

# A parsing theory of word order universals (Hawkins 1990)

Самира Ферхеев

Типология 2021 – Семинар

# Типология порядка слов составляющих

Гармонический порядок вершин и зависимых:

VO

OV

предлоги

послелог

NG

GN

NRel

RelN

...

...

# Типология порядка слов составляющих

Гармонический порядок вершин и зависимых:

VO	OV
предлоги	послелог
NG	GN
NRel	RelN
...	...

Отдельные пары характерны для определенных языковых семей ([Dunn et al. 2011](#))

# Типология порядка слов составляющих

Гармонический порядок вершин и зависимых:

VO	OV
предлоги	послелог
NG	GN
NRel	RelN
...	...

Отдельные пары характерны для определенных языковых семей (Dunn et al. 2011) / Тем не менее, экспериментальные данные показывают общее стремление к гармонии (Culbertson et al. 2020)

Hawkins (1990, 1994):

- ▶ Самые частотные паттерны гармонии < самые простые для парсинга
- ▶ Простота парсинга определяется скоростью, с которой слушающий может идентифицировать основные составляющие, и отнести их к материнскому узлу на дереве

- (1) I introduced to Mary some friends that John had brought to the party.
- (2) I introduced some friends that John had brought to the party to Mary.

- (3) I<sub>VP</sub>[introduced<sub>PP</sub>[to Mary]<sub>NP</sub>[some friends that John had brought to the party]].
- (4) I<sub>VP</sub>[introduced<sub>NP</sub>[some friends that John had brought to the party]<sub>PP</sub>[to Mary]].

- (3) I<sub>VP</sub>[introduced<sub>PP</sub>[to Mary]<sub>NP</sub>[some friends that John had brought to the party]].
- (4) I<sub>VP</sub>[introduced<sub>NP</sub>[some friends that John had brought to the party]<sub>PP</sub>[to Mary]].

Простота парсинга измеряется следующим образом: [количество обнаруженных составляющих / количество слов]



- (3) I<sub>VP</sub>[introduced<sub>PP</sub>[to Mary]<sub>NP</sub>[some friends that John had brought to the party]].
- (4) I<sub>VP</sub>[introduced<sub>NP</sub>[some friends that John had brought to the party]<sub>PP</sub>[to Mary]].

Простота парсинга измеряется следующим образом: [количество обнаруженных составляющих / количество слов]

$$(\textcolor{red}{3}) = 3/4 \text{ (75\% = 86\%)}; (\textcolor{red}{4}) = 2/11 \text{ (27\% = 47\%)}$$

# Early Immediate Constituents (EIC)

- ▶ The basic idea is that the parser prefers those orders of words that enable it to recognize all ICs of a mother node as rapidly as possible.
- ▶ The human parser prefers to maximize the left-to-right IC-to-word ratios of the phrasal nodes that it constructs.

(Hawkins 1990: 230, 233)

# Early Immediate Constituents (EIC)

Evidence for EIC comes from four major sources:

1. from real-time psycholinguistic experiments on alternative orderings of ICs in languages that permit such alternatives;
2. from text-frequency counts for these alternative orderings;
3. from native speaker acceptability judgments (which must be used with caution and in clear cases only);
4. and from the grammaticalized word orders of the world's languages (Hawkins 1990: 230)

# VP recognition domain

- (5) Joe looked up the number.
- (6) Joe looked the number up.
- (7) Joe looked the number of the ticket up.
- (8) Joe looked the number that Mary had forgotten up.

# VP recognition domain

- (5) Joe **looked up** the number.
- (6) Joe **looked** the number **up**.
- (7) Joe **looked** the number of the ticket **up**.
- (8) Joe **looked** the number that Mary had forgotten **up**.

Позиция частицы влияет на размер Constituent Recognition Domain для глагольной группы → влияет на EIC.

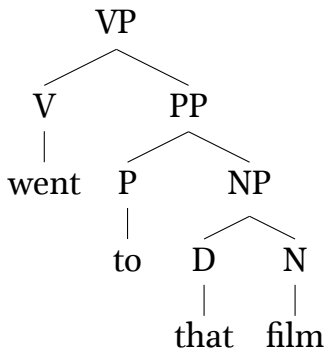
# VP recognition domain

- ▶ Чем дальше частица находится от глагола, тем меньше предложение нравится носителям (Hawkins 1990: 230)
- ▶ Расстояние частицы от глагола отрицательно коррелирует с частотностью конструкции в текстах, см. (Hawkins 1994: 180–214)
- ▶ (но выборка текстовых фрагментов очень маленькая, и не очень понятно как подобрана)

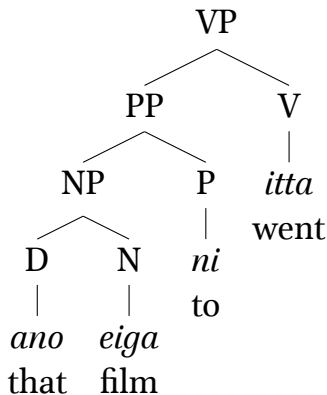
# Левое vs. правое ветвление

# Левое vs. правое ветвление

1. a.



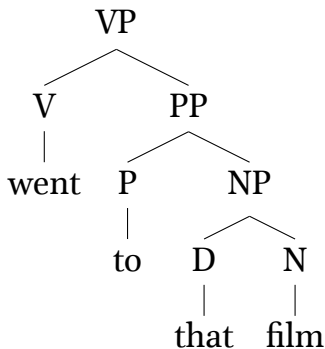
b.



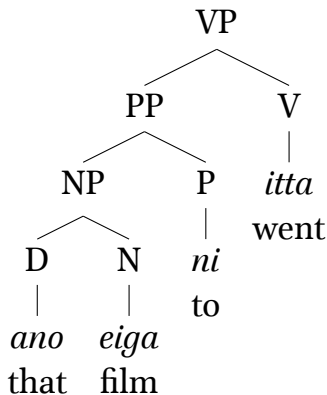


# Левое vs. правое ветвление

1. a.



b.



У обоих эффективность – 100%.

Нет доминирующего типа в языках мира.

# Ветвление и адлоги

	prepositions	postpositions
VO	+	-
OV	-	+

# Ветвление и адлоги

	prepositions	postpositions
VO	+	-
OV	-	+

Ожидаем, что языки в общем предпочитают сочетания с плюсом

# Ветвление и адлоги

	prepositions	postpositions
VO	+	-
OV	-	+

Ожидаем, что языки в общем предпочитают сочетания с плюсом

... а сочетания с минусом тоже встречаются, в результате неизвестных факторов, но значительно реже.

# Ветвление и адлоги

## Ожидание

	prepositions	postpositions
VO	+	-
OV	-	+

## Реальность<sup>1</sup>

	prepositions	postpositions
(S)VO	77%	23%
(S)OV	7%	93%

---

<sup>1</sup>Процент языков с порядком SVO или SOV + предлоги или послелогии, по выборке [Hawkins \(1983\)](#)

# Ветвление и адлоги

## Эффективность

	prepositions	postpositions
(S)VO	100%	79%
(S)OV	58%	100%

## Реальность

	prepositions	postpositions
(S)VO	77%	23%
(S)OV	7%	93%

# Тяжелые составляющие

- ▶ Более тяжелые модификаторы двигаются в конец, тогда как более легкие модификаторы предшествуют свою вершину (Hawkins 1983: 89–92)
- ▶ Heaviness hierarchy:  
Rel < Gen < Adj < Dem, Num

# Тяжелые составляющие

- ▶ Более тяжелые модификаторы двигаются в конец, тогда как более легкие модификаторы предшествуют свою вершину (Hawkins 1983: 89–92)
- ▶ Heaviness hierarchy:  
Rel < Gen < Adj < Dem, Num
- ▶ Тяжесть определяется количеством слогов, слов, и дальнейших составляющих, ср. релятивные клаузы, генитивные фразы, и указательные местоимения как однословные составляющие (Croft 2002: 70)



# Тяжелые составляющие

- (9) I introduced some friends to Mary.
- (10) I introduced some friends that John had brought to the party to Mary.

Экстрапозиция релятивной клаузы:

- (11) I introduced to Mary some friends that John had brought to the party.

# Тяжелые составляющие

- (12) Mary-ga [kinoo John-ga kekkonshi-ta]  
Mary-SBJ yesterday John-SBJ marry-PST  
to it-ta.  
COMP said-PST  
'Mary said that John got married yesterday.'  
Японский (Hawkins 1990: 231)

# Тяжелые составляющие

- (13) [kinoo John-ga kekkonshi-ta] to  
yesterday John-SBJ marry-PST COMP  
Mary-ga it-ta.  
Mary-SBJ said-PST  
'Mary said that John got married yesterday.'  
Японский (Hawkins 1990: 231)

# Тяжелые составляющие

- (13) [kinoo John-ga kekkonshi-ta] to  
yesterday John-SBJ marry-PST COMP  
Mary-ga it-ta.  
Mary-SBJ said-PST  
'Mary said that John got married yesterday.'  
Японский (Hawkins 1990: 231)

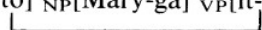
Сначала парсится зависимая клауза как отдельная, а сложная структура строится bottom-up, начиная с *to*.

# Тяжелые составляющие

$s_1[ \text{NP}[\text{Mary-ga}] \text{VP}[ \text{S}'[ \text{s}_2[\text{kinoo John-ga kekkonsi-ta}] \text{to}] \text{it-ta} ]].^1$



$s_2[ \text{S}'[ \text{s}_1[\text{Kinoo John-ga kekkonsi-ta}] \text{to}] \text{NP}[\text{Mary-ga}] \text{VP}[\text{it-ta} ]].^2$



# Тяжелые составляющие

$s_1[ \text{NP}[\text{Mary-ga}] \text{VP}[ \text{S}'[ \text{s}_2[\text{kinoo John-ga kekkonsi-ta}] \text{to}] \text{it-ta} ]]$ .<sup>1</sup>

$s_2[ \text{S}'[ \text{s}_1[\text{Kinoo John-ga kekkonsi-ta}] \text{to}] \text{NP}[\text{Mary-ga}] \text{VP}[\text{it-ta}] ]]$ .<sup>2</sup>

Общая тенденция в японском сдвинуть такие тяжелые составляющие вперёд (Dryer 1980).

# Асимметричная гармония

## Ожидание

	NRel	RelN
VO	+	-
OV	+	+

Статистика из **WALS** (исключая дополнительные варианты)

	NRel	RelN
VO	416	5
OV	113	132

# Асимметричная гармония

- ▶ У языков в общем есть тенденция к гармонии в порядке составляющих
- ▶ Существование менее оптимальных стратегий объясняется “другими факторами”
- ▶ По Хокинзу эти другие факторы тоже сводятся к EIC
- ▶ Например, RelN в целом неудобно, но если Rel можно легко узнать, например, из-за наличия какой-нибудь частицы или определенной формы предиката, то норм (Hawkins 1990: 252–257)



# Дискуссия

Почему предпочитается быстрый парсинг?

# Дискуссия

Почему предпочитается быстрый парсинг?  
– Principle of least effort?

# Дискуссия

Почему предпочитается быстрый парсинг?  
– Principle of least effort?

Почему быстрый парсинг слушающим влияет на  
производство речи?

# Литература I



Croft, William. 2002. *Typology and universals*.

Cambridge: Cambridge University Press.



Culbertson, Jennifer, Julie Franck, Guillaume Braquet, Magda Barrera Navarro & Inbal Arnon. 2020. A learning bias for word order harmony: Evidence from speakers of non-harmonic languages. *Cognition* 204. 104392.



Dryer, Matthew Synge. 1980. The positional tendencies of sentential noun phrases in universal grammar. *Canadian Journal of Linguistics* 25(2). 123–196.

# Литература II



Dunn, Michael, Simon J Greenhill, Stephen C Levinson & Russell D Gray. 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473(7345). 79–82. <https://www.nature.com/articles/nature09923>.



Hawkins, John A. 1983. *Word order universals*. New York: Academic Press.



Hawkins, John A. 1990. A parsing theory of word order universals. *Linguistic inquiry* 21(2). 223–261.

# Литература III



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