

## Eye-Tracking Measures of Processing Difficulty in Standard Arabic

Nicole Dodd<sup>1</sup>, Fatima Boush<sup>2</sup>, Tommi Leung<sup>2</sup>, Fernanda Ferreira<sup>1</sup>, Emily Morgan<sup>1</sup>

<sup>1</sup>University of California, Davis, <sup>2</sup>United Arab Emirates University

What makes some sentences more difficult to read and comprehend than others? Both working memory and expectations affect processing difficulty: memory-based constraints theories predict more difficulty when reading cognitively demanding structures (e.g., long-distance dependencies) [1], while expectation-based theories predict more difficulty when reading low-frequency structures [2, 3]. Both can contribute to processing difficulty simultaneously and have been shown to manifest in distinct processing behaviors [4]. Misaligned expectations during reading can also result in the misinterpretation of a sentence. Under good-enough and noisy-channel processing models, language users weigh the probability of a structure against the probability of noisy input and may accept a noisy yet higher probability interpretation [5, 6]. These theories have been tested cross-linguistically using subject- and object-extracted relative clauses (SRCs and ORCs) [7]. We expand this body of work by investigating Standard Arabic, an under-represented language in psycholinguistic research. Previous work found that ORCs are harder to process than SRCs in Arabic, supporting expectation-based theories [8]; however, the results did not preclude difficulty due to memory constraints, and the analysis of comprehension question accuracy suggested good-enough or noisy-channel processing with ORCs being misinterpreted as SRCs.

Our study built upon this work by conducting an eye-tracking experiment. 47 native Arabic speakers from the United Arab Emirates University read 80 sentences with an SRC/ORC manipulation (Fig 1). Comprehension questions targeted RC interpretation. We first assessed whether both memory limitations and violated expectations contributed to processing difficulty and whether we observed distinct processing behaviors for each (longer go-past reading times and increased regressive saccades respectively) [4]. We then asked whether readers were initially misreading ORCs as SRCs, or correctly reading ORCs yet accepting noisy SRC interpretations [6]. On the one hand, Arabic SRCs and ORCs differ by only a resumptive object pronoun (RP) clitic, which may be easily missed. On the other hand, SRCs are more frequent than ORCs, so readers may accept a noisy but preferred SRC interpretation.

We conducted three analyses (Table 1). First, we compared SRCs and ORCs with correct comprehension question answers to determine differences during veridical processing. We found no significant effects at the RC verb. At the RC noun phrase, ORCs were associated with significantly longer reading times and higher regression rates. This indicates processing difficulty from both memory limitations and expectations [4]; however, the difficulty was predicted to have been localized to the disambiguating region (the RC verb). Our data suggest that readers pay a processing cost when integrating the RC NP in the globally less-expected ORC structure, even though they had previously received the disambiguating information. These results are not predicted by either the memory- or expectation-based theories.

We then compared incorrect ORC items to correct ORC items, and then to correct SRC items to investigate misreading versus good-enough/noisy-channel processing (Table 2). We found no significant differences in reading times between correct and incorrect ORC items at the RC verb, suggesting that readers were not misreading the verb when misinterpreting ORCs. Rather, correct ORCs had longer reading times and higher regression rates at the RC NP, consistent with our first analysis. Our comparison of incorrect ORC items and correct SRC items revealed no significant differences at either the RC verb or NP. Taken together, our results support noisy-channel processing over misreading. Correctly interpreting an ORC is dependent on the successful integration of the updated structure at the RC NP, causing increased difficulty in ORCs versus SRCs and correct versus incorrect ORCs. Accepting a noisy SRC interpretation of an ORC skips this integration and thus does not lead to processing time costs, resulting in comparable processing difficulty between incorrect ORCs and correct SRCs.

spillover-1	matrix-VERB	rel-NP	rel-VERB	rel-PRON	matrix-NP
بالخطأ	اعترف	السناتور	هاجم	الذي	الصحفي (a)
b=il=xaðʕ-i	<ʔ>ʕ<ta>rafa	a:=si:na:tu:r	h<a:>ʒam	a:la-ði	a:=sʕahafi-u
to=DET=error-ACC	admit<3SG.M.PST>	DET=senator	attack<3SG.M.PST>	who-3SG.M	DET=reporter-NOM
SRC: "The reporter who attacked the senator admitted the error."					
بالخطأ	اعترف	السناتور	هاجمه	الذي	الصحفي (b)
b=il=xaðʕ-i	<ʔ>ʕ<ta>rafa	a:=si:na:tu:r	h<a:>ʒam=ahu	a:la-ði	a:=sʕahafi-u
to=DET=error-ACC	admit<3SG.M.PST>	DET=senator	attack<3SG.M.PST>= 3SG.M.ACC	who-3SG.M	DET=reporter-NOM
ORC: "The reporter who the senator attacked admitted the error."					

**Figure 1:** Sample (a) SRC and (b) ORC stimuli. Arabic sentences and English glosses are read right to left. The red circles indicate the disambiguating region: the relative clause verb. Arabic grammatical features entail 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 a bound relative object pronoun clitic.

Model Parameters	Cor SRC vs. Cor ORC	Cor ORC vs. Incor ORC	Cor SRC vs. Incor ORC
<i>Go-past time</i>	38.66* [-2.65, 79.61]	59.94* [8.58, 110.93]	-24.33 [-68.44, 19.02]
<i>Total fixation duration</i>	59.97* [24.67, 94.84]	76.91* [26.58, 127.94]	-27.90 [-73.32, 17.96]
<i>First pass regression</i>	0.11* [0.00, 0.23]	0.26* [0.07, 0.46]	-0.13 [-0.35, 0.06]

**Table 1:** *brms* mixed-effect model estimates and credible intervals (CI) for the main predictor variable for all three analyses at the rel-NP. Estimates marked with an asterisk are significant. We consider the model estimates as reliable if the CI does not include 0, or over 95% of the sampled posterior distribution is over or under 0 in the predicted direction.

Theory	Comparison	Predicted behavior at disambiguating region (rel-VERB)
<i>Misreading</i>	Cor ORC vs. Incor ORC	Decreased processing difficulty from skipping RP clitic
	Cor SRC vs. Incor ORC	No significant difference; incorrect ORC misread as SRC
<i>Good-enough/noisy-channel processing</i>	Cor ORC vs. Incor ORC	No significant difference; both read without skipping RP clitic
	Cor SRC vs. Incor ORC	Increased processing difficulty; incorrect ORC read similarly to correct ORC as in veridical processing

**Table 2:** Predicted behavioral outcomes if misreading versus good-enough/noisy-channel processing is occurring.

**References:** [1] Gibson (1998). *Cognition* 68(1). [2] Hale (2001). *Proceedings of NAACL '01*. [3] Levy (2008). *Cognition* 106(3). [4] Staub (2010). *Cognition* 116(1). [5] Ferreira et al. (2002). *Curr Dir Psychol Sci* 11(1). [6] Keshev & Meltzer-Asscher (2021). *Cognitive Psychology* 124. [7] Lau & Tanaka (2021). *Glossa: a journal of general linguistics* 6(1). [8] Dodd & Morgan (2022). *Proceedings of the Annual Meeting of the Cognitive Science Society* 44(44).