Expectations and Noisy-Channel Processing of Relative Clauses in Standard Arabic

HUMAN SENTENCE PROCESSING 2022

NICOLE DODD & EMILY MORGAN

(UNIVERSITY OF CALIFORNIA, DAVIS)

Theories in sentence processing

Memory-based theories (e.g., Gibson, 1998) predict greater processing difficulty when reading structures that require holding unresolved elements in memory, such as long-distance dependencies.

Expectations-based theories (e.g., Hale, 2001) predict that high-frequency structures are easier to process.

Case study: subject and object relative clauses

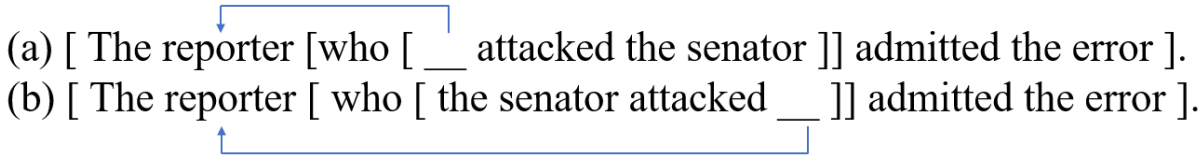


Figure 1: (a) Example SRC in English; (b) Example ORC in English. Gap dependencies are illustrated in blue.

In subject-extracted relative clauses (SRCs), the noun phrase head of the matrix clause is also the subject of the relative clause (Figure 1a).

In object-extracted relative clauses (ORCs), the noun phrase head of the matrix clause is the object of the relative clause (Figure 2b).

Processing patterns for these structures are not consistent across languages; these differences are due to word order (e.g., SVO vs SOV), clause headedness (head-initial or head-final), position of the relative clause in the sentence (pre-nominal or post-nominal), animacy, semantic roles, etc.

We contribute to this body of research by studying an under-represented language in psycholinguistic research: **Standard Arabic (SA)**.

SRCs and ORCs in SA

الصحفي الذي هاجم السناتور اعترف بالخطأ

aː=sˤaħafi-u	aːla-ði	h<aː>ʒam	aː=siːnaːtuːr	<ʔ>ʕ<ta>rafa	b=il=xaðˤʕ-i
DET=reporter-NOM	who-3SG.M	attack<3SG.M.PST>	DET=senator	admit<3SG.M.PST>	to=DET=error-ACC

SRC: “The reporter who attacked the senator admitted the error.”

الصحفي الذي هاجمه السيناتور اعترف بالخطأ

aː=sˤaħafi-u	aːla-ði	h<aː>ʒam=ahu	aː=siːnaːtuːr	<ʔ>ʕ<ta>rafa	b=il=xaðˤʕ-i
DET=reporter-NOM	who-3SG.M	attack<3SG.M.PST>=3SG.M.ACC	DET=senator	admit<3SG.M.PST>	to=DET=error-ACC

ORC: “The reporter who the senator attacked admitted the error.”

Figure 2: Sample stimuli. Arabic stimuli are read right to left, and English stimuli are read left to right. The red circles indicate the area of interest: the relative clause verb. Arabic word order is such that the only difference between an SRC and an ORC (with matched gender and number for matrix and relative clause nouns) is the inclusion of an object pronoun clitic.

Theoretical predictions

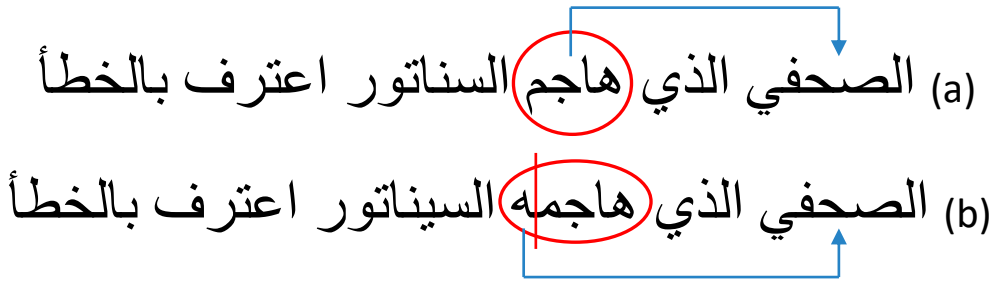


Figure 2: Arabic matrix clause subject dependency in (a) the SRC and (b) the ORC condition. The disambiguating region, the relative clause verb, is circled in red. The red vertical line on the ORC verb delineates the object pronoun clitic from the relative clause verb.

Memory-based constraints theories predict equal processing difficulty in both conditions

- Dependency length from disambiguating region to matrix clause subject is the same in both conditions

Expectations-based theories predict easier processing in SRCs

- Frequency confirmed through corpus analysis: 71% SRCs and 29% ORCs

Experiment 1: self-paced reading task

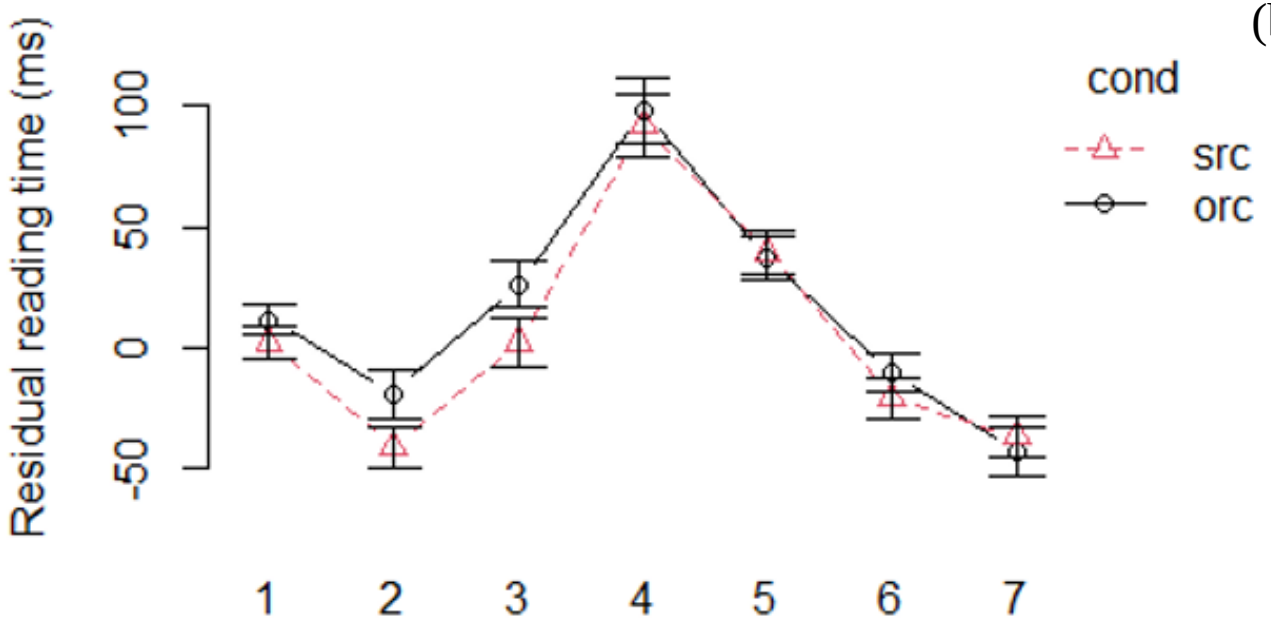
40 pairs of sentences with an SRC and ORC alternation

Native Arabic speakers (n = 48) read a total of 120 sentences

- 40 stimuli in either SRC or ORC condition + 80 filler sentences

Yes/no comprehension questions appeared after all experimental and 20 filler items

(a)



(b)

1: matrix clause subject + relative pronoun
e.g., "The reporter who" (الصحفي الذي)

2: relative clause verb + object RP (in ORC cond)
e.g., "attacked (him)" (هاجمه)

3: relative clause noun
e.g., "the senator" (السيناتور)

4: matrix clause verb
e.g., "admitted" (اعترف)

5: spillover

6: spillover

7: spillover

Figure 4: (a) Average residualized reaction times for each region by clause type; (b) Regions of interest with Arabic examples and their English gloss.

Results

SRCs are easier to process than ORCs in Standard Arabic, in line with expectations-based theories of sentence processing.

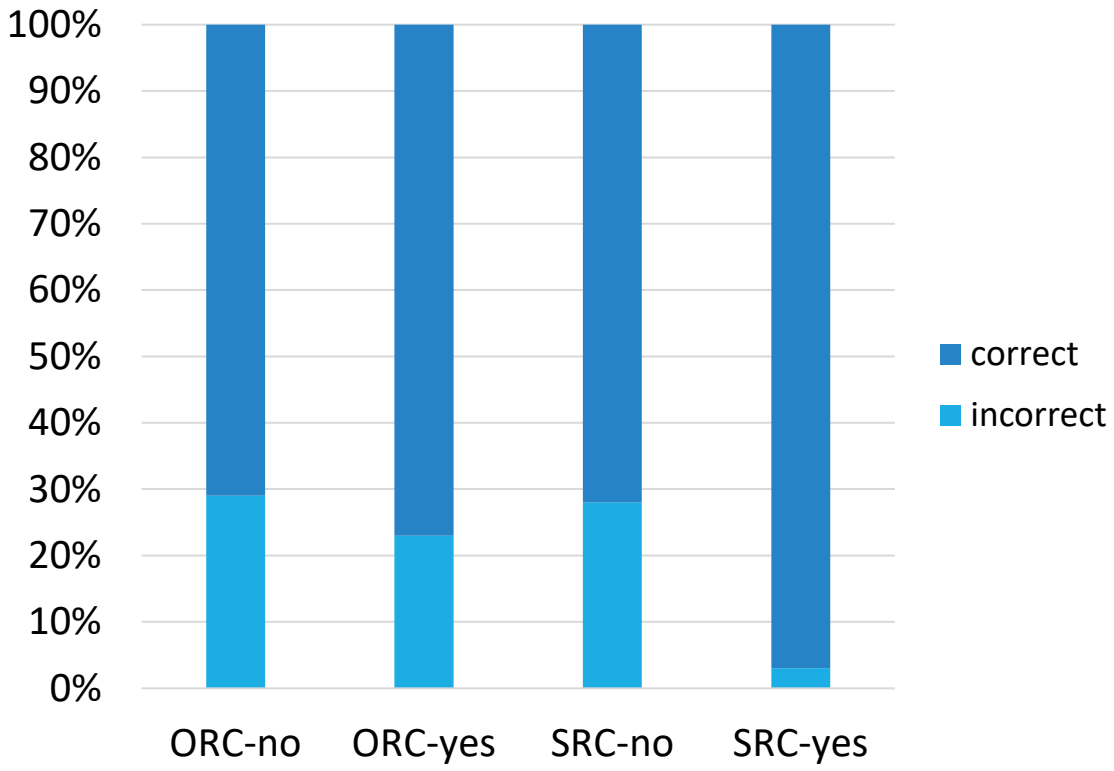


Figure 5: Proportion of correct comprehension question answers by clause type (SRC or ORC) and correct answer condition (correct answer is “yes” or “no”).

Comprehension question analysis

We analyzed answers to the comprehension questions to identify whether participants were accurately interpreting ORCs.

We found significant effects by both clause type and correct answer condition:

- more incorrect answers for ORCs
- more incorrect answers when the correct answer was “no” (i.e., false positives)
- subadditive effect for clause type and correct answer condition

This leads us to believe that readers mistakenly interpret ORCs as SRCs, *especially* when a noisy interpretation is suggested by the comprehension question via a bias for “yes” answers.

Experiment 2: recall task

Are readers skipping the resumptive pronoun in ORCs during reading, or accepting a noisy SRC interpretation given their expectations? (Keshev & Meltzer-Asscher, 2021)

We conducted a recall task using the same stimuli where participants (n = 80) read each sentence and reproduced the sentence word-for-word from memory.

Participants both misremembered ORCs as SRCs (71% of errors) and SRCs as ORCs.

This lends some support to noisy-channel processing as opposed to misreading.

Conclusions and future directions

Arabic readers have less processing difficulty when reading SRCs as compared to ORCs, lending support to expectations-based theories of sentence processing.

Questions remain as to how readers process the resumptive pronoun in the ORC condition: are they skipping the pronoun, or accepting a noisy interpretation given their expectations?

Preliminary experimental evidence lends support to noisy-channel processing with resumptive pronouns in Arabic.

Our future research will explore noisy-channel processing behavior through eye tracking, and explore how semantic, grammatical, and orthographic features influence a reader's willingness to accept a noisy interpretation.

References

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76.

Hale, J. (2001). A probabilistic earley parser as a psycholinguistic model. *Proceedings of NAACL '01*, 1–8.

Keshev, M., & Meltzer-Asscher, A. (2021). Noisy is better than rare: Comprehenders compromise subject-verb agreement to form more probable linguistic structures. *Cognitive Psychology*, *124*, 1–58.