



A Corpus-Driven Analysis of -r Dropping in Spoken Turkish

Yuji KAWAGUCHI

Tokyo University of Foreign Studies

**Corpus and Variation in Linguistic Description and Language Education
Global-COE Symposium, May 8-9, 2008**

Presentation flow

1. Corpus design for Spoken Turkish
21st Century COE “Usage-Based Linguistic Informatics”
and Multilingual Spoken Corpora
 2. Problems of –r dropping :
Present progressive suffix *-lyor*
Some factors of –r dropping
 3. –r dropping : the word *bir* “one, certain”
Correlation between *-yor* and *bir*
- Conclusion

1. Corpus design for Spoken Turkish

- The 21st Century COE (2002-2006)
Usage-Based Linguistic Informatics (UBLI)
- Multilingual Spoken Corpora (MSC)

French (Univ. Aix-Marseille, Univ. Paris 13)

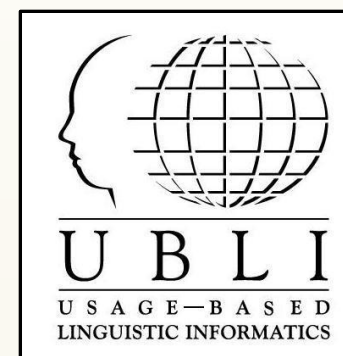
Spanish (Autonomous Univ. of Madrid)

Malay (Universiti Kebangsaan Malaysia)

Chinese (Tamkang University, Taiwan)

Russian (Moscow State University of Humanities)

Japanese



Multilingual Spoken Corpora: Turkish



Collaborated Universities

- Marmara University
- Istanbul University

http://www.coelang.tufs.ac.jp/multilingual_corpus/tr/index.php?contents_xml=top&lang=en

Recording Profiles

	Corpus	Length	Topic	Date
1	D101-03-05	28.77	Education and university	2005.6.23
2	D102	32.36	Cigarette	2005.6.23
3	D104	29.26	EU and Turkey	2005.6.23
4	D106	29.52	Television	2005.6.23
5	D107	35.42	On Turkey	2005.6.24
6	D108	22.46	Leisure	2005.6.24
7	D109	26.25	Hometown	2005.6.24
8	D110	29.12	Tourist spot	2005.6.24
9	D111	26.5	Vacation	2005.6.24
10	D112	28.18	Language Education and Turkish	2005.6.24
11	D201	5.40	Education as an occupation	2006.6.8
12	D202	38.32	Turkish culture	2006.6.21
13	D203	26.53	Test	2006.6.14
14	D204	26.53	Lecture	2006.6.14
15	D205	25.14	USA and Turkey	2006.6.21
16	D206	36.13	Students	2006.6.21
17	D207	18.38	Turkish literature and culture	2006.6.21
18	D208	30.54	Foreign language and Turkish	2006.6.21
19	D209	31.32	Poems	2006.6.29
20	D210	5.19	Leisure	2006.6.29
21	D211	15.28	Drams	2006.6.29
22	D212	26.53	Music	2006.6.29
23	D213	22.20	Vacation	2006.6.20
24	D214	5.51	Traveling abroad	2006.6.8
25	D215	24.13	French language	2006.6.8
26	D216	23.55	Istanbul	2006.6.8
27	D217	5.29	Spoken language	2006.6.8

▪ 27 free conversations recorded in June 2005 and June 2006

▪ total length:
10 h 26 m 21 s

▪ total word types:
17.417

▪ total word tokens:
93.837

Informants

		Sex	Age	Birth place	Profession
1	AHT	M	30	İstanbul	Teacher
2	AK	F	32	Elazığ	Graduate Student
3	BG	F	26	Izmir	Researcher
4	BH	F	24	İstanbul	Graduate Student
5	DH	F	26	Gebze	Teacher
6	EY	F	?	?	Student
7	FK	F	55	İstanbul	Teacher
8	HS	M	29	Kırşehir	Teacher
9	MG	M	28	Kayseri	Teacher
10	MÜ	F	40	Isparta	Student
11	NS	F	?	?	Student
12	ÖK	M	29	Bursa	Student
13	RŞŞ	M	38	Nevşehir	Teacher
14	SG	M	30	İstanbul	Teacher
15	ŞK	M	27	Sivas	Teacher
16	SY	M	36	Yozgat	Teacher
17	ÜD	M	30	Ardahan	Teacher

- 17 informants
- 9 men and 8 women
- age from 24 to 55
- born at different places
- highly educated Turks

Transcription

Symbol	Meaning
(?), (!)	Question or exclamation
-	Beginning of discourse
#, ##, ###	Short or long pause
Ooo, aaa	Long vowel or emotion: Ex. Yook ‘Oh, no’, haayır ‘Oh, yes’
(.)	Hesitation or suspension
(.....)	Dropped segment(s)
_____	Accentuated part or word
e (eee)	Filler
m (mmm)	Confirmation
{.....}	Dislocation : Ex. Hava güzel {bugün} = Bugün hava güzel. ‘It is fine today.’
[.....]	Paralinguistic elements like smile, laugh, cough, etc.
<.....>	Overlapping
.....	Relatively rapid part
%.....%	Special intonation or speech melody

Transcription rules for Spoken Turkish based on the system of Morel et al. (1998), arranged by Yuji Kawaguchi and Selim Yılmaz, 2005, cf. Yılmaz (2006)

Example

dislocation

overlapping

paralinguistic action

D206:

ŞK2- ben de iyiyim nasıl geçti {bugün sınav} (?)

BG2- # idare eder [gülümseme]

ŞK3- [gülme] umduğunu buldun mu (?)

BG3- yok hayır bulamadım

ŞK4- yok peki # nasıl ee gidiyor hayat (?) iyi mi (?) [gülme]

BG4- valla sıcaklar da biraz <zor>

BG4- okula gelip gitmek sınavlara girip çıkmak çok zorlanıyor {insan}

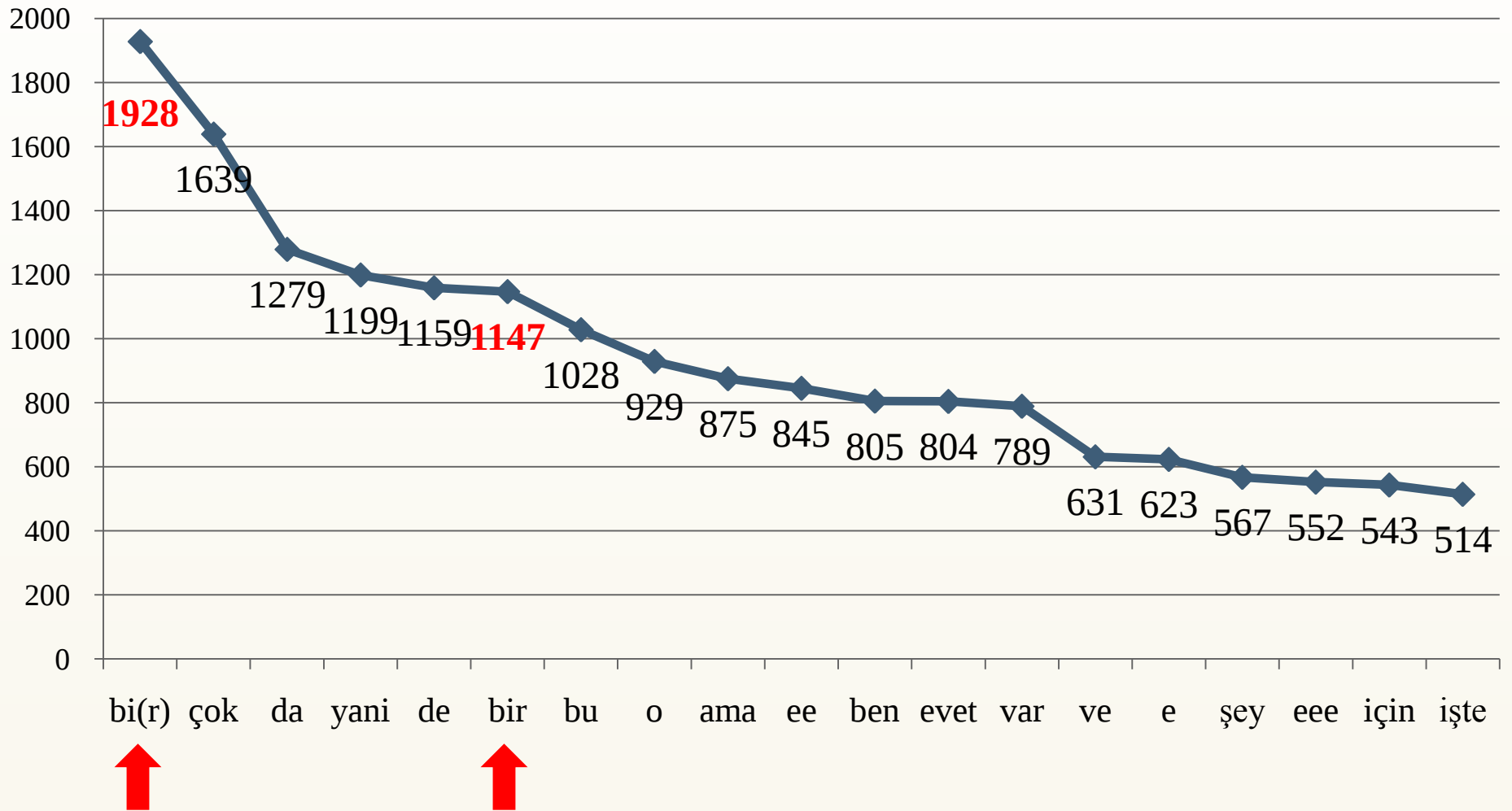
ŞK5- evet gerçekten # havalar çok sıcak <böyle insan evet >

BG5- insanda evet bir bitkinlik oluyor

ŞK6 – mm

BG6- adapte1 zorluğu oluyor

Word tokens



2. Problems of –r dropping: Present progressive suffix *-Iyor*

8.2.2. Consonant dropping

(i) /r/ dropping

The /r/ sound in the final position of the word *bir* ‘one, certain’ and in the morpheme {Iyor} is generally dropped in a rapid conversation.

bir → *bi*

geliyor → *geliyo*

/r/ dropping in the morpheme {Iyor} is also attested when followed by the second singular or plural ending:

geliyorsun → *geliyosun*

geliyorsunuz → *geliyosunuz*

(Özsoy, 2004, p.109)

Example of –r dropping

D102:

AHT 8 – ben keyif alarak öliym (= öleyim) **diyo(r)sun**?

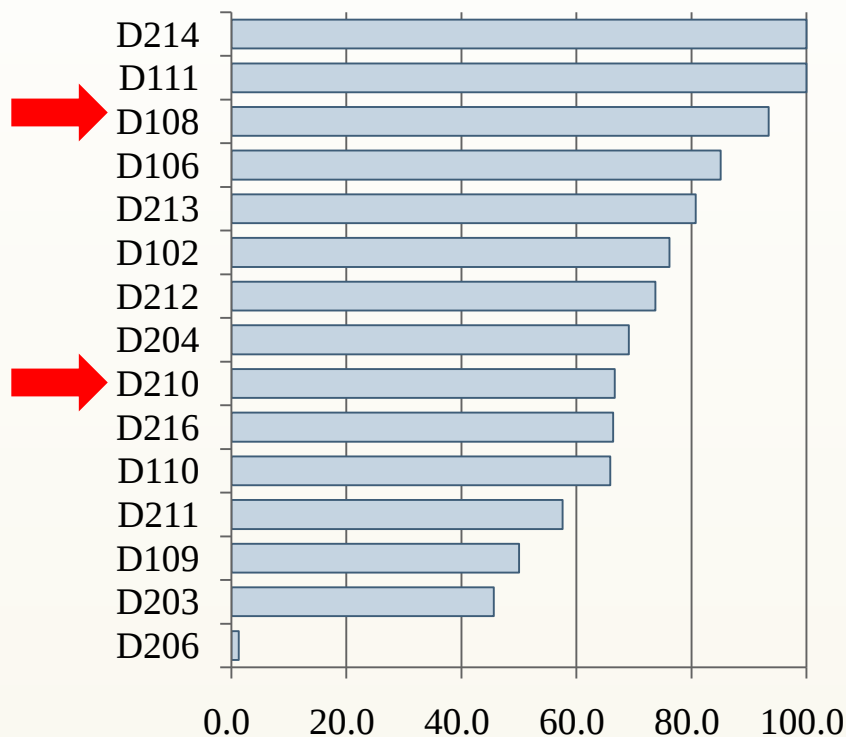


AK 8 - ben keyif alarak öliym (= öleyim) benim **bi(r)** tane şey İranlı arkadaşım vardı eee bu İran - Irak savaşında işte bombalama filan **oluyo(r)muş** bu geceleri artık hani karartmalar filan **oluyo(r)** ya? bunlar sığınağa **iniyo(r)larmış** sığınakta **bi(r)** tane arkadaşı oturmuş bayan makyaj yapmaya başlamış İşte rimel filan **sürüyo(r)muş** diğer herkes şaşkınlıkla buna bakmış İşte İranlılarda çok süslü böyle **bi(r)** şey vardır ya? # herkes demiş ki “**n’apıyosun**? (= ne yapıyorsun) bak işte bombardıman var ölücez kalacaz” sıkış sıkış oturmuşlar O da şey söylemiş “yani ölürsem de güzel öleyim hiç fark etmez benim için” bu da ölürsem de sigarayla öliym (= öleyim)

AHT 9 – üç gün fazla yaşamışım beş gün az yaşamışım hiç önemli değil

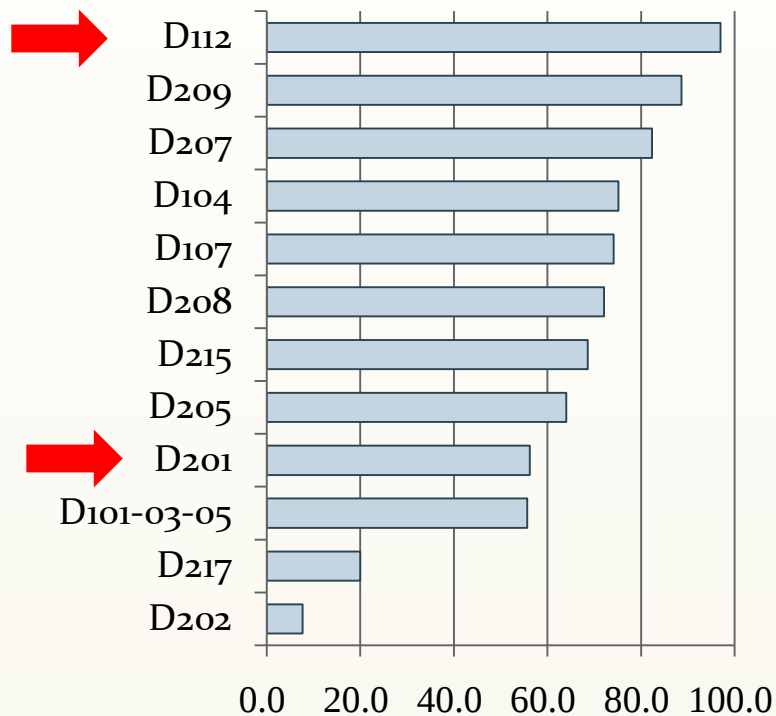
AK 9 - hiç önemli değil **bi(r)** şey ifade **etmiyo(r)**

Topic and -r dropping of *-Iyor*



general topic

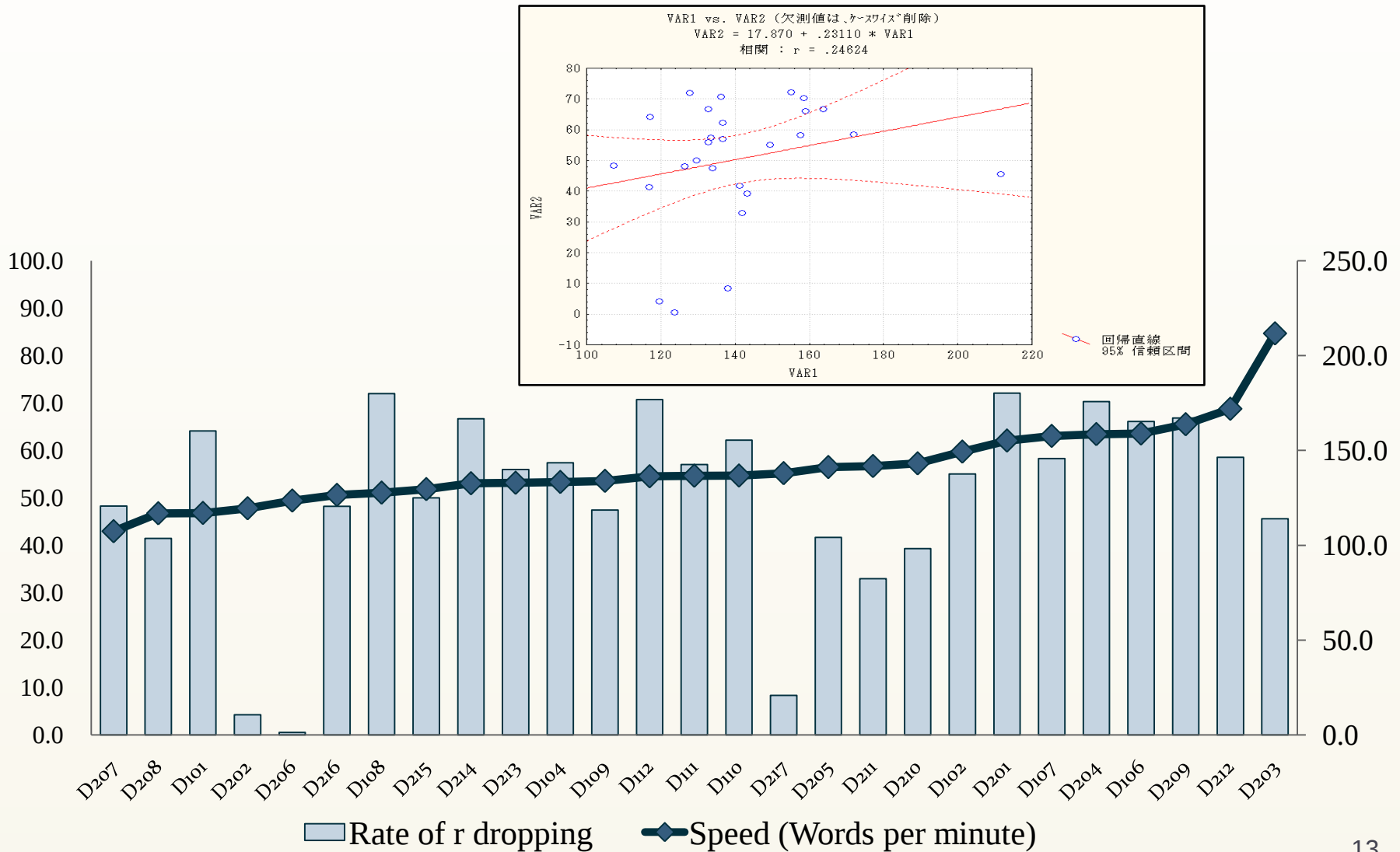
students, test, hometown, dramas, tourist spot, Istanbul, leisure, music, cigarette, vacation



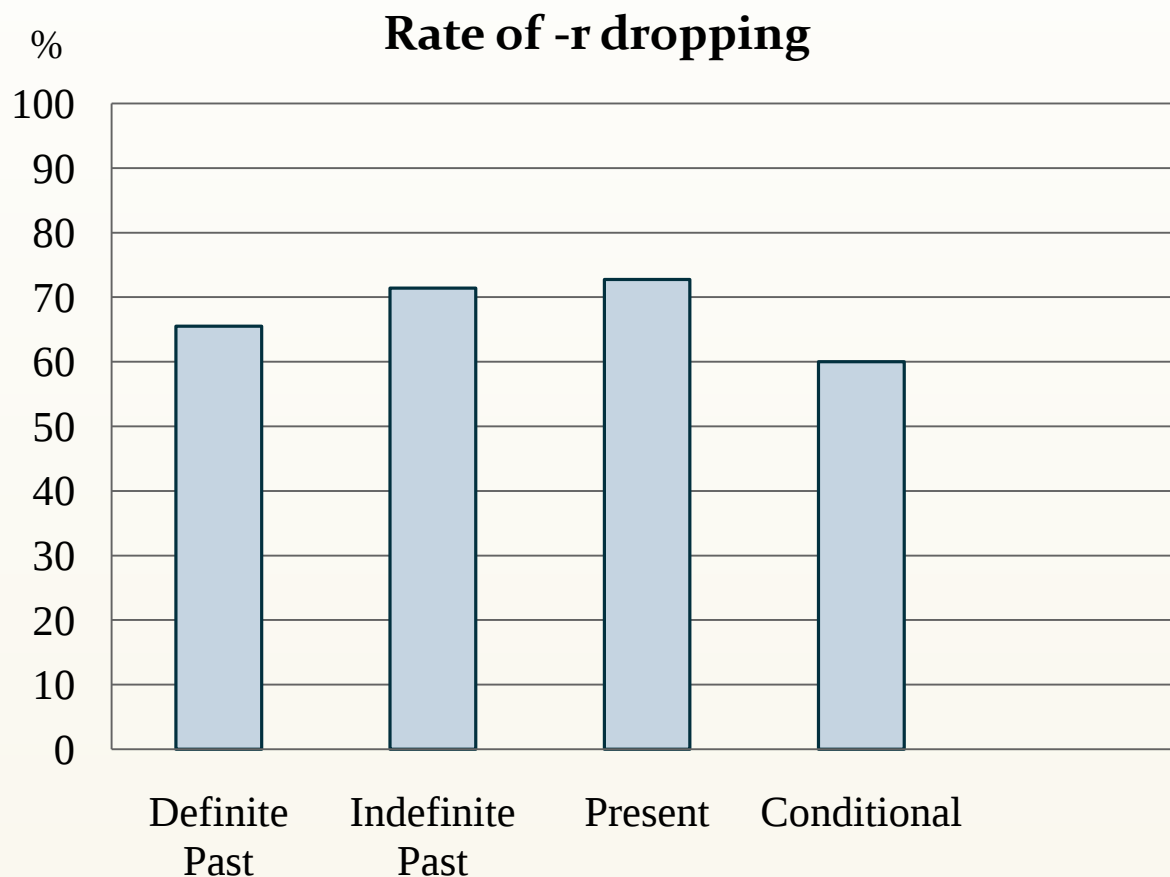
specific topic

Turkish culture, spoken language, education, EU and Turkey, Turkish literature

Speed and -r dropping of *-Iyor*



Grammatical context



		-r dropping
Definite Past	yordu	53
	yorduk	15
	yordum	43
	yordun	8
	yordunuz	3
Indefinite Past	yorlardı	9
	yorlarmış	10
	yormuş	32
	yormuşun	1
	yormuşum	2
Present	yor	1308
	yorsun	279
	yorsunuz	91
	yorlar	225
	yormu	2
Conditional	yorki	1
	yorsa	23
	yorsam	5
	yorsan	18
	yorsanız	2

Contexts of -r dropping

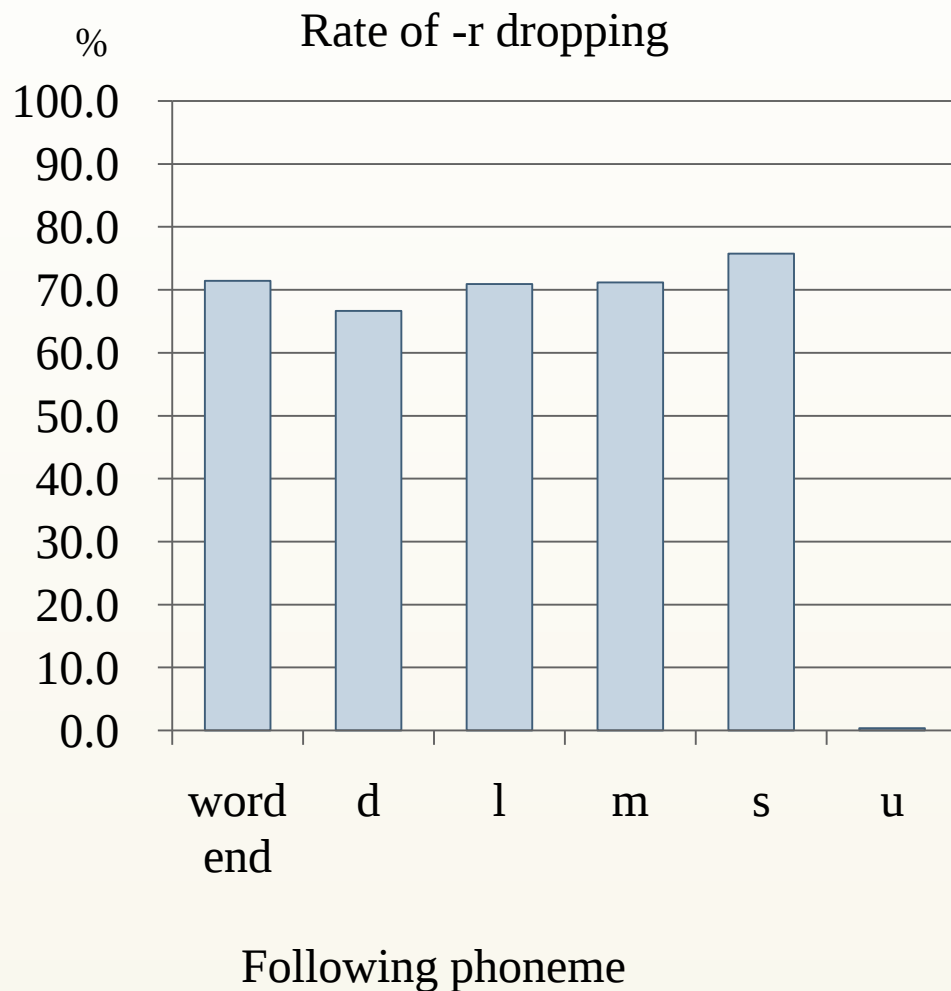
In colloquial speech the 'r' at the end of the suffix is often not pronounced **when it is followed by a suffix beginning with a consonant**: *gidiyorsun* 's/he goes/ is going' [ɟidíj**o**sun] / [ɟidíj**o**rsun], (...) or **when it is at the end of a word**: *sarıyor* 's/he wraps' [saríj**o**] / [saríj**o**ɾ]. 'r' is retained **when it is followed by a suffix beginning with a vowel**: *seviyorum* 'I love/like' [sevíjorum].

(Göksel and Kerslake, 2005, p.83.)



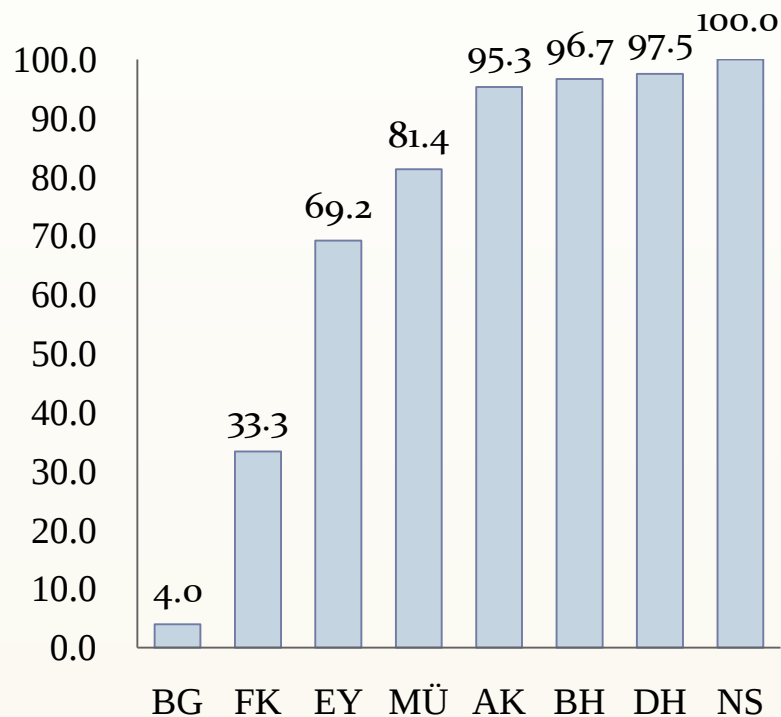
- 1) followed by a consonant → -r dropping
- 2) at word end → -r dropping
- 3) followed by a vowel → no -r dropping

Phonetic context

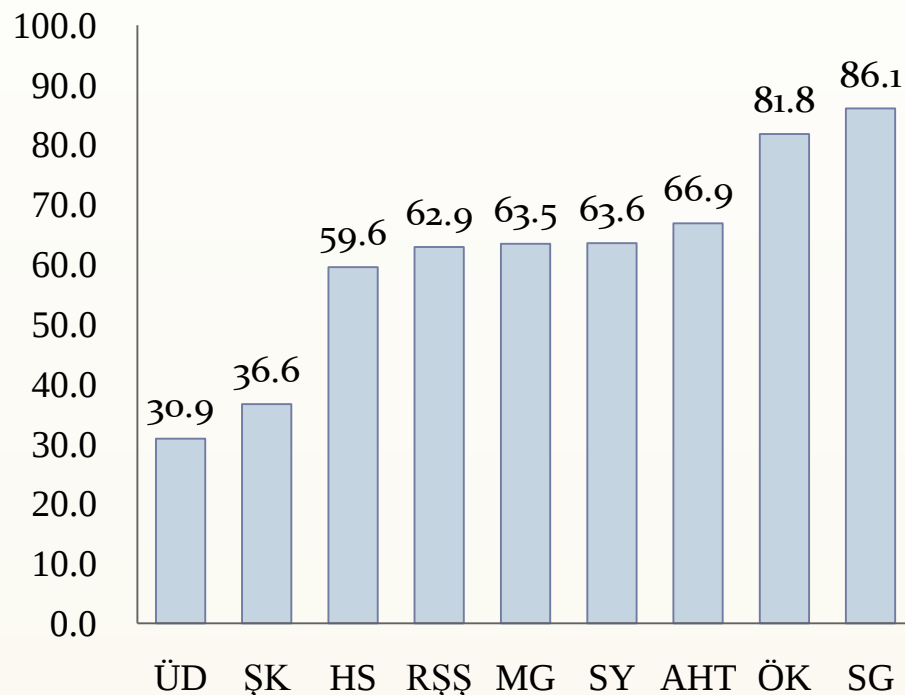


	-r dropping	no dropping
yordu	53	23
yorduk	15	14
yordum	43	15
yordun	8	8
yordunuz	3	1
yordur	6	3
yorki	1	1
yorlar	225	88
yorlardı	9	8
yorlarmış	10	4
yormu	2	1
yormuş	32	11
yormuşun	1	1
yormuşum	2	2
yorsa	23	24
yorsam	5	5
yorsan	18	1
yorsanız	2	2
yorsun	279	79
yorsunuz	91	23
yorum	4	921
yoruz	0	235
yor	1308	523

Sex and -r dropping of *-Iyor*



women ↑ ↑ ↑ ↑

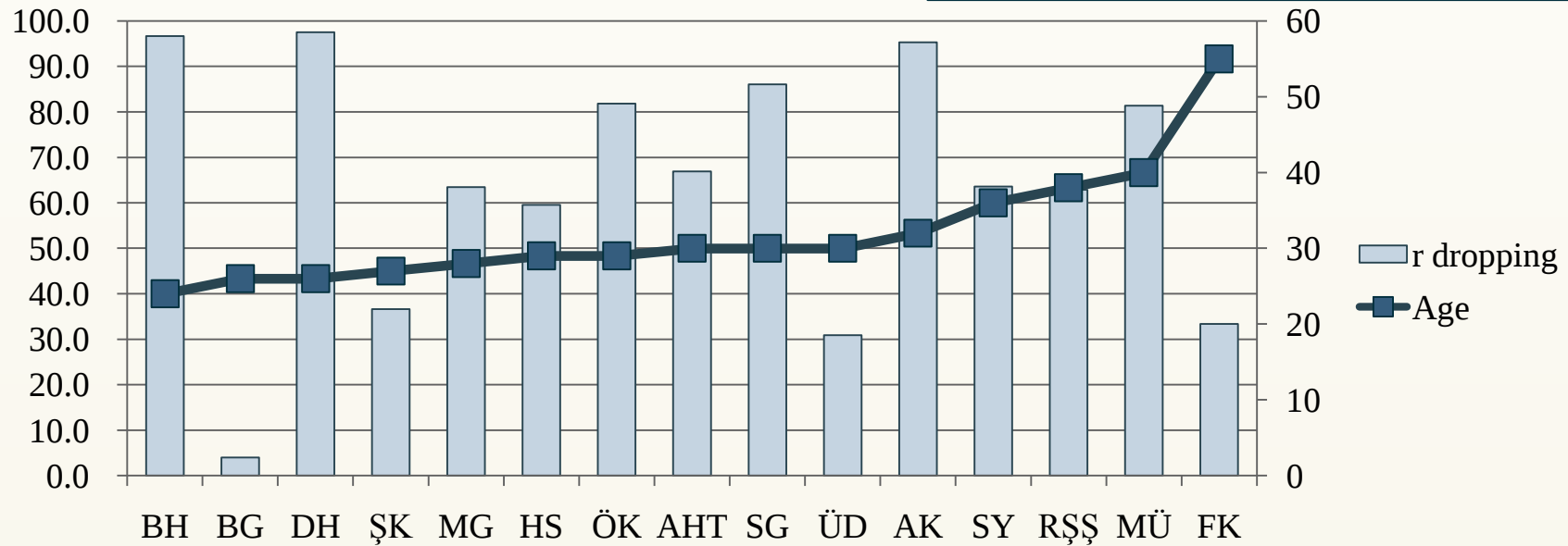
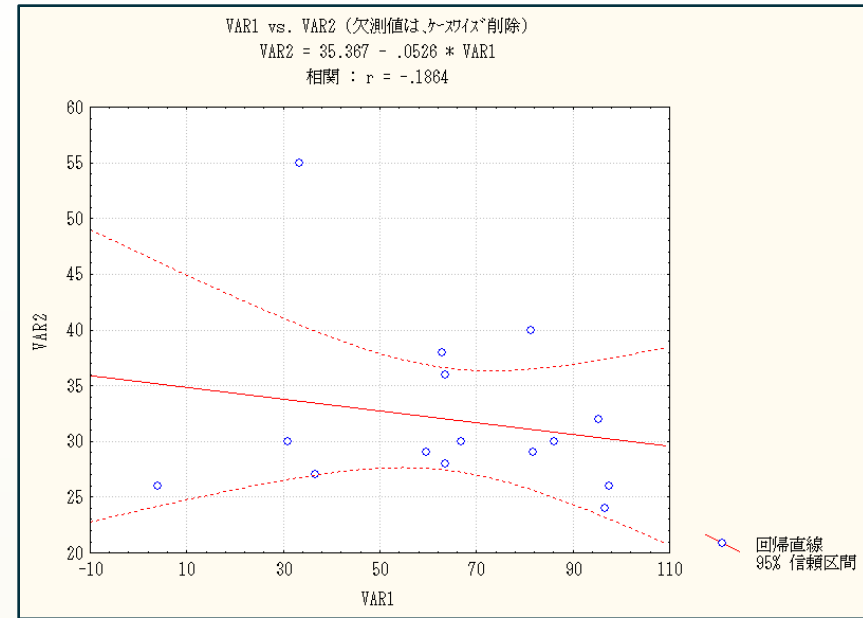


men

Statistically insignificant

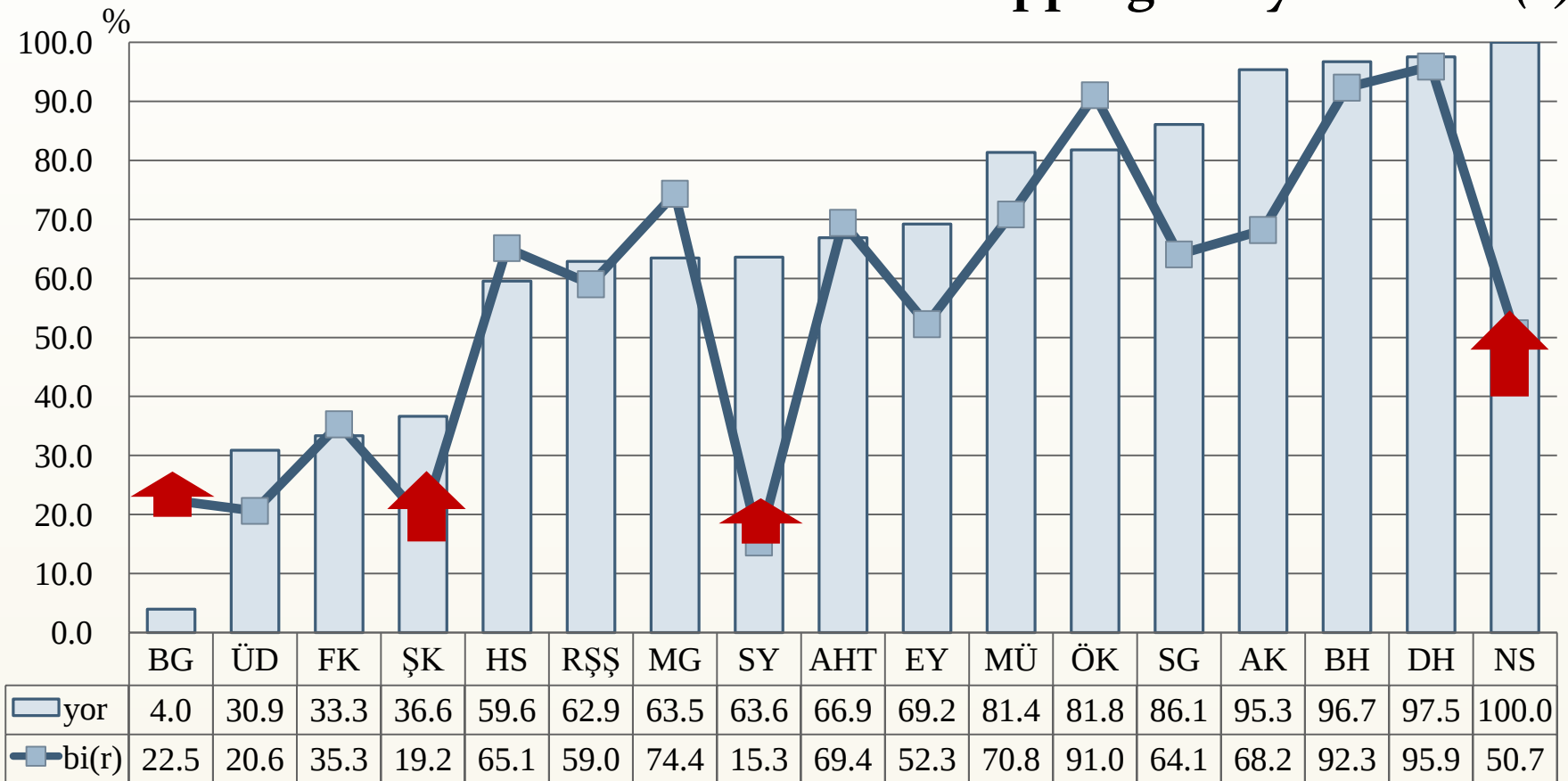
$n = 17, r = 0.16, |r| = 0.16 < 0.482 = r_{0.05}$

Age and -r dropping of *-Iyor*



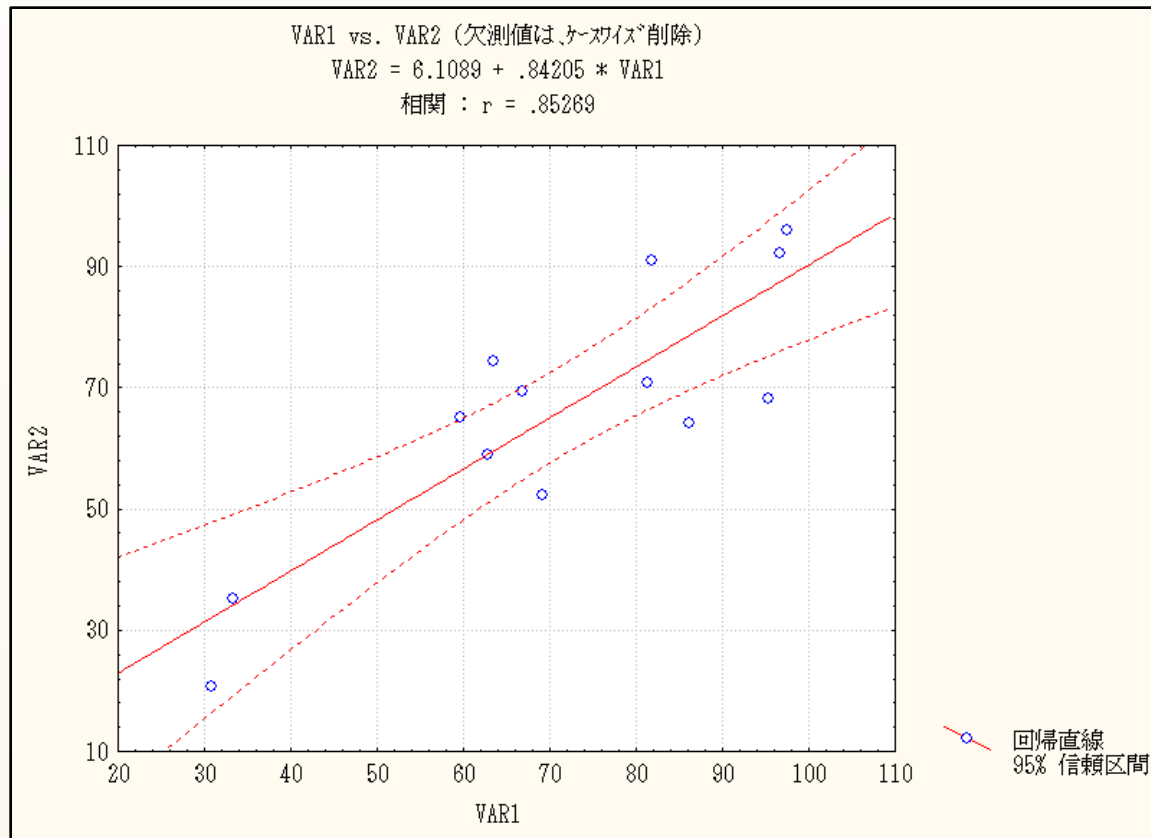
3. Individually motivated variation ?

Individual variation of -r dropping in *-lyor* and *bi(r)*



average of -r dropping: *-lyor* = 69.0 %; *bi(r)* = 61.8 %

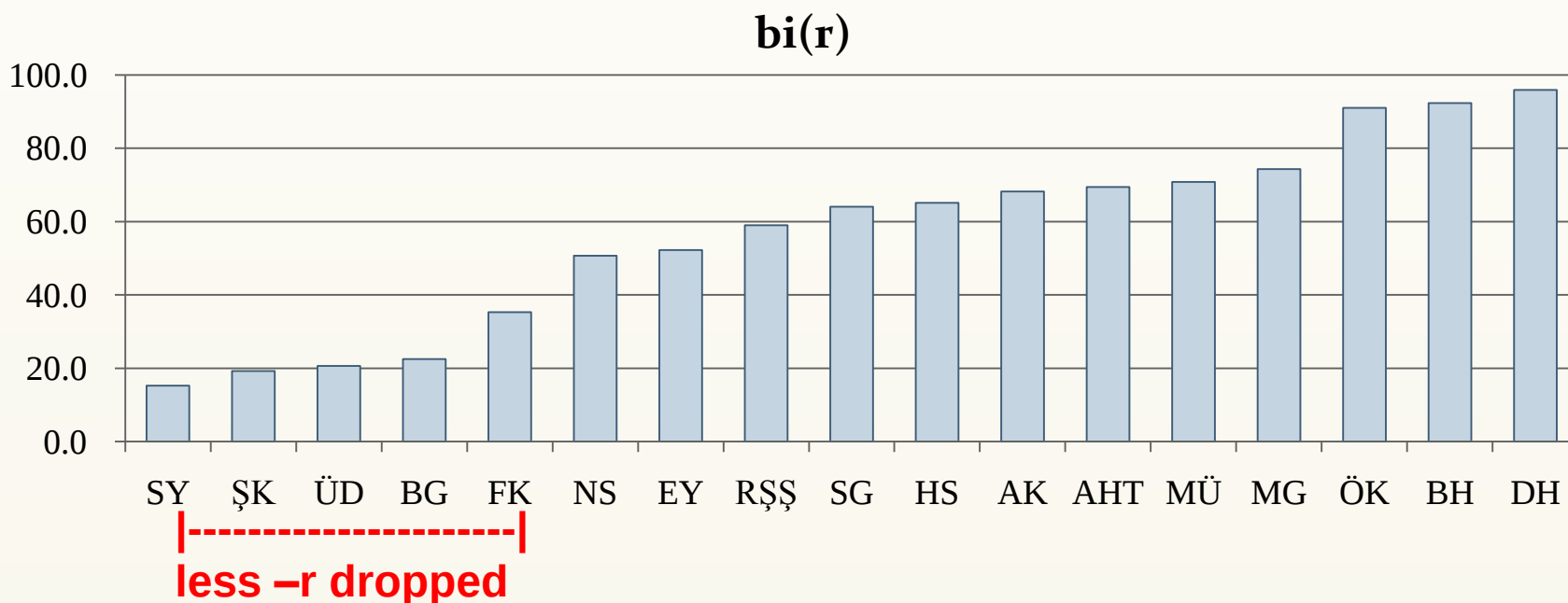
-r dropping of *-Iyor* and *bir*



	-r dropping	bi(r)
AHT	66.9	69.4
AK	95.3	68.2
BH	96.7	92.3
DH	97.5	95.9
EY	69.2	52.3
FK	33.3	35.3
HS	59.6	65.1
MG	63.5	74.4
MÜ	81.4	70.8
ÖK	81.8	91.0
RŞŞ	62.9	59.0
SG	86.1	64.1
ÜD	30.9	20.6

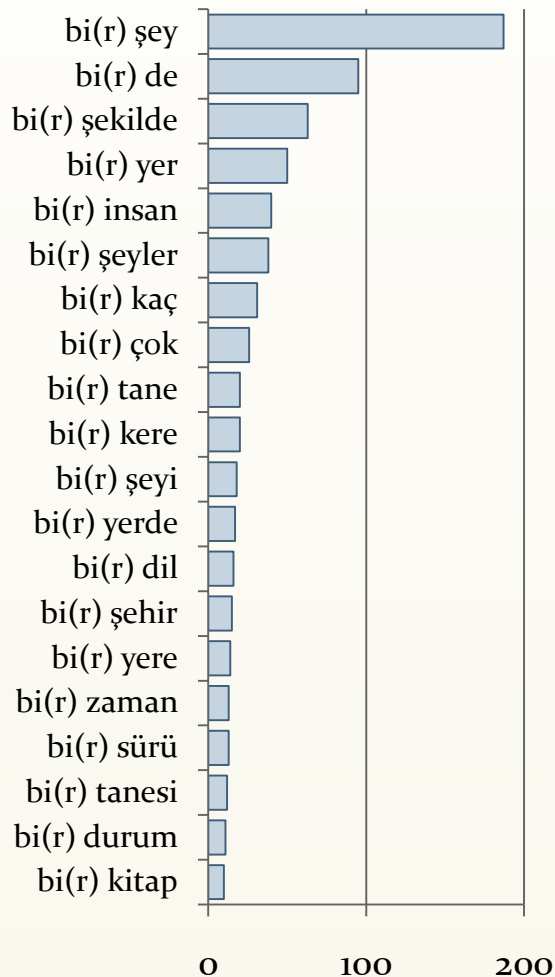
-r dropping in *bir*

- Topic, sex and age factors are all irrelevant.
- Speed and -r dropping of *bir* is correlated:
 $n = 27, r = 0.46 \quad |r| = 0.46 < 0.381 = r_{0.05}$
- Individual preference

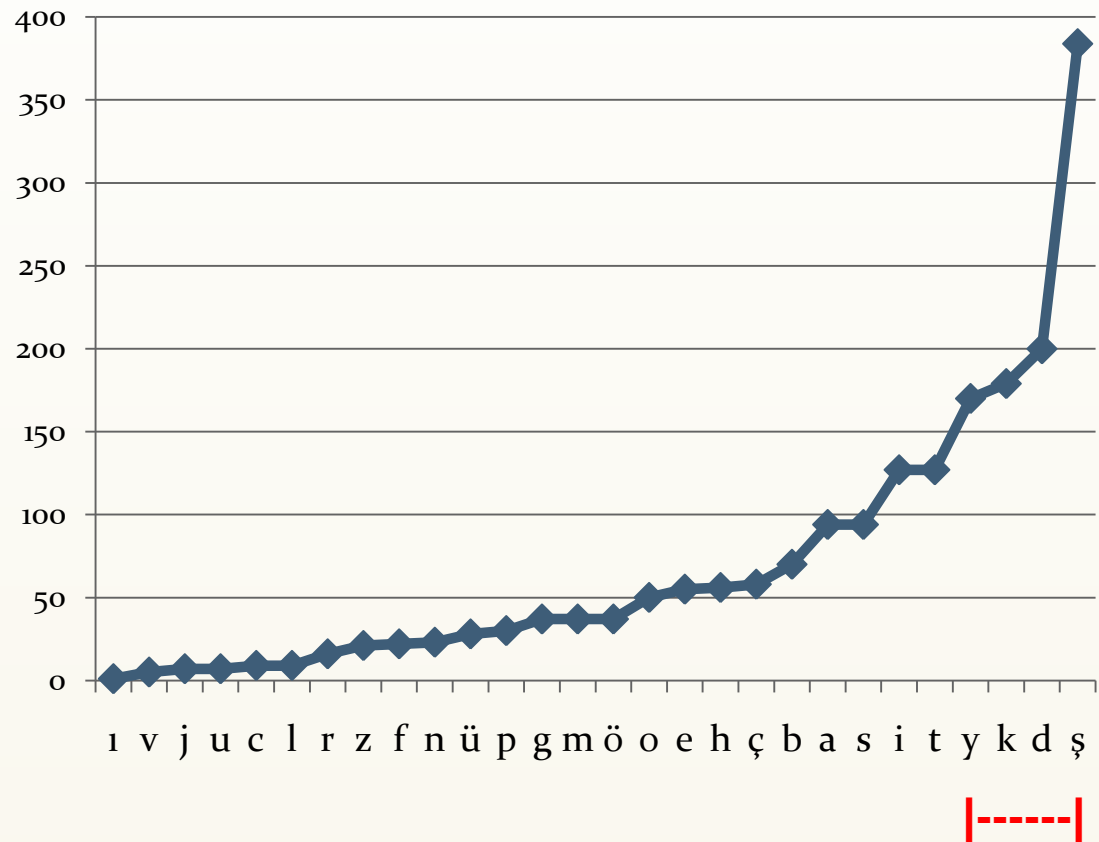


-r dropping in *bir*

Collocations of *bir*



occurrences of bi(r)



Conclusion

- Context of -r dropping
 - Iyor*
 1. Word end
 2. Before the consonants /d, l, m, s/
 - bir*
 1. Collocations: *bi(r) şey, bi(r) de, bi(r) şekil, bi(r) yer*
- Individually motivated variation
 - A tendency to avoid -r dropping for some informants
- Strong correlation in -*Iyor* and *bir*
 - Topic and age factors are irrelevant to -r dropping
- Problem of corpus size

References

- Aksan, Doğan (ed.) (1978) *Türkiye Türkçesi Gelişmeli Sesbilimi*, Ankara: Türk Dil Kurumu Yayınları.
- Delais-Roussaire, Élisabeth et Jacques Durand (éd.) (2003) *Corpus et variation en phonologie du français*, Toulouse: Presses Universitaires du Mirail.
- Demircan, Ömer (2001) *Türkçenin ses dizimi*, İstanbul: Der Yayınları.
- Gibbon, Dafydd, Roger Moore and Richard Winski (eds.) (1998) *Handbook of Standards and Resources for Spoken Language Systems*, Vol.1-4, Berlin/New York: Mouton de Gruyter.
- Göksel Aslı and Celia Kerslake (2005) *Turkish A Comprehensive Grammar*, London: Routledge.
- Kawaguchi, Yuji (2005) “Phonological system of Modern Turkish – A critical review on phonemes –”, *Journal of the Institute of Language Research*, No.10, The Institute of Language Research, Tokyo University of Foreign Studies, pp.35-62.
- (2006) “Usage-Based Approache to Linguistic Variation – Evidence from French and Turkish”, in Y. Kawaguchi et al. (eds.) *Spoken Language Corpus and Linguistic Informatics*, Amsterdam: John Benjamins, pp.247-267.
- Kornfilt, Jakline (1997) *Turkish*, London: Routledge.
- Morel, Marie-Annick et Laurent Danon-Boileau (1998) *Grammaire de l'intonation. L'exemple du français*, Paris: Ophrys.
- Özsoy, A. Sumru (1999) *Türkçe Turkish*, İstanbul: Boğaziçi Üniversitesi Yayınları.
- (2004) *Türkçenin yapısı – I*, İstanbul: Boğaziçi Üniversitesi Yayınları.
- Sarıca, Mustafa (2005) *Yeni dilbilim kuramlar ışığında Sözlü Dil Yapısı*, İstanbul: Multilingual.
- Selen, Nevin (1979) *Söyleyiş sesbilimi akustik sesbilim ve Türkiye Türkçesi*, Ankara: Türk Dil Kurumu Yayınları.
- Yılmaz, Selim (2006) “Présentation d'un corpus oral en turc contemporain”, in Yoichiro Tsuruga et al. (eds.) *Aspects of Corpus Linguistics – Spoken Language Corpora and Written Language Corpora-*, Working Papers of Linguistic Informatics, 12, 21st Century COE “Usage-Based Linguistic Informatics”, Graduate School of Tokyo University of Foreign Studies, pp.201-240.
- AntConc 3.2.1w (Windows), Laurence Anthony, Faculty of Science and Engineering Waseda University, Japan
anthony@waseda.jp.

For your attention, thank you !