有限理性视角下印地语中的依存长度最小化

SidharthRanjan (sidharth.ranjan03@gmail.com)斯
图加特大学，德国斯图加特 70174

Titus vonder Malsburg(titus.von-der-malsburg@ling.uni-stuttgart.de)斯图
加特大学，德国斯图加特 70174

arXiv:2304.11410v1 [cs.CL] 22 Apr 2023

摘要

依存长度最小化原则旨在使句子中句法相关的词语彼此靠近，人们认为这一原则普遍塑造着人类语言的结构，以实现有效沟通。然而，依存长度最小化在人类语言系统中的应用程度尚未得到充分理解。已知在动词前，将长成分置于短成分之前；在动词后，则将短成分置于长成分之前，这样可以最大限度地缩短句子的总体依存长度。本研究检验了以下假设：与全局依存长度最小化相比，仅将最短的动词前成分置于主要动词旁边，能够更好地解释印地语（一种主宾谓语言）的语序偏好。我们将这种做法称为最省力策略，因为它能以较低成本缩短动词与其所有动词前依存成分之间的依存关系。因此，这种做法符合有限理性视角：决策受"快速而节俭"的启发式支配，而非通过搜索最优解来完成。与这一观点一致，我们的结果表明，与全局最小化依存长度相比，最省力策略能更好地解释印地语-乌尔都语树库语料库中的实际语料句。此外，在区分语料句与反事实变体这一任务中，我们发现，与总依存长度相比，最靠近主要动词的成分的依存长度和成分长度，更能有效预测某个句子是否出现在语料库中。总体而言，我们的发现表明，认知资源限制在塑造自然语言方面发挥着关键作用。

关键词： 印地语；语序；句法选择；局部性；语言产出；认知建模

引言

人类工作记忆被视为一种容量有限的系统，因此人们会倾向于尽可能缩短信息的保留时间，以减轻认知负荷。Simon（1982，1990，1991）提出了决策中的有限理性理念，该理念纳入了上述认知限制，并指出人类进行理性决策的能力已适应于作出满意而非最优的选择。在语言使用中，说话者的决策会受到作出决策时所处语境的影响——信息的可用性、认知资源，以及有限的反应时间。这种张力自然给说话者带来挑战，并促使他们采用特定策略来组织句子，以实现最高效的交流。例如，选择日常话语中最常见的语序模式，或将工作记忆中更容易提取或更显著的词语／短语放在句首。在本研究中，我们考察了印地语中的语序变异是否可以被视为有限理性约束下高效沟通压力的结果。

先前研究探讨了短语长度对句子成分排序的影响（Bock & Warren, 1985; Arnold, Wasow, Losongco, & Ginstrom, 2000; Hawkins, 1994; Yamashita & Chang, 2001; Choi, 2007; Faghiri & Samvelian, 2020）。研究结果表明，SVO 语言（如英语）倾向于采用短语在前、长语在后的语言结构，即较短的短语更容易被提取，因而会先于相对不易提取或较长的短语实现（Arnold et al., 2000）。相比之下，在 SOV 语言（如日语）中，长语在前、短语在后的排序更为常见；也就是说，较长的短语往往会移至较短短语之前，因为长短语语义内容丰富、概念显著性更强，因此会先于较短短语实现（Yamashita & Chang, 2001）。在此方面，Gibson 的依存局部性理论为 SOV 和 SVO 语言中的成分排序模式提供了统一解释（Gibson, 2000）。无论何种语言，该理论都倾向于选择能够使句子整体依存长度最小化的成分语序。在动词前，长语在前、短语在后；在动词后，短语在前、长语在后的成分排序能够使句子的整体依存长度最小化（Hawkins, 2004）。

依存长度最小化（DLM）原则被认为能够通过缩短句子中句法相关词语之间的距离，减轻说话者和听话者工作记忆的认知负荷（Gibson,1991；Hudson, 1995；Gibson,1998, 2000；Ferrer-i-Cancho,2004；Temperley, 2007）。研究表明，自然语言确实如此：与随机生成的依存长度相比，自然语言往往具有更短的总体依存长度（Futrell, Mahowald,& Gibson, 2015；Liu, Xu,&Liang, 2017；Temperley & Gildea, 2018；Futrell,Levy,& Gibson,2020）。因此，DLM假设已被证明是人类语言中高效语序的一项重要特征，这种高效性源于交际效率（Gibson etal., 2019）。然而，依存长度最小化在多大程度上应用于人类语言系统，目前尚未得到充分理解。这构成了我们的主要研究动机；为此，我们将有限理性的思想应用于 DLM，并探究其在多大程度上能够解释印地语中的语序偏好。

(1) a. 母亲=ERG 市场去进行时哭泣的孩子=ACC 太妃糖 母亲=ERG 市场 去进行时 哭=PROG 孩子=ACC 太妃糖 给予-PFV.F.SG 母亲在去市场时给了哭泣的孩子=一颗糖果。 b. 母亲=ERG 哭泣的孩子=ACC 市场 去进行时 太妃糖 给予 c. 哭泣的孩子=ACC 市场去进行时 母亲=ERG 太妃糖 给予

印地语是一种属于印欧语系的印度-雅利安语，其基本语序为主宾谓（SOV），并且拥有丰富的格标记系统（Kachru，1982）。与英语相比，印地语在句内词语排列方面表现出更大的灵活性，如例1所示。例1a包含4个动词前成分，被认为是最受偏好的句法选择，因为它最初出现在语料库中；相较之下，该句还可能有所有其他变体（4! = 24）。

在本研究中，我们检验以下假设：说话者采用最省力策略来确定印地语中的语序选择，而不是在全局范围内最小化依存长度。我们将<style id='1'>最省力策略 定义为：将最短的动词前成分放置在最主要动词附近，因为这是减少动词与其所有动词前依存成分之间依存关系的最经济方式（示例见图3）。对于印地语-乌尔都语树库语料库中的每个参照句（1a），我们通过随机重新排列其动词前成分，生成了不同的反事实变体（例如，例1b至1c列举了其中几种）。随后，我们对这些变体及其对应的语料库参照句进行了定量分析。接着，我们使用逻辑回归分类器，根据依存长度、动词前成分的成分长度以及句子的总体依存长度，将语料库参照句与人工生成的变体区分开来。选择成分层面的预测变量，主要是为了在有限理性的视角下，检验最省力策略（见图3d）相对于全局最小化约束（见图3b）的有效性。随着动词前成分数量的增加（例如从3个→ 5开始），可能的变体数量也会显著增加（3! = 6 → 5! = 120），从而使说话者在进行选择时负担过重。有限理性的说话者随后会采用一种启发式策略，寻找一个令人满意的解决方案，而不一定使总依存长度达到最小（Newell & Simon，1972年；Gigerenzer & Gaissmaier，2011年）。

我们这项工作的主要贡献在于：我们使用大规模自然语料数据，而非构造的实验刺激材料，展示了依存长度最小化对印地语词序选择的影响，并提供了对认知理论发展至关重要的跨语言证据（Jaeger & Norcliffe，2009年）。

本文其余部分安排如下。我们首先简要介绍用于检验该假设的数据和方法，随后呈现本研究的结果。最后，我们将讨论研究发现的更广泛意义，并总结全文及展望未来的研究方向。

方法

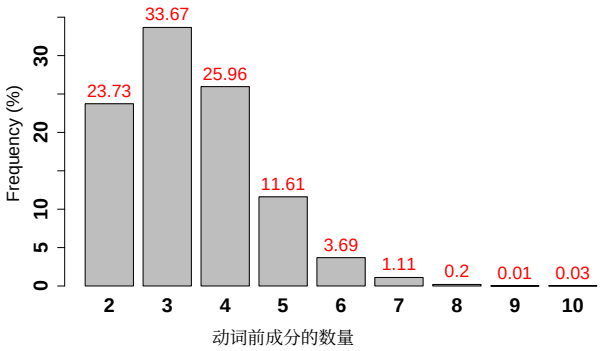
我们的数据集由 7,586 个陈述句组成，这些句子取自印地语-乌尔都语树库（HUTB）中属于新闻领域的书面文本语料库（Bhatt 等，2009年）。对于语料库中的每个参照句，我们通过随机排列动词前成分来创建人工变体；这些成分的中心词在其投射依存树中是根动词的直接子节点（见图 3）。对于变体超过 100 个的情况（为便于计算而任意设定的上限[1] ），我们从中随机选取 99 个。我们的变体生成方法为数据集中的参照句共生成了 184818 个相互竞争的变体。示例见例 1。图 1 展示了具有不同数量动词前成分的参照句-变体句对所占的百分比。由于动词前成分达到 7 个或更多的句子数据点不足，我们仅展示最多包含 6 个动词前短语的句子的语料库分析结果。不过，对于本文后文介绍的二元分类任务，我们使用整个数据集来计算结果 [2]。

语料库分析

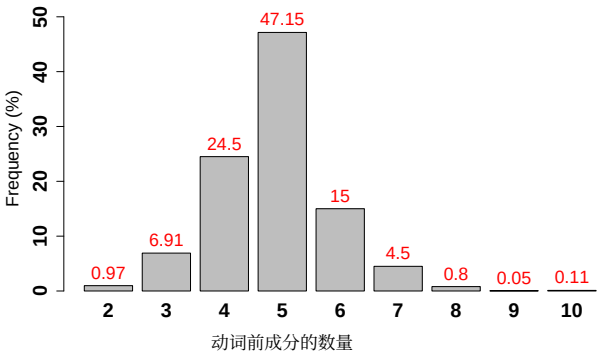
此前对印地语的语料库研究发现，动词前的成分语序以长在前、短在后为主，且中心语与依存项之间的距离较小（Sharma, Futrell, & Husain, 2020；Ranjan, Rajkumar, & Agarwal, 2022）。为检验我们的最省力策略假设，我们考察了所有动词前成分在缩短句子整体依存长度方面的作用。与先前研究一致，我们也发现长在前、短在后的成分语序在印地语中占主导地位（Ranjan et al., 2022）。此外，我们还观察到，在动词前成分数量不一的语料库句子中，长在前、短在后和短在前、长在后的成分分布广泛（见图2）。至关重要的是，图2中的图表表明，最靠近主要动词的成分有很强的倾向，成为句中所有动词前成分里最短的一个。这与最省力原则的预测一致，因为将最短的成分移至主要动词旁边，不仅能缩短该成分的依存长度，还能缩短主要动词所有其他动词前依存项的依存长度，从而实现有利的成本效益比。

有趣的是，图 2 还显示，将较短成分移至主要动词附近的压力会随着动词前成分数量的增加而增大，这表明说话者可能会动态平衡语言产出的成本与理解者的认知限制。[3] 为进一步检验说话者是否采用最省力策略（即将最短成分置于主要动词旁），我们比较了参照句（语料库中实际出现的句子）的总依存长度与四种成分语序不同的替代句。参照句出现在

[1]更高或更低的截断值不会影响我们的结果。[2]我们还仅使用语法变体进行了类似分析，发现所得结果与下文报告的结果一致。[3]句子长度与动词前成分数量之间的皮尔逊相关系数为 0.45。



(a) 参照句（7586 个数据点）



（b）变体句（184818 个数据点）

图1：包含不同数量动词前成分的参考句和变体句所占百分比

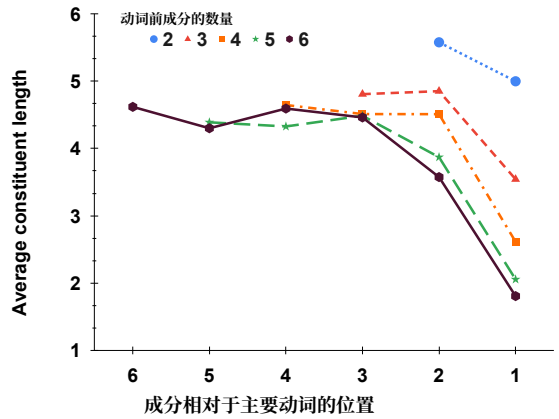


图2：语料库中包含2至6个成分的句子里，动词前成分的平均成分长度，分别统计

语料库中），这四种替代句如图 3 所示，以下以例句 1a 为例：

1. **升序排列**：按照动词前成分的成分长度递增的顺序排列。这种排列方式会使句子的依存长度达到最大值（见图 3a）。
2. **降序排列**：按照动词前成分的成分长度递减的顺序排列。这种排列方式使句子的依存长度达到全局最小值（见图 3b）。
3. **随机排列**：随机排列动词前成分（见图 3c）。4. **最小努力排列**：首先将动词前成分随机排列，然后只需将最短的成分移动到主要动词旁边（见图 3d）。

图 4 中的结果表明，对于动词前成分为三个或更多的句子，语料库中句子的依存长度总体上与最省力解的依存长度相符。只有动词前成分为两个的句子偏离了最省力语序。这可能是因为在这些情况下，记忆负担很低，系统根本不会优化依存长度，甚至不会采用最省力启发式。与以往研究一致，该图还证实，语料库中的参照句平均而言比随机词序的句子具有更短的依存长度（Liu, 2008；Gildea & Temperley, 2010；Futrell et al., 2015, 2020）。

这些结果进一步证实了这样一种观点：印地语中的依存长度最小化，以及由此决定的成分语序，受最省力策略支配。这与决策的有限理性观点一致。根据这一观点，我们会认为，说话者在规划句子时只会考虑有限的搜索空间（Gigerenzer、Hertwig 和 Pachur，2011年）。

在下一节中，我们进一步通过一项决策任务来检验并验证最省力假设，目标是区分参照句与随机生成的变体句（Gildea & Temperley, 2010; Temperley & Gildea, 2018）。

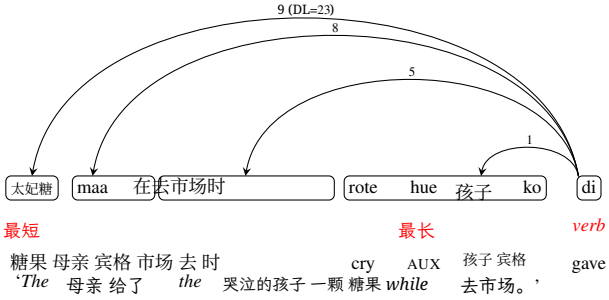
计算机模拟

我们设置了一项二元分类任务，以区分参照句与人工生成的变体句。原始数据集存在显著的类别不平衡，变体句有 184818 个，而参照句仅有 7586 个。为解决这一不平衡问题，我们采用了 Joachims（2002）提出的方法，将判断一个句子属于参照句还是变体句的分类任务，转化为将参照句与其各个变体句进行比较并排序的任务（成对排序）；具体做法是使用变体句与其对应的参照句之间的特征向量差值来训练分类器（见公式 1 和 2）。

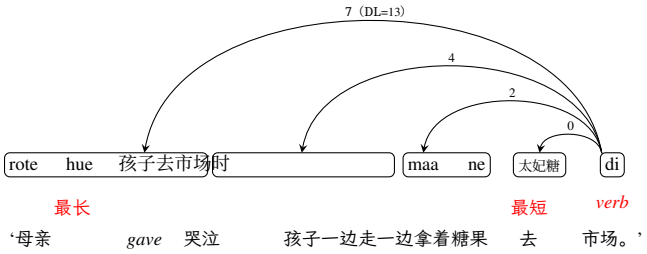
$$w \cdot \phi(reference) > w \cdot \phi(variant) \quad (1)$$

$$w \cdot (\phi(reference) - \phi(variant)) > 0 \quad (2)$$

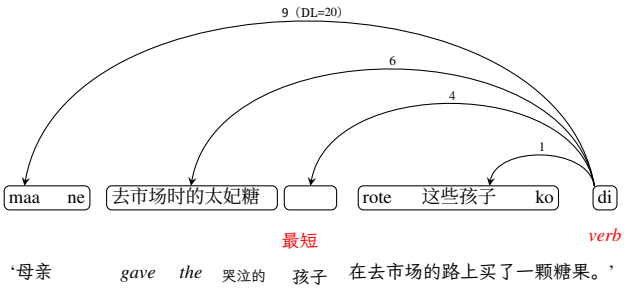
公式 1 表示标准二元分类器的目标函数：语料库中的参照句相较于其对应的变体句更受偏好。分类器必须确定一个特征权重（w），使得 w 与参照特征向量（ϕ（$reference$））的点积大于 w 与变体特征向量（ϕ（$variant$））的点积。该目标函数可以改写为公式 2，其中 w与特征向量之差的点积必须大于零。



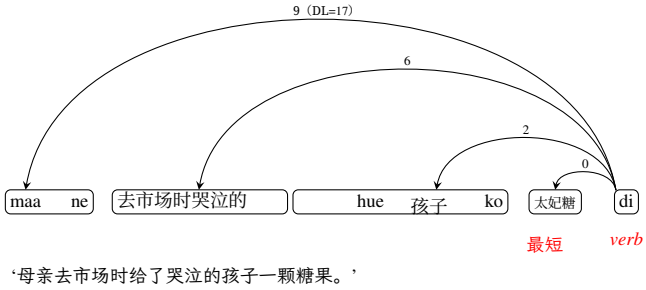
(a) 成分长度升序



（b）成分长度降序排列



（c）成分长度的随机顺序



（d）成分长度的最省力语序

图3：印地语中心语后置结构的依存长度与成分排序；图中仅展示主要动词依存关系；结构的总依存长度（DL）标注在各子图上方；各依存弧上方标注成分依存长度

数据集经过构建，使每个变体句都与其对应的参照句配对，并在这些配对中平衡句子的顺序。例如，示例 1 将生成（1a-1b）和（1c- 1a）。特征向量经过相减处理，并为每个转换后的数据点分配二元标签。“参照句-变体句”配对被标记为“1”，“变体句-参照句”配对被标记为“0”，从而平衡了此前不平衡的分类任务。经过 Joachims 的转换后，我们的分类任务总计包含 184818 个数据点。更多详情请参见 Ranjan 等（2022）。

我们使用 R 中的 glm 函数来检验开篇提出的假设（因变量 ~自变量）：选择 ~ δ 依存长度 和 选择 ~ δ 成分长度。这里，选择是一个二元选择因变量（1 表示偏好参照句，0 表示偏好变体句）。Delta（δ）指参照句与其配对变体的特征向量之间的差异。若语料库中的参照句优于与其配对的变体，分类器则预测标签“1”；若变体优于其配对参照句，则预测“0”。所有自变量均被标准化为z分数，即，将预测变量的取值（以其均值为中心）除以其标准差。我们估计了自变量

如下（示意见图 3）：

1. 总依存长度：句子依存树中中心语与依存项之间距离的总和。中心语与依存项之间的距离通过计算两者之间的间隔词数来估计。

2 成分依存长度：动词前成分的中心语与该成分所依存的主要词根之间跨越的词数。

3. 成分长度：成分中的总词数。

回归分析

本节检验我们的主要假设：在依存长度最小化的约束下，印地语说话者会采用最省力策略（即将最短的成分移至最靠近主要动词的位置），以优化语序选择。图 1b 展示了动词前成分数量不同的变体分布。由于我们的分类数据集主要由含有 5 个动词前成分的参考变体对构成，我们使用逻辑回归模型，在仅包含 5 个动词前成分的数据集转换版本（87143 个数据点）上检验最省力假说。我们估计了全部 5 个动词前成分的依存长度和成分长度，并将它们作为预测变量纳入回归模型（选择 ~ δ成分 1 +δ成分2 + δ成分3 + δ成分4 + δ成分5）。不过，在此之前

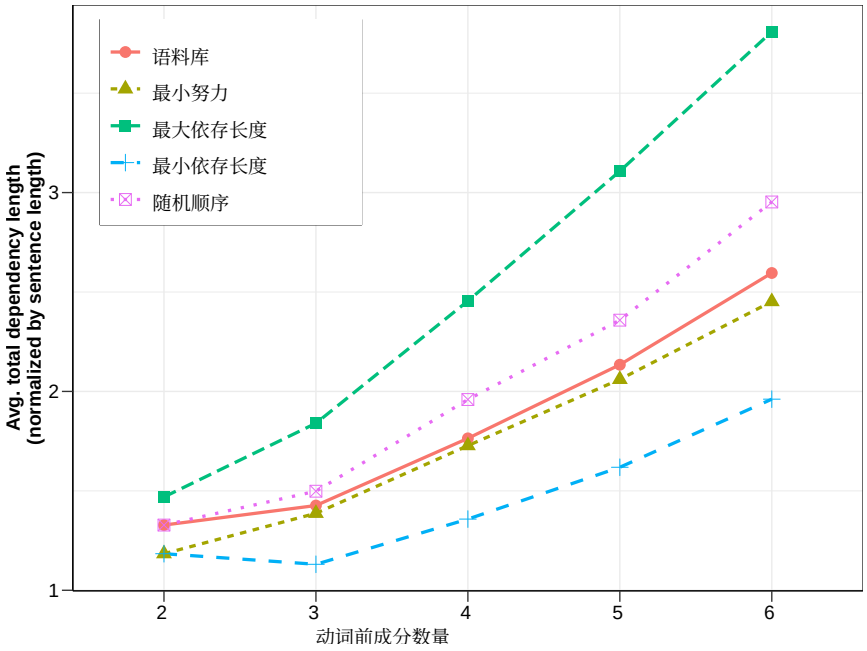


图 4：不同成分语序及不同数量动词前成分的平均总依存长度（即，总依存长度除以句子中的词数）（每种语序各有 7586 个数据点）

我们首先采用递归特征消除交叉验证算法（Guyon、Weston、Barnhill 和 Vapnik，2002；RFECV），从五个动词前成分中找出最佳预测变量，然后报告使用所得最优特征进行回归的结果。

表 1 和表 2 展示了回归分析的结果。表 1 和表 2 中为负的回归系数表明，与配对变体相比，参照句的成分层级依存长度和成分长度往往更短。模型中距离动词最近的最后两个动词前成分（成分4 和 成分5）的回归系数绝对值大于其他动词前成分，这表明，相较于配对变体中的成分，它们在缩短参照句的总体依存长度方面发挥着更大的作用。值得注意的是，对于含有 2 至 6 个动词前成分的参照句—变体句配对，我们的结果始终一致，这表明我们的发现并非仅适用于罕见结构，也适用于印地语自然语料库中的常见结构。总之，这些发现表明，说话者确实会运用最省力策略来优化语序，使距离动词最近的动词前成分更短、依存长度也更小。

预测变量	估计值	标准误	z 值
截距	**0.018**	0.008	1.98
1st 成分的依存长度	-0.004	0.012	-0.35
2nd 成分的依存长度	**0.088**	0.014	6.20
3rd 成分的依存长度	**-0.147**	0.015	-10.12
4th 成分的依存长度	**-0.529**	0.016	-32.90
5th 成分的依存长度	**-2.720**	0.023	-118.03

表 1：以动词前成分的依存长度（deplen）为预测变量的回归模型（87143 个数据点）；显著预测变量以粗体标示，$p < 0.001$

分类分析

基于上一节的见解，我们将最后两个动词前成分的依存长度和成分长度（最接近主要动词的成分：倒数第二个和最后一个成分）作为预测变量纳入分类模型，用于在反事实变体中预测参照句。我们使用完整的数据集（184818个参考变体对），并进行10折交叉验证，以评估模型的分类准确率，即模型正确预测配对变体中的指称句的数据点所占百分比。表3和表4展示了模型的预测性能。如表3所示，就单项分类准确率而言，最后一个成分的依存长度是表现最佳的预测变量（72.70%），从而验证了所提出的最省力原则的有效性。有趣的是，该度量的表现也超过了句子的总体依存长度度量（62.69%）。

在包含最后一个成分的依存长度的基线模型基础上，加入倒数第二个成分的依存长度后，分类准确率显著提高了 4.47%（使用 McNemar 双尾检验，$p < 0.001$）。如表 4 所示，将最后两个成分的成分长度作为分类器的预测变量时，我们也发现了类似结果。这些发现支持这样一种观点：说话者采用最省力策略，

预测变量	估计值	标准误	z 值
截距	-0.003	0.008	-0.41
1st 成分长度	**-0.083**	0.009	-8.44
3rd 成分长度	**0.058**	0.010	5.72
4th 成分长度	**-0.148**	0.009	-15.26
5th 成分长度	**-1.549**	0.016	-97.82

表 2：以动词前成分长度为预测变量的回归模型（87143 个数据点）；显著的预测变量以粗体标示，p $<$ 0.001；成分长度 =成分中的词数

预测变量	准确率
总依存长度	62.69
倒数第二个动词前成分的依存长度	68.48***
最后一个动词前成分的依存长度	72.70***
最后一个 + 2倒数第二个动词前成分的依存长度	77.17***

表3：以依存长度为预测变量的不同模型在完整数据集上的预测准确率（184818个数据点；deplen = 依存长度；与上一行比较的麦克内马尔双尾显著性检验：*** $p < 0.001$）

根据依存长度最小化衡量，在决定印地语语序偏好方面。

讨论与结论

我们的主要发现表明，印地语说话者通过将最短短语放在最靠近主要动词的位置来选择最优语序，这与有限理性决策观的核心理念一致。这种省力策略不仅能缩短整体依存长度，还能在庞大的搜索空间中实现选择性搜索，而在该空间中寻找最优解十分困难。因此，句子规划系统采用了一种简单而强大的启发式策略。此外，如图 2 所示，随着句子长度增加，缩短最靠近主要动词的成分所受到的压力也随之增大。我们还观察到，语料库中原本出现的句子，其依存长度分布与所提出的省力解法相似，这表明有限理性在选择恰当的印地语语序方面确有成效（见图 4）。此外，HUTB 语料库中的句子整体依存长度低于随机词序，这与早期研究的结论一致（Liu, 2008；Futrell 等，2015、2020）。

仅使用成分长度（70.28%），以及仅使用最接近主要动词的最后两个动词前成分的依存长度（77.17%），来区分参照句与人工生成的变体的两个不同分类模型的预测性能，凸显了所提出的最省力策略在塑造成分语序方面的重要性（Liu et al., 2017；Tem- perley & Gildea, 2018；Futrell et al., 2020）。此外，值得注意的是，总体依存长度（62.69%）在预测语料库参照句方面的表现较差（在

预测变量	准确率
倒数第二个动词前成分长度	54.35
最后一个动词前成分长度	69.62***
最后一个 + 2倒数第二个动词前成分长度	70.28***

表4：以成分长度为预测变量的不同模型的预测准确率（184818 个数据点；与上一行相比的麦克内马尔双尾检验：*** $p < 0.001$）

变体中）相比之下，最省力策略（最后一个动词前成分的依存长度 = 72.70%；最后一个动词前成分长度 = 69.62%）表现更好。作为未来工作的一个部分，我们计划考察所提出的最省力度量在不同语系的多种语言中的有效性。研究不同语言在实施有限理性策略的方式上可能存在何种差异，可以进一步深入了解语言与认知之间的关系。

我们的研究更广泛的启示是，语序选择可以被视为在有限理性的约束下，为实现高效沟通而承受各种压力的结果（Gigerenzer & Goldstein, 1996; Gigerenzer, 2015）。说话者试图尽量缩短依存长度，是应对认知资源有限和任务复杂性的一种节俭策略。鉴于我们往往缺乏精确解决问题所需的认知和计算资源，因此倾向于采用近似方法作为更可行的策略，从而形成有限理性（Simon, 1990; Gigerenzer & Selten, 2002）。Simon 指出，人们会运用以下机制，在有限理性的约束下处理复杂问题并作出决策——模式识别、启发式搜索，以及外推。同样地，在语言运用中，可以通过使用最常见的语序模式、在可能的语序中进行选择性搜索，将句法相关词语安排得彼此接近，或将显著／可预测的词语置于句首，以减轻记忆负荷，从而实现高效沟通。此外，说话者可以利用语言的某些特性，例如语序灵活性；这些特性能够为听者提供有助于提升理解的线索。从听者的角度看，语序灵活性可能构成挑战，而依靠各种语言特定线索，可能是他们在有限理性的约束下应对这一挑战的一种理性策略。因此，我们关于有限理性的研究结果和解释，既适用于说话者视角，也适用于听者视角。对这些视角进行深入考察，将有助于推动我们更好地理解这一问题。若要对一般意义上的语言产出提出更确切的论断，还需要考察最省力原则在口语数据中的效力。

尽管有限理性可能无法全面解释语言的各个方面，但它或许能解释某些方面，例如依存长度最小化的程度、语序模式、决策过程中的个体差异，以及参与其中的认知机制

语言使用。未来的研究需要考察不同因素（如语境、任务要求和个体认知差异）如何影响说话者和听者运用有限理性策略。

总的来说，我们的结果表明，在有限理性的约束下，依存长度最小化是印地语语序模式的重要预测变量。

致谢

我们谨向匿名的 CogSci-2023 评审者表示感谢，感谢他们提出的宝贵建议和反馈。

参考文献

Arnold, J. E., Wasow, T., Losongco, A., & Ginstrom, R.（2000）。显著性与新颖性：结构复杂性和话语地位对成分排序的影响。*Language*，*76*，28–55。取自 https://www.jstor.org/stable/417392 Bhatt, R., Narasimhan, B., Palmer, M., Rambow, O., Sharma, D. M., & Xia, F.（2009年）。用于印地语/乌尔都语的多表征、多层树库。载于第三届语言学标注研讨会论文集（第 186–189 页）。美国宾夕法尼亚州斯特劳兹堡：美国计算语言学协会。取自 http://dl.acm.org/ citation.cfm?id=1698381.1698417 Bock, J. K., & Warren, R. K.（1985）。句子生成中的概念可及性与句法结构。《认知》，*21*，47–67。取自 https://www.sciencedirect .com/science/article/pii/001002778590023X Choi, H.-w.（2007）。长度与顺序：韩语与格—宾格结构的语料库研究。话语与认知，*14*（3），207–227。取自 https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE00986720 Faghiri, P., & Samvelian, P.（2020）。主宾谓语言中的语序偏好与短语长度的影响：来自波斯语句子生成的证据。*Glossa*：普通语言学期刊（*2016—2021*），*5*（1），86。Ferrer-i-Cancho, R.（2004）。句法上相互连接的词之间的欧几里得距离。*Physical Review E*，*70*（5），056135。取自 https://core .ac.uk/download/pdf/ 286456939.pdf Futrell, R., Levy, R. P., & Gibson, E.（2020）。依存局部性作为语序的解释原则。*Language*，*96*（2），371–412。取自 https://sites.socsci.uci.edu/~rfutrell/papers/futrell2020dependency.pdf Futrell, R., Mahowald, K., & Gibson, E.（2015）。37 种语言中依存长度最小化的大规模证据。美国国家科学院院刊，*112*（33），10336-10341。取自 http://www.pnas.org/content/112/33/10336.abstract doi: 10.1073/pnas.1502134112 Gibson, E.（1991）。人类语言处理的计算理论：记忆限制与处理崩溃。卡内基梅隆大学。取自 摘自 http :// tedlab .mit .edu / tedlab website / researchpapers/Gibson 1991 PhDthesis.pdf Gibson, E. (1998). 语言复杂性：句法依存关系的局部性。《认知》，*68*，1–76。获取自 https://doi.org/10.1016/S0010-0277(98)00034-1 Gibson, E. (2000). 依存局部性理论：一种基于距离的语言复杂性理论。载于 A. Marantz、Y. Miyashita 和 W. O’Neil（编），图像、语言与大脑：首届心智表征项目研讨会论文集。马萨诸塞州剑桥：麻省理工学院出版社。获取自 http://www .ling .uni -potsdam .de/~vasishth/Papers/Gibson -Cognition2000.pdf Gibson, E., Futrell, R., Piandadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019, 5月1日). 效率如何塑造人类语言。《认知科学趋势》，*23*(5)，389-407。获取自 https:// doi.org/10.1016/j.tics.2019.02.003 doi: 10.1016/ j.tics.2019.02.003 Gigerenzer, G. (2015). 《简单理性：现实世界中的决策》。《进化与认知》。Gigerenzer, G., & Gaissmaier, W. (2011). 启发式决策。《心理学年度评论》，*62*，451–482。Gigerenzer, G., & Goldstein, D. G. (1996). 以快速而节俭的方式推理：有限理性模型。《心理学评论》，*103*(4)，650。Gigerenzer, G., Hertwig, R. E., & Pachur, T. E. (2011).启发式：适应性行为的基础。牛津大学出版社。Gigerenzer, G., & Selten, R. (2002). 有限理性：适应性工具箱。麻省理工学院出版社。Gildea, D., & Temperley, D. (2010). 语法是否会使依存长度最小化？《认知科学》，*34*(2)，286–310。Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). 使用支持向量机进行癌症分类的基因选择。《机器学习》，*46*，389–422。获取自 https://link .springer .com/article/10 .1023/A: 1012487302797 Hawkins, J. A. (1994). 语序与句法成分的表现理论。纽约：剑桥大学出版社。Hawkins, J. A. (2004). 语法中的效率与复杂性。牛津大学出版社。Hudson, R. A. (1995). 测量句法难度。伦敦大学学院。获取自 http://dickhudson .com / wp -content / uploads / 2013 / 07 / Difficulty .pdf Jaeger, T. F., & Norcliffe, E. (2009). 句子生成的跨语言研究：现状与行动倡议。《语言与语言学指南》，*3*(4)，866–887。获取自 http://dx.doi.org/10.1111/j.1749 -818X.2009.00147.x Joachims, T. (2002). 利用点击数据优化搜索引擎。载于第八届美国计算机协会 *SIGKDD* 国际知识发现与数据挖掘会议论文集（第133–142页）。美国纽约州纽约：美国计算机协会。获取自 http://doi .acm .org/10 .1145/775047

.775067 doi: 10.1145/775047.775067 Kachru, Y. (1982). 印地语-乌尔都语和波斯语中的并列动词。《南亚评论》，*6*(3)，117-126。获取自 https://doi.org/10.1080/02759527.1982.11933096 doi: 10.1080/02759527.1982.11933096 Liu, H. (2008). 依存距离作为衡量语言理解难度的指标。《认知科学杂志》，*9*(2)，159-191。获取自 http://www .lingviko . net/JCS.pdf Liu, H., Xu, C., & Liang, J. (2017). 依存距离：研究自然语言句法模式的新视角。《生命物理评论》，*21*，171 - 193。获取自 http://www.sciencedirect.com/science/article/pii/S1571064517300532 doi: https://doi.org/10.1016/j.plrev.2017.03.002 Newell, A., & Simon, H. A. (1972). 《人类问题求解》。新泽西州恩格尔伍德克利夫斯：Prentice-Hall。Ranjan, S., Rajkumar, R., & Agarwal, S. (2022). 印地语动词前成分排序中的局部性与预期效应。《认知》，*223*，104959。获取自 https://www.sciencedirect.com/science/article/pii/S0010027721003826 doi: https://doi.org/10.1016/j.cognition.2021.104959 Sharma, K., Futrell, R., & Husain, S. (2020). 印地语中动词依存成分的语序由什么决定？理解与产出中的效率效应。载于认知建模与计算语言学研讨会论文集（第1–10页）。获取自 https://aclanthology.org/2020.cmcl-1.1/ Simon, H. A. (1982). 有限理性模型，第1卷和第2卷。经济分析与公共政策，麻省理工学院出版社，马萨诸塞州剑桥。Simon, H. A. (1990). 人类行为的不变量。《心理学年度评论》，*41*(1)，1–20。获取自 https://www .annualreviews .org/doi/10.1146/annurev.ps.41.020190.000245 Simon, H. A. (1991). 认知架构与理性分析：评论。载于 K. VanLehn（编），智能架构：卡内基梅隆大学第二十二届认知研讨会（第25–39页）。Lawrence Erlbaum Associates, Inc. 获取自 https://apps.dtic.mil/sti/citations/ADA219199 Temperley, D. (2007). 书面英语中依存长度的最小化。《认知》，*105*(2)，300–333。获取自 http :// www .sciencedirect .com / science / article / B6T24 -4M7CDMS -2 / 2 / e095449f6439b30003822a5838e53786 doi: DOI:10.1016/j.cognition.2006.09.011 Temperley, D., & Gildea, D. (2018). 最小化句法依存长度：类型学/认知普遍性？《语言学年度评论》。获取自 https:// www .cs . rochester .edu/u/gildea/pubs/temperley - gildea-ar18.pdf Yamashita, H., & Chang, F. (2001). 句首优先偏好在中心语后置语言中的表现。《认知》*nition*，*81*。取自 https://www.sciencedirect . com/science/article/pii/S0010027701001214