A Dependency-Based Analysis on the Choices of Constituent Ordering

English Locative Inversion as an Example

CHEN Yixi[1] YAN Jianwei[2]

Department of Linguistics, Zhejiang University

2023-12-02

[1]Email: canjikhei1030@outlook.com
[2]Email: yanjianwei@aliyun.com

Dependency Distance Minimization (DDM)

The efforts of comprehension and production play an important role in explaining syntactic choices in human languages. When expressing a similar meaning, language users tend to structure their sentences to minimize the processing difficulty (Gibson et al., 2019; Hawkins, 1994, 2004; Temperley, 2007; Zipf, 1949).

Dependency Distance Minimization (DDM)

One of the important factors for syntactic choices is the tendency toward **dependency distance minimization** (DDM), which has been proven to be an effective predictor of constituent order (Ferrer-i-Cancho, 2008; Futrell et al., 2020; Rajkumar et al., 2016; Temperley, 2007).

$$DD_i = GN_i - DN_i \quad (1)$$

$$MDD = \frac{1}{n-1}\sum_{i=1}^{n}|DD_i| \quad (2)$$

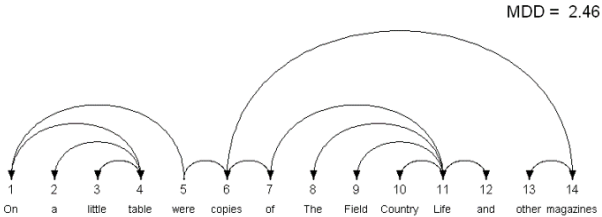

Figure 1. Dependency structure of the sample sentence

Short-Before-Long Preference

The **short-before-long preference** (Behaghel, 1909; Wasow, 2002), the tendency to order constituents from short to long, is advantageous to minimize MDD in head-initial languages (Hawkins, 1994; Temperley, 2007).

The ordering of constituents is constrained by its **heaviness**, a metric associated with

- the length or relative length (Hawkins, 2004; L. Stallings & MacDonald, 2011; L. M. Stallings et al., 1998)
- the syntactic complexity (Ross, 1967; Wasow, 1997; Wasow & Arnold, 2003)

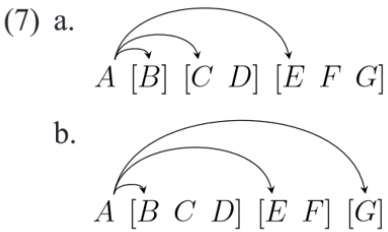

Figure 2. Short-before-long preference in dependency structure (Futrell et al., 2020, p.379)

Locative Inversion in English

Locative inversion refers to the inversion where the locative PP is placed in the preverbal position while the NP is after the verb.

- "Mismatch of role, function and category" (Bresnan, 1994)
  - NP: the role of subject with respect to agreement in a postverbal position
- Restricted choice of the verb (Birner, 1994; Culicover & Levine, 2001; Levin & Hovav, 1995)
  - Copula *be*, unaccusatives, unergatives, passives
  - Light inversion (limited to unaccusatives, see Culicover & Levine (2001)) and heavy inversion

(1) a. On a little table were copies of *The Field Country Life* and other magazines. (Copula *be*)
b. From the darkness beside him came a sudden wild shout of laughter. (Unaccusative, non-agent subject)
c. Behind him on the steps of the little office sat old man Arthur. (Unergative, agent subject)

Research Questions

Constituent order satisfying the short-before-long preference may be *non-canonical*, which renders some constructions, though causing less processing difficulty, less frequent in language usage.

The current study **examines the choices of constituent ordering in English locative inversion under the framework of dependency grammar**. We intend to answer whether the non-canonical ordering of NP and PP aligns with the short-before-long preference and to what degree such choice of constituent ordering lessens the processing difficulty.

Questions

1. Does locative inversion reflect the short-before-long preference in the ordering of NP and PP?
2. Can locative inversion reduce MDD compared to its canonical ordering alternation?

Data Collection

Language materials were retrieved from

- Brown in NLTK Corpora (N of token = 3,314,357)
- BNC (XML edition) on CQP Web (N of token = 112,102,325)

The procedure of data collection and pocessing:

- Searching (by regex and simple query)
- Manual filtering (locative inversion sentences within 50 words)
- Splitting and lemmatizing (by SpaCy 3.7.2)
- Dependency parsing (by SpaCy 3.7.2)

Table 1. General information on the self-built corpus

	Brown	**BNC**	**Total**
be	57	434	491
passive	7	16	23
unaccusative	17	35	52
unergative	20	44	64
Total	201	529	630

Data Extraction

(2) a. [$_{\text{PP}}$ On a little table] were [$_{\text{NP}}$ copies of *The Field Country Life* and other magazines].
(locative inversion sentence, IS thereafter)

b. [$_{\text{NP}}$ Copies of *The Field Country Life* and other magazines] were [$_{\text{PP}}$ on a little table].
(altered ordering sentence, AS thereafter)

- $LEN = 14$
- $MDD_{IS} = 2.46$
- $MDD_{AS} = 3.08$
- $HVLEN = \frac{LEN_{NP}}{LEN_{PP}} = \frac{9}{4} = 2.25$
- $HVMDD = \frac{MDD_{NP}}{MDD_{PP}} = \frac{2.33}{2} = 1.17$

Heavier NP in English Locative Inversion

Preliminary result

HVLEN and HVMDD were significantly larger than one, suggesting that NP in English locative inversion was significantly heavier than PP.

When not considering the difference among verb classes, NPs in English locative inversion sentences were **significantly heavier** than PPs, whether measured by HVLEN ($V = 178,054.5$, $p < .001$, $r_{rb} = .95$) or by HVMDD ($V = 125,744$, $p < .001$, $r_{rb} = .39$).

When considering the difference among verb classes, NPs in English locative inversion sentences were also significantly heavier than PPs ($p < .001$), **except in sentences with passive verb phrases when measured by HVMDD** ($p = .43$, $r_{rb} = .03$).

Heavier NP in English Locative Inversion

Light inversion (LI), in which the postverbal NP element can be phonologically and structurally extremely simple, possibly consisting of a single name, [...], is restricted to unaccusatives (Culicover & Levine, 2001, p. 283).

Although relative heaviness measured by LEN and MDD both captured the heavier NP in English locative inversion, HVMDD could better reflect **the distinction between light inversion** (related to unaccusative verbs) **and heavy inversion**.

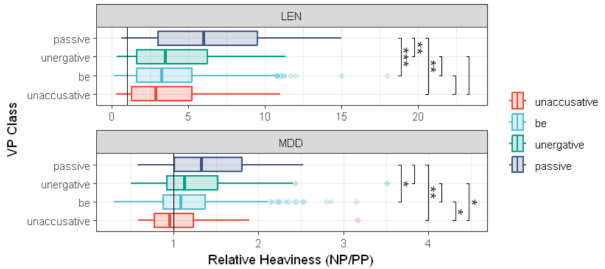

Figure 3. Relative heaviness of four verb classes measured by MDD and by LEN

Locative Inversion with Shorter MDD

Preliminary result

(Only) when NP in English locative inversion was heavier than PP, its postverbial position is preferred to shorten MDD.

Compared to its canonical ordering alternation (NP-VP-PP), English locative inversion, whose verb belongs to copula *be* ($p < .001$, r_{rb} = -.24), passives ($p = .02$, r_{rb} = -.49) and unaccusatives ($p = .02$, r_{rb} = -.32), yields **significantly smaller MDD**.

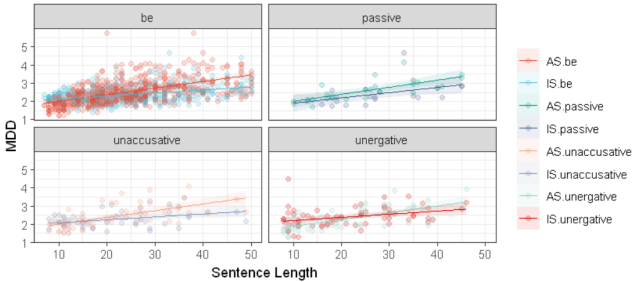

Figure 4. Comparison English locative inversion (IS) and its canonical ordering alternation (AS) in terms of MDD

Locative Inversion with Shorter MDD

When measuring relative heaviness with LEN, MDD **was significantly shortened only in English locative inversion with heavier NP** and with copula *be* ($p < .001$), unaccusative ($p = .01$) or passive ($p = .02$) as the verb.

For locative inversion sentences without heavier NPs, their MDDs **failed to be significantly shortened** compared to their canonical ordering alternations ($p > .75$).

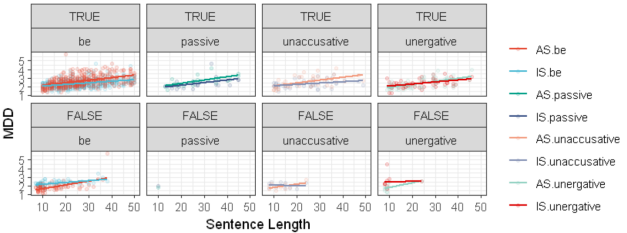

Figure 5. MDDs of English locative inversion (IS) and its canonical ordering alternation (AS) (grouped by whether HVLEN is greater than one)

Locative Inversion with Shorter MDD

Similar results were observed when measuring relative heaviness with MDD.

Another observation was **the lower frequencies of locative inversion sentences without heavier NPs** in English. A possible explanation is that the short-before-long preference motivates the postposing of the heavier NP; but for less heavier NPs, inversion is not so urgent and thus less frequent.

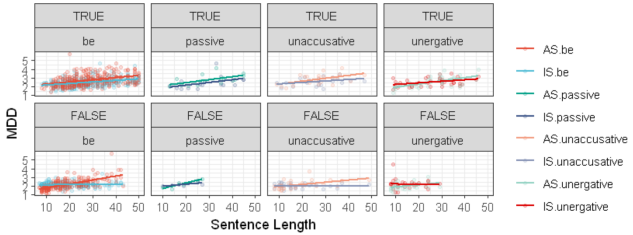

Figure 6. MDDs of English locative inversion (IS) and its canonical ordering alternation (AS) (grouped by whether HVMDD is greater than one)

Conclusion

Questions

1. Does locative inversion reflect the short-before-long preference in the ordering of NP and PP?
2. Can locative inversion reduce MDD compared to its canonical ordering alternation?

NP in English locative inversion was found to be significantly heavier than PP under most cases (except the case where the verb phrase is passive) in terms of either constituent length or syntactic complexity. In this sense, the postverbial position of NP **aligns with the short-before-long preference** in head-initial languages.

Specifically, results suggest that English locative inversion with unaccusative verbs has significantly smaller relative heaviness compared to the construction with other verb classes. Such finding is consistent with **the observation of light inversion** in Culicover & Levine (2001).

Conclusion

Questions

1. Does locative inversion reflect the short-before-long preference in the ordering of NP and PP?
2. Can locative inversion reduce MDD compared to its canonical ordering alternation?

English locative inversion **with heavier NPs presented significantly shorter MDD** compared to its canonical order alternation (except inversion whose verb is unergative). It corroborates the findings that placing easy constituents before relatively complex ones is advantageous to lessen the processing difficulty (Hawkins, 1994; MacDonald, 2013; Temperley, 2007).

In contrast, **no significant tendency towards shortening MDD** was observed for locative inversion sentences whose NPs were not heavier. They were also found **less frequent** than inversion sentences with heavier NPs.

Limitations

- Canonical ordering sentences (naturally occurring)
  - -> Is the low frequency of English locative inversion sentences caused by their greater MDDs compared to *naturally occurring* canonical ordering sentences?
- Explanation of the unergative verb class
  - -> Why English locative inversion sentences with unergatives show *no significant tendency* to shorten MDD?

Thanks for listening!

References I

Behaghel, O. (1909, December). Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen*. doi: 10.1515/9783110242652.110

Birner, B. (1994, June). Information Status and Word-Order - an Analysis of English Inversion. *LANGUAGE*, *70*(2), 233–259. doi: 10.2307/415828

Bresnan, J. (1994, March). Locative Inversion and the Architecture of Universal Grammar. *LANGUAGE*, *70*(1), 72–131. doi: 10.2307/416741

Culicover, P. W., & Levine, R. D. (2001, May). Stylistic inversion in English: A reconsideration. *NATURAL LANGUAGE & LINGUISTIC THEORY*, *19*(2), 283–310. doi: 10.1023/A:1010646417840

Ferrer-i-Cancho, R. (2008, June). SOME WORD ORDER BIASES FROM LIMITED BRAIN RESOURCES: A MATHEMATICAL APPROACH. *Advances in Complex Systems*, *11*(03), 393–414. doi: 10.1142/S0219525908001702

Futrell, R., Levy, R. P., & Gibson, E. (2020, June). Dependency Locality as an Explanatory Principle for Word Order. *Language*, *96*(2), 371–412. doi: 10.1353/lan.2020.0024

References II

Gibson, E., Futrell, R., Piantadosi, S. P., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019, May). How Efficiency Shapes Human Language. *Trends in Cognitive Sciences*, *23*(5), 389–407. doi: 10.1016/j.tics.2019.02.003

Hawkins, J. A. (1994). *A Performance Theory of Order and Constituency*. Cambridge University Press.

Hawkins, J. A. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Levin, B., & Hovav, M. R. (1995). *Unaccusativity: At the syntax-lexical semantics interface* (No. 26). Cambridge, Mass: MIT Press.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*. doi: 10.3389/fpsyg.2013.00226

Rajkumar, R., van Schijndel, M., White, M., & Schuler, W. (2016, October). Investigating locality effects and surprisal in written English syntactic choice phenomena. *Cognition*, *155*, 204–232. doi: 10.1016/j.cognition.2016.06.008

References III

Ross, J. R. (1967). *Constraints on variables in syntax*. (Thesis). Massachusetts Institute of Technology.

Stallings, L., & MacDonald, M. (2011). It's not Just the "Heavy NP": Relative Phrase Length Modulates the Production of Heavy-NP Shift. *Journal of Psycholinguistic Research*. doi: 10.1007/s10936-010-9163-x

Stallings, L. M., MacDonald, M. C., & O'Seaghdha, P. G. (1998, October). Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *JOURNAL OF MEMORY AND LANGUAGE*, *39*(3), 392–417. doi: 10.1006/jmla.1998.2586

Temperley, D. (2007, November). Minimization of dependency length in written English. *Cognition*, *105*(2), 300–333. doi: 10.1016/j.cognition.2006.09.011

Wasow, T. (1997, March). Remarks on grammatical weight. *Language Variation and Change*, *9*(1), 81–105. doi: 10.1017/S0954394500001800

Wasow, T. (2002). *Postverbal behavior* (No. no. 145). Stanford, Calif: CSLI Publications.

References IV

Wasow, T., & Arnold, J. (2003, December). Post-verbal constituent ordering in English. In G. Rohdenburg & B. Mondorf (Eds.), *Determinants of Grammatical Variation in English* (pp. 119–154). DE GRUYTER MOUTON. doi: 10.1515/9783110900019.119

Zipf, G. K. (1949). *Human behavior and the principle of least effort.* Oxford, England: Addison-Wesley Press.