoustics, Speech, and Signal Processing. ICASSP*, pages 517–520. IEEE.

Kristina Gulordava and Paola Merlo. 2015. Structural and lexical factors in adjective placement in complex noun phrases across romance languages. In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*, pages 247–257.

Kristina Gulordava, 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 (Short Papers)*, pages 477–482.

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

John A. Hawkins. 1999. The relative order of prepositional phrases in english: Going beyond manner–place–time. *Language variation and change*, 11(3):231–266.

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

John A. Hawkins. 2014. *Cross-linguistic variation and efficiency*. OUP Oxford.

Franny Hsiao and Edward Gibson. 2003. Processing relative clauses in chinese. *Cognition*, 90(1):3–27.

Yaling Hsiao and Maryellen C MacDonald. 2013. Experience and generalization in a connectionist model of mandarin chinese relative clause processing. *Frontiers in psychology*, 4:767.

Shenai Hu, Anna Gavarró, Mirta Vernice, and Maria Teresa Guasti. 2016. The acquisition of chinese relative clauses: contrasting two theoretical approaches. *Journal of Child Language*, 43(1):1–21.

Richard Hudson. 1995. Measuring syntactic difficulty. *Manuscript, University College, London*.

Ray S Jackendoff. 1977. *X-bar syntax: A study of phrase structure*. Linguistic inquiry monographs 2. Cambridge, Massachusetts: MIT Press.

T. Florian Jaeger and Elisabeth J. Norcliffe. 2009. The cross-linguistic study of sentence production. *Language and Linguistics Compass*, 3(4):866–887.

Lena Jäger, Zhong Chen, Qiang Li, Chien-Jer Charles Lin, and Shravan Vasishth. 2015. The subject-relative advantage in chinese: Evidence for expectation-based processing. *Journal of Memory and Language*, 79:97–120.

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

Lars Konieczny 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*, pages 13–17.

Lars Konieczny. 2000. Locality and parsing complexity. *Journal of psycholinguistic research*, 29(6):627–645.

Henry Kučera and Winthrop Nelson Francis. 1967. *Computational analysis of present-day American English*. Dartmouth Publishing Group.

Roger Levy and T. Florian Jaeger. 2007. Speakers optimize information density through syntactic reduction. In *Proceedings of the 20th Conference on Neural Information Processing Systems (NIPS)*.

Roger Levy. 2013. Memory and surprisal in human sentence comprehension. In Roger P. G. van Gompel, editor, *Sentence Processing*, page 78–114. Hove: Psychology Press.

Haitao Liu, 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.

Arne Lohmann and Tayo Takada. 2014. Order in np conjuncts in spoken english and japanese. *Lingua*, 152:48–64.

Barbara Lohse, John A Hawkins, and Thomas Wasow. 2004. Domain minimization in english verb-particle constructions. *Language*, 80(2):238–261.

Brian MacWhinney, Andrej Malchukov, and Edith Moravcsik. 2014. *Competing motivations in grammar and usage*. OUP Oxford.

Mitchell P. Marcus, Mary Ann Marcinkiewicz, and Beatrice Santorini. 1993. Building a large annotated corpus of english: The penn treebank. *Computational linguistics*, 19(2):313–330.

Mitchell Marcus, Grace Kim, Mary Ann Marcinkiewicz, Robert MacIntyre, Ann Bies, Mark Ferguson, Karen Katz, and Britta Schasberger. 1994. The penn treebank: annotating predicate argument structure. In *Proceedings of the workshop on Human Language Technology*, pages 114–119. Association for Computational Linguistics.

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

Paola Merlo and Eva Esteve Ferrer. 2006. The notion of argument in prepositional phrase attachment. *Computational Linguistics*, 32(3):341–378.

Emily Morgan and Roger Levy. 2015. Modeling idiosyncratic preferences: How generative knowledge and expression frequency jointly determine language structure. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, page 1649–1654.

Carl Pollard and Ivan A Sag. 1994. *Head-driven phrase structure grammar*. University of Chicago Press.

Randolph Quirk, Sidney Greenbaum, and Geoffrey Leech. 1985. *A Comprehensive Grammar of the English Language*. Longman, London.

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

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

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

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

Russell S Tomlin. 1986. *Basic word order: Functional principles*. London: Croom Helm.

Shravan Vasishth and Richard L. Lewis. 2006. Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, 82(4):767–794.

Shravan Vasishth, Zhong Chen, Qiang Li, and Gueilan Guo. 2013. Processing chinese relative clauses: Evidence for the subject-relative advantage. *PloS one*, 8(10):e77006.

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

Thomas Wasow, T. Florian Jaeger, and David Orr. 2011. Lexical variation in relativizer frequency. *Expecting the unexpected: Exceptions in grammar*, pages 175–96.

Thomas Wasow. 1997a. End-weight from the speaker's perspective. *Journal of Psycholinguistic research*, (3):347–361.

Thomas Wasow. 1997b. Remarks on grammatical weight. *Language variation and change*, (1):81–105.

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

Naiwen Xue, Fei Xia, Fu-Dong Chiou, and Marta Palmer. 2005. The penn chinese treebank: Phrase structure annotation of a large corpus. *Natural language engineering*, 11(2):207–238.

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

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