Expectations and Noisy-Channel Processing of Relative Clauses in Arabic

Nicole Dodd (ncdodd@ucdavis.edu)
Department of Linguistics, University of California, Davis
One Shields Avenue, Davis, CA 95616 USA

Emily Morgan (eimorgan@ucdavis.edu)
Department of Linguistics, University of California, Davis
One Shields Avenue, Davis, CA 95616 USA

Abstract

Some sentences are hard to read, and we don't fully understand why. Memory-based and expectation-based constraints both attempt to explain sentence processing difficulties, and decades of sentence processing literature have found evidence in support of both theories. We further investigate theories of sentence processing by exploring subject- and object-extracted relative clause processing in Standard Arabic. We conducted a self-paced reading task and found that SRCs are easier to process than ORCs in Arabic, in line with expectation-based theories. A follow-up analysis of comprehension question answers revealed that when suggested with the possibility of a noisy interpretation, readers preferentially accept an SRC interpretation over an ORC interpretation. Our future research will explore these findings and test the threshold for acceptance of noisy interpretations.

Keywords: sentence processing; noisy channel processing; Standard Arabic; resumptive pronouns

Introduction

An outstanding question in psycholinguistic research is what makes sentences more difficult to process and why. Two main types of theories – memory-based theories (e.g., Gibson, 1998; Gibson, 2000) and expectation-based constraints theories (e.g., MacDonald et al., 1994; Hale, 2001; Levy, 2008a) – aim to explain these difficulties.

Memory-based theories propose that syntactic structures with longer dependencies – or structures that utilize more working memory during incremental processing – are more difficult to process. Humans have limited computational resources, and when constituents with incomplete dependencies are maintained in memory at length, this incurs higher processing costs. An additional processing cost is also paid upon integrating the dependency with the existing structure of the sentence. This phenomenon is formalized in the Dependency Locality Theory (Gibson, 2000), which states that the cost of processing an element is directly proportional to the length of its dependency. Thus, memory-based theories predict more processing difficulty in structures with long dependencies.

Expectation-based theories, on the other hand, posit that items that are low frequency in context are more difficult to process. During incremental processing, readers use grammatical and semantic contextual cues to predict upcoming words and the sentence's overall structure. Processing difficulty arises when a reader encounters an unexpected word, or a word that violates their expectations for the resulting syntactic parse. When expectations are violated, the reader pays a processing cost proportional to the difficulty of updating their expectations. Many expectation-based theories operationalize this cost using surprisal theory, calculated as the negative log-probability of a word given previous context (Hale, 2001; Levy, 2008a). So, expectation-based theories predict more processing difficulty in structures that are low frequency or have a low probability in context.

Violated expectations not only cause increased processing difficulty, but can also lead to the acceptance of the wrong overall interpretation of the sentence. This is the case in models of rational noisy-channel processing (Levy, 2008b; Levy, 2011). Language input takes place in noisy circumstances – such as human error and competing environmental conditions – and this noise affects language processing strategies. Noisy-channel processing theories thus suggest that language users weigh the probability of a given sentence structure against the probability of noisy input during sentence processing. In cases where different syntactic structures are possible but one is higher probability than the other, a reader may assume noise in the input and make a number of "edits" to a sentence to arrive at the higher-probability interpretation. In such cases, readers both experience increased processing difficulty when encountering violated expectations, and accept the wrong, but more probable, interpretation of the sentence.

One popular structure for testing sentence processing theories is the relative clause (RC); more specifically, the subject- and object-extracted relative clause. In subject-extracted relative clauses (SRC), the noun phrase head of the matrix clause is also the subject of the RC; in object-extracted relative clauses (ORC), the noun phrase head of the matrix clause is the object of the RC (see Figure 1).

(a) [The reporter [who [the senator attacked]] admitted the error].
(b) [The reporter [who [attacked the senator]] admitted the error].

Figure 1: (a) Example ORC in English; (b) Example SRC in English. Dependencies between the RC verb and matrix clause subject are illustrated in blue.

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

Early studies assumed a cross-linguistic tendency for more processing difficulty when reading ORCs compared to SRCs, based on evidence from English (e.g., King & Just, 1991), German (Schriefers et al., 1995), Dutch (Mak et al., 2002), Japanese (Ueno & Garnsey, 2008), and Korean (Kwon et al., 2010). Subsequent studies, however, found that this "subject advantage" in RC processing was not universal; in Chinese (Hsiao & Gibson, 2003) and Basque (Carreiras et al., 2010), SRCs were *harder* to process than ORCs. Further, findings from previous cross-linguistic studies do not conclusively support one processing theory over the other. Some studies (e.g., Konieczny & Doring, 2003) have found evidence directly in support of and in contradiction to one theory, while others (e.g., Staub, 2010) have found evidence that both constraints contribute to processing costs.

Many of the differences in these cross-linguistic findings can be attributed to typological factors such as word order (e.g., SVO vs. SOV), clause-headedness (head-initial vs. head-final), RC position in a sentence (pre-nominal vs. post-nominal), and the use of resumptive pronouns (RPs) in gap processing (Lau & Tanaka, 2021). Languages with different typological features are not evenly represented in previous research; for example, more research has been done on SVO and SOV languages than VSO languages. Thus, future research must focus on typologically under-represented languages to better understand the cross-linguistic principles that govern sentence processing. Our research takes steps towards diversifying this body of research by studying a morphosyntactically-complex language that is under-represented in psycholinguistic literature: Standard Arabic.

Standard Arabic

Standard Arabic (SA) is a Semitic language written right-to-left and uses alternating SVO and VSO word order. SA is mainly used in official governmental or media domains, so native Arabic speakers typically learn both the Standard dialect and a regional dialect for every-day communication.

Arabic RCs exhibit linguistic properties that are under-represented in sentence processing literature, namely VSO word order and the use of grammaticalized resumptive pronouns (RPs). A sample SRC and ORC demonstrating these features are shown in Figure 3. In our stimuli, RPs appear in the ORC condition as a bound clitic on the RC verb. Many regional dialects require the use of resumptive object pronoun clitics; SA requires them only in sentences with an indefinite matrix subject, but they are still preferred. Previous literature on RP processing has shown inconclusive evidence as to whether RPs help or hinder processing and comprehension (Meltzer-Asscher, 2021); however, RPs are grammaticalized in SA and provide syntactic information that can aid in processing. Our stimuli utilize these linguistic properties in order to further explore the effect of these features on patterns of sentence processing.

While SA and its regional dialects are the sixth most spoken language group in the world, psycholinguistic research in Arabic is sparse (Hermena, 2016). Thus, our research aims to both diversify existing language processing literature, and augment existing research in Arabic psycholinguistics. The present study explores processing difficulty in SRCs and ORCs for native SA speakers. We conducted a self-paced reading task in which participants were presented with stimuli using a self-paced word-by-word reading paradigm. Reading times (RTs) were collected for each word and used as a proxy for processing difficulty. Overall processing difficulty was determined by comparing residualized RTs (i.e., RTs controlled for word length) across clause condition (SRC vs. ORC) to identify which construction had longer average RTs. The findings from our self-paced reading task demonstrated an interesting trend in relative clause comprehension; therefore, we additionally conducted a follow-up recall task to explore the possibility of noisy-channel processing in ORCs in SA.

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

Memory-based theories grounded in the DLT predict comparable processing times for SRCs and ORCs in Arabic; because the RP in the ORC condition is cliticized to the relative clause verb, the dependency between the disambiguating region (the relative clause verb) and the matrix clause subject have the same total length in both conditions (see Figure 2). To determine what expectation-based theories would predict, we conducted a corpus analysis using the Penn Arabic Treebank Part 3 v 3.2 corpus (Maamouri et al., 2010). We used two search parameters: RCs with an explicit الذي\التي ("who") relative pronoun, and RCs that used either a الذي\التي ("who") or the more general ما ("that") relative pronoun. We identified 2,928 RCs with an explicit "who" relative pronoun, of which 71% were SRCs. When including the general "that" relative pronoun, we identified 7,268 RCs, of which 79% were SRCs. Expectation-based theories, then, predict shorter processing times in SRCs.

Methods

Participants

48 native Standard Arabic speakers (mean age: 27; sd: 7.33) were recruited from Prolific. Participants were paid $7.80/hour for their participation. Of our 48 participants, 44 reported that they were fluent in a regional dialect in addition to Standard Arabic. Any participant who scored lower than 75% accuracy on the comprehension questions was excluded from the analysis.

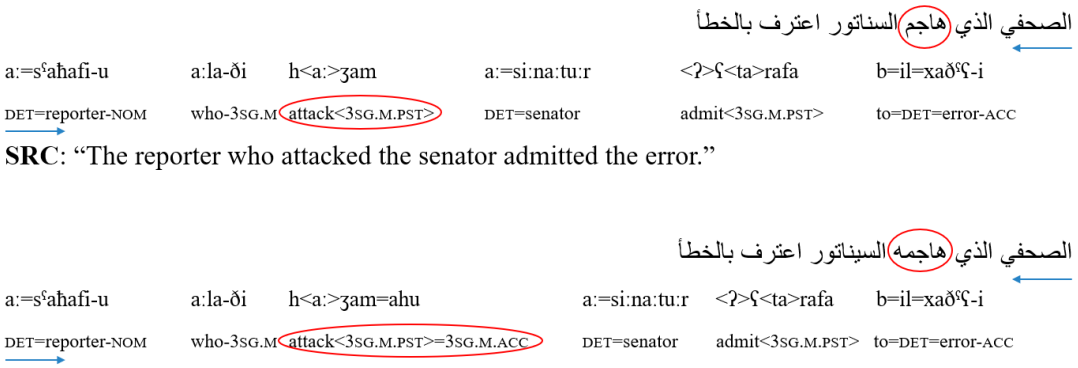

Figure 3: Sample stimuli. Arabic sentences are read right to left, and English are read left to right. The red circles indicate the area of interest: the relative clause verb.

Materials

The stimuli are designed to take advantage of the flexible word order in Arabic to minimize variation between clause conditions. For each sentence, the matrix clause is SVO, and the RC is VSO. This word order was selected with the help of native speaker input – when presented with both a SVO and a VSO embedded relative clause, the native speaker preferred the VSO word order. This intuition was then confirmed with a frequency analysis from the Penn Arabic Treebank Part 3 v 3.2 corpus, in which we identified that 98% of RCs in the corpus used VSO word order.

Given this word order, readers first read the matrix clause subject, followed by the relative pronoun, and then the RC verb in both conditions. The key difference between the SRC and the ORC condition is the presence of the resumptive object pronoun in the ORC condition as a bound clitic on the RC verb (see Figure 3).

Stimuli are adapted and translated from previous studies on relative clause processing (Gordon et al., 2001; Traxler et al., 2002; Staub, 2010). Arabic nouns, verbs, and pronouns are marked for both number and gender; thus, we matched the matrix and RC nouns on number and gender so that the head of the RC would not be disambiguated by number and gender marking. All nouns are animate to control for animacy effects. The ORC condition of each stimulus includes an object RP bound to the RC verb; these are not required for definite clauses in SA, but are preferred. Finally, all stimuli were presented in a non-diacritized format, as is standard for written publications in SA.

A norming study was conducted to confirm that the subject and object of each RC were equally plausible in both clause conditions (e.g., "the reporter attacked the senator" is as plausible as "the senator attacked the reporter"). Native SA speakers (n = 80; mean age: 28; sd: 6.72) were recruited through Prolific and asked to rate the plausibility of each sentence on a Likert scale (1 = highly implausible, 7 = highly plausible). Participants for the stimuli norming task and the self-paced reading task did not overlap. Plausibility ratings were collected for both the full stimuli sentences (e.g., "The reporter who attacked the senator admitted the error" and "The reporter who the senator attacked admitted the error") and the relative clauses as simplified transitive sentences (e.g., "The reporter attacked the senator" and "The senator attacked the reporter"). The study also included implausible distractor sentences (e.g., "The laptop angrily finished the homework") as an attention check. Four stimuli were excluded after a paired t-test revealed substantial discrepancies between plausibility ratings in the SRC and ORC conditions for those items, and one stimulus was excluded for low overall ratings. After exclusion, we had 40 stimuli sentences. The mean plausibility rating for the full stimuli sentences was 6.11 (sd: 0.49) for SRCs and 6.00 (sd: 0.57) for ORCs, and the mean plausibility rating for the simplified transitive sentences was 6.18 (sd: 0.46) for SRCs and 6.08 (sd: 0.49) for ORCs.

In addition to the 40 target sentences, 80 unrelated filler sentences were included. Comprehension questions appeared after all 40 stimuli sentences and 20 filler sentences. Of the 40 stimuli comprehension questions, half of the questions targeted comprehension of the relative clause (e.g., "Did the reporter attack the senator?"), and the other half targeted comprehension of the sentence overall (e.g., "Did the reporter admit the error?"). 20 general comprehension questions appeared after approximately one third of the filler sentences. "Yes" and "no" correct answers were balanced within question type. In total, each participant read 120 sentences (40 target sentences (20 for each clause type) + 80 filler sentences = 120 total sentences) and answered 60 comprehension questions. Experimental items were counterbalanced in a Latin square design.

Procedure

Subjects were told that they would be reading sentences and answering comprehension questions. All experimental instructions were given in Arabic. Prior to the start of the experiment, participants saw two practice stimuli and answered one practice comprehension question. Each sentence was presented word-by-word using a subject-paced paradigm in which participants used the spacebar to advance through the sentence. Each word in the sentence was presented in isolation with no option to move backward in the

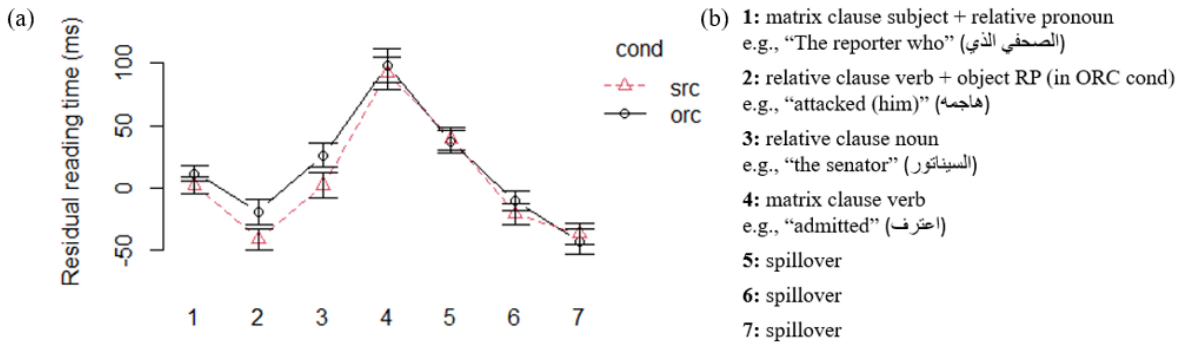

Figure 5: (a) Average residualized RTs for each region by clause type (after data preprocessing); (b) Regions of interest with Arabic examples and their English gloss.

sentence. RTs were collected at each key stroke indicating the appearance of the next word in the sentence. The experiment took between 30 and 45 minutes on average.

Analysis

Each sentence was divided into target regions for our analysis (see Figure 4). The matrix clause noun and relative clause pronoun are grouped into one region as they did not differ across clause type condition, and were not a region of interest. Regions of interest included the RC verb, including the RP in the ORC condition (Region 2), the RC noun (Region 3), and the matrix clause verb (Region 4). These are our regions of interest as we expect to see higher processing costs within the relative clause in cases of higher processing difficulty, plus one spillover region (the matrix clause verb). We also analyzed three additional spillover regions (Regions 5 through 7).

[الصحفي الذي]$_1$ [هاجمه]$_2$ [السيناتور]$_3$ [اعترف]$_4$ [بالخطأ]$_5$

[The reporter who]$_1$ [attacked (him)]$_2$ [the senator]$_3$ [admitted]$_4$ [the error]$_5$

1: matrix clause subject + relative pronoun
2: relative clause verb + object pronoun (where applicable)
3: relative clause noun phrase
4: matrix clause verb
5: spillover region
6: spillover region (where applicable)
7: spillover region (where applicable)

Figure 4: Target regions for RT analysis.

Prior to our analysis, RTs shorter than 100 ms and longer than 2,000 ms were excluded, and RTs were residualized within-subject to control for word length (Ferreira and Clifton, 1986). To determine whether SRCs or ORCs were more difficult to process, we first summed residualized RTs for each trial across a pre-selected region of interest (Regions 2-4) and asked whether reading times varied by clause type. In a post hoc analysis, we analyzed RTs for each individual region (Regions 1-7) by clause type. For each analysis, we fit linear mixed effects models to the data using the brms package in R (Bürkner, 2018). We modeled residual RTs with clause type (SRC or ORC) as a sum-coded fixed effect. Effects are considered to be reliable if their 95% Credible Interval does not include 0. We included the maximal random effects structure justified by the design (Barr et al., 2013), namely random intercepts and random slopes of clause type both by participant and by item.

For each analysis, we ran two models: one that included all RT data, and one that included RT data only from items in which participants got the comprehension question correct.

Results

Reading Times

Average residualized RTs in each region by clause type are plotted in Figure 5, and raw RTs, trending in similar directions as the residualized RTs, are additionally shown in Figure 6. Negative residual RTs indicate shorter processing times given word length, and positive residual RTs indicate longer processing times given word length.

Main region of interest We first fit two models to RT data summed across our region of interest (Regions 2-4) – one with all RT data, and one with only RT data from items with correct comprehension question answers. Both models showed an effect for clause type on RTs; SRCs tended to have shorter RTs compared to ORCs. This effect, however, was only reliable in the model with all data (All data: β = -62.92; se: 26.72; CI: [-113.47, -11.14]; Correct comprehension question data: β = -61.03; se: 31.93; CI: [-126.28, 0.53]).

Individual regions In a post-hoc analysis, we additionally fit individual models to each region (Regions 1-7), also for both groups of data (all vs. correct comprehension question data). Models for Regions 2 and 3 (RC noun and RC verb respectively) showed an effect for clause type that matched trends from our region of interest models. This estimate was reliable in the models with all data (Region 2: β = -32.40; se: 13.09; CI: [-59.02, -6.26]; Region 3: β = -32.00; se: 14.87; CI: [-62.07, -1.92]), and was only reliable in Region 2 in the model with correct comprehension question data (β = -29.35; se: 14.89; CI: [-57.93, -0.54]).

Comprehension Questions

Our initial review of comprehension question accuracy revealed higher-than-expected error rates. Specifically, we

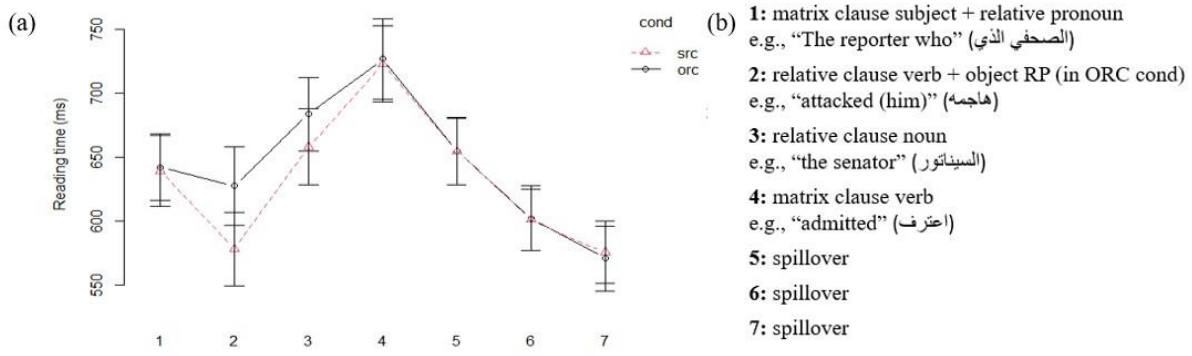

Figure 6: (a) Average raw RTs for each region by clause type (after data preprocessing); (b) Regions of interest with Arabic examples and their English gloss.

observed that participants performed markedly worse on comprehension questions that targeted clausal comprehension (i.e., the correct SRC or ORC interpretation) than on questions that targeted overall sentence comprehension. To investigate this pattern, we analyzed comprehension question answers by clause type (SRC vs. ORC) and question type ("Yes" correct answer vs. "No" correct answer).

80 total comprehension questions were used in the experiment: 40 that targeted relative clausal comprehension (20 stimuli x 2 conditions (SRC or ORC) = 40 questions), 20 that targeted overall sentence comprehension, and 20 that targeted comprehension of the filler material. Of the 80 total questions, 10 questions (6 targeting ORC clausal comprehension, 3 targeting SRC clausal comprehension, 1 filler) were excluded from further analyses. 3 were excluded for issues related to poor translation, and the remaining 7 were excluded for lower than chance (50%) correctness rate. For our analysis, we focused only on questions that targeted RC comprehension.

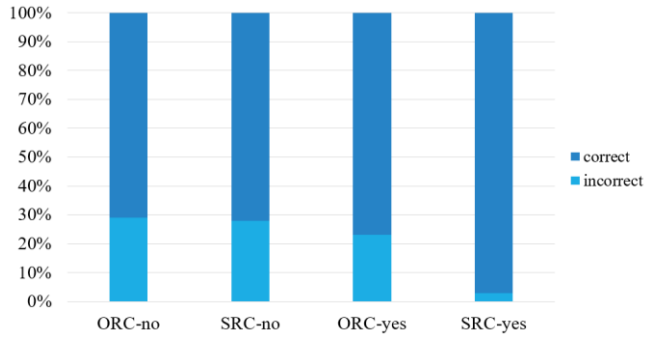

Figure 7: Proportion of correct answers by clause type and correct answer condition.

We first plotted the proportion of correct comprehension question answers by clause type and correct answer condition, shown in Figure 7. Whereas the "no" condition was comparable across clause type, there appeared to be a substantial discrepancy between clause type for the "yes" condition. Specifically, readers were more likely to respond "yes" when the correct answer was "no" in the ORC condition. This would entail that after reading an ORC stimulus (e.g., "The reporter who the senator attacked admitted the error."), if prompted with a possible SRC interpretation by the comprehension question (e.g., "Did the reporter attack the senator?"), a reader was more likely to accept that interpretation and answer "yes" when the correct answer was "no."

To evaluate the statistical reliability of these findings, a logistic mixed effects model was fit to the data with correctness as the dependent variable and clause type and correct answer as fixed effects. We also included an interaction between fixed effects and used the maximal random effects structure by subject and item.

Model estimates showed a reliable effect of both clause type and correct answer condition on correctness. First, comprehension questions for ORC stimuli were correlated with incorrect answers (β = -0.72; se: 0.18; CI: [-1.08, -0.39]). Additionally, questions whose correct answer was "no" were correlated with incorrect answers (β = -0.76; se: 0.20; CI: [-1.18, -0.41]). There was an additional reliable effect for the interaction between clause and question type (β = 0.67, se: 0.18, CI: [0.33, 1.04]), resulting in a subadditive effect for our predictor variables.

Discussion

We set out to determine whether SRCs or ORCs are easier to process in Standard Arabic. We found that SRCs are read faster in SA, supporting predictions from expectation-based theories of sentence processing. While these findings do not necessarily preclude memory-based processing costs, we find stronger support for expectations-based processing costs. Our models showed consistent effects for clause type across our main region of interest (the RC pronoun, RC verb, and RC noun), as well as individually at the RC pronoun and verb.

We then conducted an analysis of comprehension question data, which showed more incorrect answers for ORC sentences, as well as more incorrect answers for questions whose correct answer was "no" (i.e., a "yes" bias). There was an additional subadditive effect for clause and correct answer type, meaning that clause and correct answer type alone do not explain the trends in our data. Given these findings, we

hypothesize that some readers mistakenly interpret ORCs as SRCs while reading.

We believe there are two possible reasons for this misinterpretation. First, it is possible that the resumptive pronoun clitic is short enough that it is missed during reading. Our native speaker consultants observed that the object pronoun clitic is easy to miss in the ORC condition. On the other hand, it is also possible that readers see the object pronoun, yet reject an ORC interpretation in favor of a higher-frequency SRC interpretation. This hypothesis is in line with noisy-channel processing theories, which state that readers will compare the probability of a sentence's structure against the probability that noise corrupted the input (Levy, 2008b). Recent research in Hebrew relative clauses – a typologically similar language to Arabic – has demonstrated that readers prefer high-frequency, grammatically incorrect interpretations of sentences to their grammatical but infrequent counterparts, suggesting that expectations strongly modulate processing (Keshev & Meltzer-Asscher, 2021). Further, previous reading studies have found that re-reading does not improve comprehension accuracy (Christianson et al., 2017); thus, even when given the opportunity to re-read a RC verb with an object pronoun clitic, it is unlikely that readers will update their understanding of the sentence. These findings lead us to believe that readers are engaging in noisy-channel processing while reading ORCs in Arabic.

We have taken the first steps towards understanding these processing trends by conducting a recall task. Using the same stimuli, we asked participants (n = 80) to read each sentence and then reproduce the sentence word-for-word from memory. If readers are misreading ORCs by skipping the RP on the relative clause verb, we expect to see unidirectional errors of ORCs reproduced as SRCs; the skipping of the RP would simply result in an SRC reading and interpretation. However, if ORCs are misremembered as SRCs *and* SRCs are misremembered as ORCs, we believe that this lends stronger evidence to a noisy-channel processing framework; ORCs may be noisily interpreted as SRCs as they are the more frequent structure, but SRCs may also be noisily interpreted as ORCs if semantic expectations outweigh syntactic expectations. Overall, misremembrance rates for our recall task were low (<4%); however, we found that participants both misremembered ORCs as SRCs (71% of errors) and SRCs as ORCs. We thus interpret this as potential support for noisy-channel processing. These outcomes raise further questions about the strength of semantic, grammatical, and orthographic features in influencing a reader's willingness to accept a noisy interpretation. We plan to conduct an eye-tracking experiment next to ask how these features interact, particularly in the case of reading resumptive pronoun clitics.

The goal of the current study was to better understand both general sentence processing patterns, and processing patterns specific to Standard Arabic. Specifically, we tested whether SRCs or ORCs are easier to process in SA, and which processing theory best explained these patterns. Our results find support for expectation-based theories, but questions remain about the interaction of expectations and noisy-channel processing in sentence processing. We aim to explore these questions in future eye-tracking studies, where the temporal granularity of eye movements will allow us to better understand observed processing behaviors.

Acknowledgements

We would like to acknowledge the valuable native speaker assistance with stimuli design from Manar Al-Shatarat and Shayma Hassouna, and contributions to stimuli design and data processing by our undergraduate research assistants, Mariam Ali and Reem Suleiman.

References

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology, 4:328*.

Bürkner, P. (2018). Advanced Bayesian multilevel modeling with the R package brms. *R Journal, 10*(1), 395-411.

Carreiras, M., Duñabeitia, J., Vergara, M., de la Cruz-Pavía, I., & Laka, I. (2010). Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition, 115*(1), 79-92.

Christianson, K., Luke, S. G., Hussey, E. K., & Wochna, K. L. (2017). Why reread? Evidence from garden-path and local coherence structures. *Quarterly Journal of Experimental Psychology*, *70*(7), 1380–1405.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*(3), 348-368.

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

Gibson, E. (2000). The Dependency Locality Theory: A Distance-Based Theory of Linguistic Complexity. *Image, Language, Brain*, 95-126.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory Interference during Language Processing. *Journal of Experimental Psychology: Learning Memory and Cognition*, *27*(6), 1411–1423.

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

Hermena, E. (2016). Aspects of Word and Sentence Processing during Reading Arabic: Evidence from Eye Movements.

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition, 90*(1), 3-27.

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.

King, J., & Just, M. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*(5), 580-602.

Konieczny, L., & Döring, P. (2003). Anticipation of clause-final heads: Evidence from eye-tracking and SRNs. *Proceedings of the ICCS/ASCS-2003 Joint International Conference on Cognitive Science*, 1-6.

Kwon, N., Lee, Y., Gordon, P., Kluender, R., & Polinsky, M. (2010). Cognitive and Linguistic Factors Affecting Subject/Object Asymmetry: An Eye-Tracking Study of Prenominal Relative Clauses in Korean. *Language, 86*(3), 546-582.

Lau, E., & Tanaka, N. (2021). The subject advantage in relative clauses: A review. *Glossa: A Journal of General Linguistics*, *6*(1), 1–34.

Levy, R. (2008a). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126-1177.

Levy, R. (2008b). A noisy-channel model of rational human sentence comprehension under uncertain input. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 234–243).

Levy, R. (2011). Integrating surprisal and uncertain-input models in online sentence comprehension: formal techniques and empirical results. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 1055–1065.

Maamouri, M., Bies, A., Kulick, S., Krouna, S., Gaddeche, F., & Zaghouani, W. (2010). Arabic Treebank: Part 3 v 3.2 LDC2010T08. *Philadelphia: Linguistic Data Consortium*.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical Nature of Syntactic Ambiguity Resolution. *Psychological Review, 101*(4), 676–703.

Mak, W., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language, 47*(1), 50-68.

Meltzer-Asscher, A. (2021). Resumptive Pronouns in Language Comprehension and Production. *Annual Review of Linguistics, 7*, 177-194.

Schriefers, H., Friederici, A., & Kuhn, K. (1995). The Processing of Locally Ambiguous Relative Clauses in German. *Journal of Memory and Language, 34*, 499-520.

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition, 116*(1), 71-86.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, *47*(1), 69–90.

Ueno, M., & Garnsey, S. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes, 23*(5), 646-688.