

# Juxtapozice „z – do“ v Arabštině

Kvantitativní přístup

Jiří Milička

Ústav srovnávací jazykovědy

# Jak do začalo?

První data pro výzkum byla shromážděna automaticky.

Ale když se ukázalo, že téma je zajímavé, začalo být jasné, že bez manuální práce to nepůjde.

# Úvod

ذهبت من بيتي إلى بيته.

Šel jsem ze svého domu do jeho domu.

ذهبت إلى بيته من بيتي.

Šel jsem do jeho domu ze svého domu.

- Nusus fi an-nahwi al-Arabiyyi 1, 2 (Texty o arabské gramatice). (8. – 14. sdoletí)
- Wright, W.: A Grammar of The Arabic Language (1896)
- Gulli, A. - Carter, M. - Badawi, M. A Modern written Arabic: Comprehensive Grammar. (2004)

- Sgall, P. – Hajičová, E. – Buráňová, E.: Topic – Comment in Czech. Prague 1980.

الآن أنا في بيتي ، سأذهب من بيتي إلى بيته.

Jsem ve svém domě. Půjdu ze svého domu do jeho domu.

الآن أنا في بيته ، ذهبت إلى بيته من بيتي.

Jsem v jeho domě. Přišel jsem do jeho domu ze svého domu.

الآن أنا في بيتي ، من بيتي سأذهب إلى بيته.

Jsem ve svém domě. Ze svého domu půjdu do jeho domu.

# Korpusy, které máme k dispozici:

- CLARA: Corpus Linguae Arabicae (synchronní, cca 40 000 000 tokenů)
- CLAUDIA: Corpus Linguae Arabicae Universalis Diachronis (diachronní, cca 300 000 000 tokenů)

Náhodně vybrané vzorky

1. Najít všechny výskyty *z a do* v rámci jedné věty (což je v arabštině obtížnější než v evropských jazycích, ale pořád se do dá zvládnout automaticky)
2. Eliminovat ty výskyty, kde *z a do* nejsou syntakticky závislé na stejném slově (to už bylo třeba udělat ručně)
3. Oanotovat každý výskyt podle různých vlastností věty a jejího kontextu.



Souhlaska\_pred

- Prvym  
 Druhym  
 Obema  
 Zadnym

Hlavni

- Min ila  
 Ila min

Zanr

- Proza  
 Poesie

delka1segmentu\_slabik

delka1segmentu\_slov

Delka2segmentu\_slabik

Delka2segmentu\_slov

Fraze

- Neni  
 Min x ila x  
 Min x ila uchra

Znamy\_z\_kontextu

- Prvy  
 Druhy  
 Oba  
 Zadny

Urceno

- Prvni  
 Druhy  
 Oba  
 Zadny

ClenUrcity

- Prvni  
 Druhy  
 Oboji  
 Zadny

Zajmeno

- Prvni  
 Druhy  
 Oboji  
 Zadny

Vyznam

- Mistne  
 Casove  
 Osoba  
 Abstraktne  
 Ostatni

Obecnost

- 1 Obecnější  
 2 Obecnější  
 Stejne  
 Nerozlišitelne

Negace\_slovesa

- Pozitivni  
 Negativni

Druh

- Sloveso  
 Jmeno

PoziceVuciSlovesu

- Za  
 Pred

Rod

- Muzsky  
 Zensky

Zpusob

- Indikativ  
 Subjunktiv  
 Apokopat

Cas

- Perfektum  
 Impefektum  
 Imperativ

Osoba

1. os  
 2. os  
 3. os

Podmet

- Bez podmetem  
 S podmetem

Semantika

- Prijit  
 Odejit  
 Zustat  
 Jine

Cislo

- Sg.  
 Pl.  
 Dual

pad

- Nom  
 Gen  
 Dat

والنظر \* إلى الحديقة \* من أعلى ..

فة التي تفصل بين حافة النافذة والأرض ليست كافية . أجابت كاتبة بمزاج حاد . عندي \* من الغيظ ما يكفي لأموت حتى لو قفزت \* من فوق علبة كبريت . هذه الطريقة أسميها انتحاراً ، وهي عمل غير شجاع وفيها خروج على الاتفاق . إن المزاج الحاد ، يعتبر \* من فصيحة الأدوات الحادة . واستعماله غير مشروع في pq اتفاقيتنا ، يمكنك الصعود \* إلى بيت أحد الجيران في الطابق الثالث وشرب فنجان \* من pq القهوة على شرفتهم ، والنظر \* إلى الحديقة \* من أعلى ...

# Které vlastnosti nejvíc ovlivňují distribuci opozice $z - do / do - z$ ?

- Informační dok
- Fonologická struktura
- Gramatické kategorie závislých větných členů
- Gramatické kategorie řídicího členu
- Význam závislých členů
- Význam řídicího členu
- Další

$Z - do$  je signifikantně častější než  $do - z$ .

$Z - do$	175
$Do - z$	26

Podle binomické distribuce můžeme tvrdit:

$$P < 0.0001$$

# Informační struktura (1)

- Je konstrukce z-do / do-z v tématu, nebo v rematické části?

ذهبت من بيتي إلى بيته.

Šel jsem ze svého domu do jeho domu.

**VS.**

من بيتي ذهبت إلى بيته.

Ze svého domu jsem šel do jeho domu.

(V arabštině „agramatické“)

# Informační struktura (1)

V rámci našich dat (201 výskyt) se konstrukce *z-do / do-z* nevyskytovala před řídicím členem **ani jednou**.

...zdá se že je *do* v arabštině skutečně agramatické.

# Informační struktura (2)

Je závislý člen znám z předchozího kontextu?

الآن أنا في بيتي ، سأذهب من بيتي إلى بيته.

Jsem teď ve svém domě. Půjdu ze svého domu do jeho domu.

**vs.**

الآن أنا في بيته ، ذهبت إلى بيته من بيتي.

Jsem teď v jeho domě. Šel jsem do jeho domu ze svého domu.

# Informační struktura (2)

Z (známý) do (známý)	26
Z (známý) do (neznámý)	21
Z (neznámý) do (známý)	0
Z (neznámý) do (neznámý)	127
Do (známý) z (známý)	8
Do (známý) z (neznámý)	11
Do (neznámý) z (známý)	2
Do (neznámý) z (neznámý)	5

# Informační struktura (2)

	<b>Z+známý do + neznámý</b>	<b>Z+neznámý do+známý</b>	<b>Celkem</b>
<b>Z - do</b>	21	0	21
<b>Do - z</b>	2	11	13
<b>Celkem</b>	23	11	34

Fisherův test:

$P < 0.0001$



- Existuje signifikantní tendence k tomu, že jestliže je argument se „z“ známý z předchozího kontextu, a argument s „do“ je neznámý, pak argument se „z“ bude uveden jako první a vice versa.

# Fonologická struktura (1)

Předchází předložku souhláska?

ذَهَبْتُ مِنْ بَيْتِي إِلَى بَيْتِهِ.

Šel jsem ze svého domu do jeho domu.

**VS.**

ذَهَبَتْ مِنْ بَيْتِي إِلَى بَيْتِهِ.

Šla z mého domu do jeho domu.

# Fonologická struktura (1)

	<b>Cons. + z</b>	<b>Voc + z</b>	
	<b>Voc. + do</b>	<b>Cons. + do</b>	<b>Celkem</b>
<b>Z - do</b>	13	36	49
<b>Do - z</b>	3	2	5
<b>Celkem</b>	16	38	54

Fisherův test:

$P = 0.15$

# Fonologická struktura (2)

Následuje po předložce určitý člen?

ذهبت من السوق إلى بيته.

Šel jsem z toho trhu do jeho domu.

**VS.**

ذهبت من بيتي إلى السوق.

I šel jsem z mého domu na ten trh.

# Fonologická struktura(2)

	<b>Z + al-</b>	<b>Z + non (al-)</b>	
	<b>do + non (al-)</b>	<b>do + al-</b>	<b>Celkem</b>
<b>Z - do</b>	16	22	38
<b>Do - z</b>	2	6	8
<b>Celkem</b>	18	28	46

Fisherův test:

$P = 0.45$

# Gramatické kategorie závislých větných členů

Jsou závislé větné členy určeny?

ذهبت من السوق إلى بيت.

Šel jsem z toho trhu do nějakého domu.

**VS.**

ذهبت من بيت إلى السوق.

Šel jsem z nějakého domu na ten trh.

# Gramatické kategorie závislých větných členů

	<b>Z + urč. do + neurč.</b>	<b>Z + neurč. do + urč.</b>	<b>Celkem</b>	
<b>Z - do</b>		8	8	12
<b>Do - z</b>		4	2	10
<b>Celkem</b>		16	6	22

Fisherův test:

$$P = 0.65$$

# Gramatické kategorie řídicího členu

Je řídicí člen sloveso, nebo jméno?

ذهبت من السوق إلى بيتي.

Šel jsem z trhu do svého domu.

**VS.**

انا ذاهب من بيتي إلى السوق.

Jsem jdoucí ze svého domu na trh.



# Gramatické kategorie řídicího členu

	<b>Sloveso</b>	<b>Jméno</b>	<b>Celkem</b>
<b>Z - do</b>	110	65	175
<b>Do - z</b>	19	7	26
<b>Celkem</b>	129	72	201

Fisherův test:

$P = 0.38$

# Gramatické kategorie řídicího členu

- Zkoumáno mnoho kategorií – čas, způsob, mód, osoba, rod (gender)...
- Pád u jmen
- Presence/absence subjektu, negace.
- **Nic z toho se neukázalo jako významně ovlivňující distribuci zkoumaného jevu.**

# Význam závislých členů (1)

1. Z Francie do Egypta. من فرنسا إلى مصر.
2. Od neděle do pátku. من الأحد إلى الجمعة.
3. Od 9 do 10 kg. من 9 إلى 10 كيلو.
4. Od Muhammada k Alímu. من محمد إلى علي.

**Bez signifikantních výsledků, ale (2) a (3) vypadají slibně**

## Význam závislých členů (2)

- Který argument je obecnější?

• انتقل من دمشق إلى اوربا.

- Přestěhoval se z Damšku do Evropy.

**vs.**

انتقلوا إلى دمشق من كل أنحاء العالم.

Přestěhovali se do Damašku ze všech koutů světa.

# Význam závislých členů (2)

	<b>Z</b> <b>+ obecnější</b>	<b>Do</b> <b>+ obecnější</b>	<b>Celkem</b>
<b>Z - do</b>	9	26	35
<b>Do - z</b>	16	4	20
<b>Celkem</b>	25	30	22

Fisherův test:

$$P = 0.0002$$

## Význam závislých členů (3)

- Je argument složen z osobního zájme?

ذَهَبْتُ مِنْ بَيْتِي إِلَيْهِ.

Šel jsem ze svého domu k němu.

**VS.**

ذَهَبْتُ إِلَيْهِ مِنْ بَيْتِي.

Šla k němu z mého domu.

# Význam závislých členů (3)

	<b>Z + pron. do + non pron.</b>	<b>Z + non pron. Do + pron.</b>	<b>Celkem</b>
<b>Z - do</b>	9	0	9
<b>Do - z</b>	0	16	16
<b>Celkem</b>	9	16	25

Fisherův test:

$P < 0.0001$

# Význam závislých členů (4)

- Který argument je delší?
- Druhý argument má významně více slabik i slov než první.



	<b>Průměrná délka</b>			
	Ve slovech		Ve slabikách	
	1	2	1	2
Z - do	1.41	1.99	4.38	5.95
Do - z	0.42	2.08	1.73	6.58
Celkem	4.10	6.13	1.30	2.03
P (t-test)	<0.00001		<0.00001	

Použit Studentův test (není ideální)

# Význam řídícího členu

- z – do

انتقل

خط خروج سرب نقل

توجه مضى دخل هنز طري لقي

تأمل

كض خرج

تأمل

عربي قفز

ممدود انصرف هاجر وزع امتد

موصل

تنقل نقل

استظهر رش ممتد

انتقال تناقل

حمل أجب

طار اتجه هرب قذف عابر هرب

أرسل فرار

طريق

سبابط تخاطف آتي إياب طلع

متجه

سبل قادم

مد أخذ سقط

أنتقل

جاب تأرجح مندلع انتقال

ولد عمق

نظر

- do – z

انتقل

جلب

نزل لجوء

متجه

قادم وصل

ممتدتين

دخل

أشار

تراامي

مر

تناهي

اتي

عاد جاء

نظر

Je třeba více dat

# Další...

- Některé fráze mají pevný slovosled

من المهد إلى اللحد

Od kolébky do hrobu

من الخليج إلى المحيط

Od Zálivu k Oceánu

(500k výskytů na Google, opačná varianta má pouze jeden výskyt)

Samozřejmě nemůžeme prozkoumat všechny vlastnosti (pokud vás napadne nějaká důležitá, na kterou jsem zapomněl, nenechávejte si ji pro sebe).

# Závěr

- *Z – do* je frekventovanější než *do – z*
- Druhý argument tíhne k tomu být:
  - kratší
  - méně obecný
  - známý z předchozího textu
  - složený z předložky a zájmena
- Pevný slovosled je také v některých frázích

# Pro další výzkum...

- Více dat
- Více různých předložek
- Více jazyků (pro češtinu se zdá, že budou závěry podobné, včetně vzájemného poměru obou konstrukcí)
- Pořádná PCA (component analysis)

Díky!