

Case interference and phrase length effects in processing Turkish center-embeddings



Özge Bakay & Nazik Dinçtopal Deniz
Boğaziçi University, Turkey
CUNY Conference, University of Pennsylvania
March 4-6 2021



- This study investigates doubly center-embedded structures with relative clauses (2-CE-RCs) in Turkish.
 - *E.g., The rat that the cat that the dog chased ate died.* (Chomsky & Miller, 1963, p. 286)
- 2-CE-RCs are reported to be extremely difficult to process due to (among many other factors)
 - nesting of the same type of clauses (e.g., Yngve, 1960)
 - the number of incomplete syntactic dependencies (Chomsky & Miller, 1963)
 - working memory limitations (Gibson, 1998; 2000)
 - similarity-based interference (Lewis & Vasishth, 2005)
 - non-optimal prosodic phrase lengths (Fodor, 2013)
- **Similarity-based interference:** easier processing of 2-CE-RCs when the similarity among the subject NPs is decreased
 - e.g., in terms of their case marking
- **Prosodic phrase lengths:** easier processing of 2-CE-RCs with **optimal** and **balanced** phrase lengths
 - NP1 || RC1 || VP1 in English
 - NP1 || NP6 || VP1 in Turkish

The Present Study



- Turkish, a highly inflectional and head-final language, can provide further data on similarity-based case interference and prosodic phrase lengths.
- An eye-tracking experiment was conducted to investigate:
 - if syntactic interference, phonological interference and/or prosodic phrase lengths influence the real-time processing of Turkish center-embeddings,
 - when (if any) the effects of syntactic interference, phonological interference and prosodic phrase lengths would reveal,
 - if there is an interaction between syntactic interference, phonological interference and prosodic phrase lengths.

- Experimental sentences were Turkish 2-CE-RCs nested as a complement clause inside a matrix clause:

Ø	[Marangoz-lar-ın	[nakliyeciler-in	[kiracı-nın	gri	koltuğ-u	yerleştirdiğ-i	oda-ya]
<i>Pro</i>	carpenter-PL-GEN	mover-PL-GEN	tenant-GEN	gray	sofa-ACC	place-PART-3SG	room-DAT
	NP1	NP2	NP3			VP3	
taşıdıkları	dolab-ı]	kurdukları-nı]	biliyorum.				
move-PART-3PL	cabinet-ACC	set.up-VN-3PL-ACC	know-PROG-1SG				
VP2		VP1					

'I know that the carpenters set up the cabinet that the movers moved to the room that the tenant placed the gray sofa in.'

- Experimental sentences were Turkish 2-CE-RCs nested as a complement clause inside a matrix clause:

Ø	[Marangoz-lar-ın	[nakliyeciler-in	[kiracı-nın	gri	koltuğ-u	yerleştirdiğ-i	oda-ya]
Pro	carpenter-PL-GEN	mover-PL-GEN	tenant-GEN	gray	sofa-ACC	place-PART-3SG	room-DAT
	NP1	NP2	NP3			VP3	
taşıdıkları	dolab-ı]	kurdukları-nı]		bil-iyor-um.			
move-PART-3PL	cabinet-ACC	set.up-VN-3PL-ACC		know-PROG-1SG			
VP2		VP1					

'I know that the carpenters set up the cabinet that the movers moved to the room that the tenant placed the gray sofa in.'

Semantic support between the subjects and their verbs
to ease the processing of the already difficult structure
to prevent other plausible combinations of subjects and verbs

- Six conditions manipulating
 - Case interference: syntactic & phonological interference through case marking on the subject and object NPs
 - high syntactic – high phonological
 - low syntactic – high phonological
 - low syntactic – low phonological
 - Phrase lengths: (non-)optimal/(un-)balanced phrasing through phrase lengths
 - lengths encouraging optimal/balanced phrasing
 - lengths discouraging optimal/balanced phrasing
- 24 experimental sentences & 72 fillers, each followed by a comprehension question
- Task: read the sentences and answer the comprehension questions
- Participants: 43 native speakers of Turkish

Data analysis & Results



[Marangozlar] [nakliyecilerin] [kiracının] oldukça geniş gri koltuğu ... yerleştirdiği] odaya taşıdıkları] dolabı kurdular] sandım...
NP1 NP2 NP3 VP3 VP2 VP1 matrix V



critical region



spill-over region

Overall accuracy > 70% ($M = 84\%$)

Critical region:

Discouraging > **Encouraging** ($t's \geq 3$, $z = 2.51$: in first fixation dur., regression path dur., rereading dur., total dur. & pro. of regression out)

In both Discouraging and Encouraging Phrase Lengths Conditions:

HighSyn-HighPhono > **LowSyn-HighPhono** ($t's \geq 2.31$, in gaze dur., regression path dur., rereading dur., & total dur.)

HighSyn-HighPhono > **LowSyn-LowPhono**

($t's \geq 2.32$, $z = 2.63$: in gaze dur., regression path dur., rereading dur., total dur. & pro. of regression out)

LowSynLowPhono = **LowSynHighPhono**

Spill-over region:

Discouraging > **Encouraging** ($t = 2.52$ in regression path dur.)

In both Discouraging and Encouraging Phrase Lengths Conditions:

HighSyn-HighPhono = **LowSyn-HighPhono**

HighSyn-HighPhono = **LowSyn-LowPhono**

LowSynLowPhono = **LowSynHighPhono**

Mixed effects linear/logistic regression models

- Prosodic phrase lengths & case interference as fixed effects
- Subjects & items as random effects

Summary & Conclusions

Syntactic interference:

- Decreasing syntactic interference eased the processing of Turkish CEs
- This was the case regardless of phonological interference

Phonological interference:

- Did not affect the processing difficulty of Turkish CEs. **WHY?**
 - the complex structure of experimental stimuli, i.e., 2-CE-RCs nested inside another clause (e.g., Nakayama et al., 2005)
 - no grammatically-permissible conditions available that controlled for syntactic int. while manipulating phonological int.
 - phonological interference may not cause similarity-based retrieval interference effects at all (e.g., Obata et al., 2010)

Prosodic phrase lengths:

- Prosodic phrase lengths encouraging the optimal phrasing of the sentence eased the processing of Turkish CEs
- Prosodic phrase lengths effects persisted to the spill-over region. **WHY?**
 - prosodic information may be used during reanalysis (Marcus & Hindle, 1990)
 - the final interpretations may be determined by phrase lengths
 - an end-of-sentence acceptability judgment task to investigate this possibility

References

- Bakay, Ö. (2020). Processing Turkish center-embeddings: an investigation of case interference and prosodic phrase lengths. Unpublished MA thesis, Boğaziçi University, Istanbul.
- Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 269-321). New York; London; Sydney: Wiley.
- Fodor, J. D. (2013). Pronouncing and comprehending center-embedded sentences. In S. Montserrat, I. Laka, & M. K. Tanenhaus (Eds.), *Language down the garden path: the cognitive and biological basis for linguistic structures* (pp. 206-228). Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199677139.003
- Gibson, E. (1998). Syntactic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1-76. doi:10.1016/s0010-0277(98)00034-1
- Gibson, E. (2000). The dependency locality theory: a distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, and W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press. doi: 10.7551/mitpress/3654.001.0001
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3), 375-419. doi:10.1207/s15516709cog0000_25
- Marcus, M., & Hindle, D. (1990). Description theory and intonation boundaries. In G. Altmann (Ed.), *Cognitive models of speech processing: computational and psycholinguistic perspectives* (pp. 483-512). Cambridge, MA: MIT Press.
- Nakayama, M., Lee, S., & Lewis, R. L. (2005). Difficulty of processing Japanese and Korean center-embedding constructions. *Studies in Language Sciences*, 4, 99- 118.
- Obata, M., Lewis, R. L., Epstein, S., Bartek, B., & Boland, J. (2010). Featural analysis and short-term memory retrieval in on-line parsing: Evidence for syntactic, but not phonological, similarity-based interference. *Proceedings of NELS 41: Conference of the North East Linguistics Society*. Philadelphia.
- Yngve, V. H. (1960). A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society*, 104(5), 444–466.

THANK YOU!

Norming study for phrase lengths in Turkish

- A norming study was conducted to examine Turkish speakers' preferences for the prosodic phrasing of Turkish center-embeddings.
- 6 linguistically-informed judges were asked to insert two-three commas in the sentences.
- For the prosodic phrasing of the sentences as NP1 || NP6 || VP1;
 - all the judges placed a comma after NP1, in all the sentences (100%),
 - the number of commas after NP6 was 48 (33%),
 - two judges also placed a comma after NP2, in a total of 44 sentences (30.5%).
- There was no reliable difference between the sentences encouraging or discouraging the optimal phrasing.

Materials (Encouraging Phrase Lengths)

a. ENC, HighSyn-HighPhono

∅ [İşinin ehli marangoz-lar-in] || [nakliyeciler-in] [kiracı-nın gri koltuğ-u yerleştirdiği] ...
Pro expert carpenter-PL-GEN mover-PL-GEN tenant-GEN gray sofa-ACC place-PART-3SG.POSS
NP1 NP6

b. ENC, LowSyn-HighPhono

∅ [İşinin ehli marangoz-lar-∅] || [nakliyeciler-in] [kiracı-nın koltuğ-un minder-leri-ni yerleştirdiği] ...
Pro expert carpenter-PL-NOM mover-PL-GEN tenant-GEN sofa-GEN cushion-3PL.POSS-ACC place-PART-3SG.POSS
NP1 NP6

c. ENC, LowSyn-LowPhono

∅ [İşinin ehli marangoz-lar-∅] || [nakliyeciler-in] [kiracı-nın gri koltuğ-u yerleştirdiği] ...
Pro expert carpenter-PL-NOM mover-PL-GEN tenant-GEN gray sofa-ACC place-PART-3SG.POSS
NP1 NP6

... oda-ya taşıdıkları dolab-ı || dikkatli şekilde kurduklarını/kurduklarını san-dım.
room-DAT move-PART-3PL.POSS closet-ACC carefully set.up-VN-3PL.POSS-ACC/set.up-PF-3PL think-PF-1SG
VP1

Followed by a content-neutral two-word follow-up sentence □ Yakında taşınacaklarmış.

'I thought that the expert carpenters carefully set up the closet that the movers moved to the room where the tenant placed the gray sofa. I heard they would move in soon.'

Materials (Discouraging Phrase Lengths)

d. DISC, HighSyn-HighPhono

∅ [Marangoz-lar-in] || [nakliyeciler-in] [kiracı-nın] oldukça geniş gri koltuğ-u büyük özen-le ...
Pro carpenter-PL-GEN mover-PL-GEN tenant-GEN extremely large gray sofa-ACC great care-with
NP1 NP6

e. DISC, LowSyn-HighPhono

∅ [Marangoz-lar-∅] || [nakliyeciler-in] [kiracı-nın] koltuğ-un oldukça geniş minder-leri-ni büyük özen-le ...
Pro carpenter-PL-NOM mover-PL-GEN tenant-GEN sofa-GEN extremely large cushion-3PL.POSS-ACC great care-with
NP1 NP6

f. DISC, LowSyn-LowPhono

∅ [Marangoz-lar-∅] || [nakliyeciler-in] [kiracı-nın] oldukça geniş gri koltuğ-u büyük özen-le ...
Pro carpenter-PL-NOM mover-PL-GEN tenant-GEN extremely large gray sofa-ACC great care-with
NP1 NP6

... yerleştirdi-i] oda-ya taşıdıkları dolab-ı || kurdukları-nı/kurdu-lar] san-dı-m...
place-PART-3SG.POSS room-DAT move-PART-3PL.POSS closet-ACC set.up-VN-3PL.POSS-ACC/set.up-PF-3PL think-PF-1SG
VP1

Followed by a content-neutral two-word follow-up sentence □ Yakında taşınacaklarmış.

'I thought that the carpenters set up the closet that the movers moved to the room where the tenant placed the extremely large gray sofa with great care. I heard they would move in soon.'

Data Analysis

Six eye-tracking measures:

Rayner et al., 1989)

- First fixation duration (FD)
- Gaze duration (GD)
- Regression path duration (RPD)
- Rereading duration (RRD)
- Total duration (TD)
- Probability of regression out (PRO)



Early processes of comprehension



Later processes of comprehension

• Two sets of analyses:

1st set: Prosodic Phrase Lengths & Case Interference (comparison of HighSyn–HighPhono and LowSyn–HighPhono to LowSyn–LowPhono)

2nd set: Prosodic Phrase Lengths & Case Interference (comparison of LowSyn–HighPhono to HighSyn–HighPhono)

Mixed effects linear regression models (for FD, GD, RPD, RRD and TD) or logistic regression models (for PRO)

(Bates et al., 2015)

- Prosodic phrase lengths & case interference as fixed effects
- Subjects & items as random effects
- Random slopes for subjects and/or items by the predictor variables when the model allowed

(Barr et al., 2013)

R packages

- *lme4* (Bates & Maechler, 2013)
- *languageR* (Baayen, 2013)
- *influence.ME* (Nieuwenhuis et al., 2012)

Descriptive Results (Case Interference)

Critical Region

	FFD	GD	RPD	RRD	TD	PRO
HighSyn-HighPhono, ENC	231 (6)	380 (14)	428 (19)	315 (31)	661 (31)	.10 (.02)
LowSyn-HighPhono, ENC	233 (5)	318 (12)	363 (17)	253 (26)	565 (28)	.11 (.02)
LowSyn-LowPhono, ENC	238 (7)	333 (12)	418 (24)	275 (30)	602 (30)	.15 (.03)
HighSyn-HighPhono, DISC	239 (6)	409 (16)	554 (27)	434 (36)	793 (35)	.20 (.03)
LowSyn-HighPhono, DISC	247 (8)	347 (14)	462 (23)	327 (32)	633 (31)	.21 (.03)
LowSyn-LowPhono, DISC	239 (6)	324 (13)	417 (21)	352 (32)	648 (31)	.19 (.03)

Spill-over Region

	FFD	GD	RPD	RRD	TD	PRO
HighSyn-HighPhono, ENC	244 (7)	283 (10)	373 (27)	146 (18)	408 (19)	.15 (.03)
LowSyn-HighPhono, ENC	238 (6)	267 (9)	332 (18)	180 (23)	439 (24)	.19 (.03)
LowSyn-LowPhono, ENC	239 (7)	290 (11)	378 (24)	178 (26)	447 (25)	.19 (.04)
HighSyn-HighPhono, DISC	247 (7)	280 (19)	434 (26)	210 (27)	463 (24)	.26 (.04)
LowSyn-HighPhono, DISC	249 (7)	297 (11)	416 (27)	170 (22)	466 (26)	.26 (.04)
LowSyn-LowPhono, DISC	251 (8)	278 (11)	416 (27)	192 (22)	465 (22)	.24 (.04)

Descriptive Results (Phrase Lengths)

Critical Region

	FFD	GD	RPD	RRD	TD	PRO
HighSyn-HighPhono, ENC	231 (6)	380 (14)	428 (19)	315 (31)	661 (31)	.10 (.02)
LowSyn-HighPhono, ENC	233 (5)	318 (12)	363 (17)	253 (26)	565 (28)	.11 (.02)
LowSyn-LowPhono, ENC	238 (7)	333 (12)	418 (24)	275 (30)	602 (30)	.15 (.03)
HighSyn-HighPhono, DISC	239 (6)	409 (16)	554 (27)	434 (36)	793 (35)	.20 (.03)
LowSyn-HighPhono, DISC	247 (8)	347 (14)	462 (23)	327 (32)	633 (31)	.21 (.03)
LowSyn-LowPhono, DISC	239 (6)	324 (13)	417 (21)	352 (32)	648 (31)	.19 (.03)

Spill-over Region

	FFD	GD	RPD	RRD	TD	PRO
HighSyn-HighPhono, ENC	244 (7)	283 (10)	373 (27)	146 (18)	408 (19)	.15 (.03)
LowSyn-HighPhono, ENC	238 (6)	267 (9)	332 (18)	180 (23)	439 (24)	.19 (.03)
LowSyn-LowPhono, ENC	239 (7)	290 (11)	378 (24)	178 (26)	447 (25)	.19 (.04)
HighSyn-HighPhono, DISC	247 (7)	280 (19)	434 (26)	210 (27)	463 (24)	.26 (.04)
LowSyn-HighPhono, DISC	249 (7)	297 (11)	416 (27)	170 (22)	466 (26)	.26 (.04)
LowSyn-LowPhono, DISC	251 (8)	278 (11)	416 (27)	192 (22)	465 (22)	.24 (.04)