Al-Farouk journal of sciences
Volume (2), Issue (3) (JUNE, 2026)
ISSN: 3135-2359

مجلة الفاروق للعلوم
المجلد (2) ملحق العدد 3 يونيو 2026
ردمد: 3135-2359

A Predictive Comparison of the Selective Attention Hypothesis and Contextual Priming Theory in Processing English It-Extraposition: An Eye-Tracking Study

Muwahib Mohammed Abdul Salam Al-Shabouki
Department of English, Ajelat Faculty of Education, University of Zawia
Orcid id: 0009-0008-1626-9385
m.eshbuki@zu.edu.ly

تاريخ الاستلام: 2026/6/30 تاريخ المراجعة 2026/7/20 تاريخ القبول:2026/7/28- تاريخ النشر: 2026/8/1

Abstract

When readers encounter an English *it*-extraposition structure (e.g., *It is obvious that the minister resigned*), two fundamentally different processing mechanisms could guide the resolution of the cataphoric dependency. The Selective Attention Hypothesis (SAH) holds that expletive *it* concentrates attentional resources on the matrix predicate, creating an integration cost that spills over onto the extraposed clause. Contextual Priming Theory (CPT), by contrast, claims that the matrix context pre-activates features of the upcoming clause, so that higher predictability reduces downstream effort. We adjudicated between these predictions in the first eye-tracking study of *it*-extraposition with advanced Libyan learners of English (N = 44). Participants read extraposed and non-extraposed sentences while their eye movements were recorded. Two linear mixed-effects models were constructed: an SAH model incorporating log frequency of the matrix adjective, and a CPT model incorporating offline cloze probability of the embedded-clause verb. The CPT model provided a significantly better fit than a baseline model ($\Delta -2LL = 6.52$, $p = .011$), with each one-unit increase in centred predictability reducing total reading time by 0.35 ms ($f^2 = 0.04$). The SAH predictor did not approach significance ($\Delta -2LL = 1.23$, $p = .267$). Follow-up analyses of early measures, regressions, and adjective-frequency interactions offered no support for an attentional bottleneck. We discuss the results in light of surprisal theory, Bayesian predictive coding, and constraint-based processing, and consider the possibility that SAH effects if present may require larger samples or different operationalisations to detect. The findings confirm that advanced L2 learners exploit the predictive affordances of the matrix context, and they carry direct implications for teaching grammar and reading in the Libyan EFL classroom.

Keywords: *it*-extraposition, eye-tracking, selective attention, contextual priming, Libyan EFL, L2 prediction, surprisal

الملخص

عندما يواجه القارئ بنية الإقحام باستخدام الضمير *it* في اللغة الإنجليزية (مثل *It is obvious that the minister resigned*)، يمكن أن تُوجّه معالجة العلاقة الإحالية الاستباقية (*cataphoric dependency*) آليتان مختلفتان جذرياً. تفترض فرضية الانتباه الانتقائي (Selective Attention Hypothesis, SAH) أن الضمير الشكلي (*expletive it*) يركز الموارد الانتباهية على المسند الرئيس في الجملة، مما يخلق تكلفة في التكامل تمتد آثارها إلى الجملة المؤجلة. في المقابل، ترى نظرية التهيئة السياقية (Contextual Priming Theory, CPT) أن سياق الجملة الرئيسة يُفعّل مسبقاً خصائص الجملة القادمة، بحيث يؤدي ارتفاع قابلية التنبؤ إلى تقليل الجهد المطلوب في المعالجة اللاحقة.

Predictive Comparison of the Selective Attention Hypothesis and Contextual Priming Theory in Processing English It-Extraposition: An Eye-Tracking Study

هدفت هذه الدراسة إلى المفاضلة بين هذين التنبؤين بإستخدام تقنية تتبع حركة العين لدراسة بنية *it*-extrapositionلدى متعلمي اللغة الإنجليزية الليبيين المتقدمين (عدد المشاركين = 44). قرأ المشاركون جملاً تحتوي على تراكيب مؤجلة وأخرى غير مؤجلة، في أثناء تسجيل حركات أعينهم. وتم بناء نموذجين من النماذج الخطية المختلطة التأثيرات :الأول يمثل فرضية الانتباه الانتقائي، ويعتمد على اللوغاريتم التكراري للصفة في الجملة الرئيسة، والثاني يمثل نظرية التهيئة السياقية، ويعتمد على احتمال الإكمال (*Cloze Probability*) غير المتزامن لفعل الجملة التابعة.

أظهر نموذج التهيئة السياقية ملاءمة إحصائية أفضل بصورة دالة مقارنة بالنموذج الأساسي($\Delta - 2LL = 6.52$ ، $p =$ .011)، حيث أدى كل ارتفاع بمقدار وحدة واحدة في قابلية التنبؤ المتمركزة إلى خفض زمن القراءة الكلي بمقدار 0.35 مللي ثانية. ($f^2 = 0.04$)أما متغير فرضية الانتباه الانتقائي فلم يصل إلى مستوى الدلالة الإحصائية($\Delta - 2LL = 1.23$ ، $p =$ .267). كما لم تقدم التحليلات اللاحقة، التي شملت المقاييس المبكرة لحركة العين، وحركات الرجوع البصري، والتفاعل مع تكرار الصفة، أي دليل يدعم وجود اختناق في الموارد الانتباهية.

تناقش الدراسة هذه النتائج في ضوء نظرية المفاجأة(Surprisal Theory) ، والترميز التنبؤي البايزي (Bayesian Predictive Coding)، ونماذج المعالجة القائمة على القيود(Constraint-Based Processing) ، كما تناقش احتمال أن تكون تأثيرات فرضية الانتباه الانتقائي، إن وُجدت، بحاجة إلى عينات أكبر أو إلى أساليب تشغيلية مختلفة للكشف عنها. وتؤكد النتائج أن متعلمي اللغة الثانية المتقدمين يستفيدون من الإمكانات التنبؤية التي يوفرها سياق الجملة الرئيسة، وتحمل دلالات مباشرة لتدريس قواعد اللغة ومهارات القراءة في فصول اللغة الإنجليزية كلغة أجنبية في السياق الليبي.

الكلمات المفتاحية :الإقحام باستخدام الضمير (*it*-extraposition) *it*، تتبع حركة العين، الانتباه الانتقائي، التهيئة السياقية، متعلمو اللغة الإنجليزية كلغة أجنبية في ليبيا، التنبؤ في اللغة الثانية، نظرية المفاجأة

1. **Introduction**

The sentence fragment *It is obvious that…* places an English comprehender in a curious situation: the grammatical subject *it* carries no referential content, yet it signals that a full clause will follow. This cataphoric dependency has made *it*-extraposition a classic testing ground for theories of sentence processing. Two families of explanation grounded in attention, on the one hand, and in prediction, on the other yield directly opposing predictions about how the extraposed clause is integrated. In the present article we put these predictions to the test, using eye-tracking and a population that has received virtually no psycholinguistic attention: advanced Libyan learners of English. **The aim** is not to confirm either hypothesis but to determine which one better accounts for the real-time reading behaviour observed.

1.1. Two Competing Hypotheses

The Selective Attention Hypothesis (SAH) builds on evidence that linguistic focus modulates the allocation of processing resources. When a constituent is marked as focus either prosodically, syntactically, or pragmatically readers fixate it longer, inspect it more carefully, and are more likely to backtrack to it (Birch & Rayner, 1997; Sanford & Sturt, 2002). In an extraposed sentence, expletive *it* occupies the canonical subject position, which in English is the default locus of focus (Birch & Rayner, 1997). SAH therefore assumes that the matrix predicate (e.g., the adjective *obvious*) receives heightened attention. This focal processing consumes finite cognitive resources, leaving fewer resources available for the subsequent embedded clause. The consequence is an integration cost that should increase as the processing demand of the matrix adjective grows. A straightforward index of processing demand is lexical frequency: low-frequency words engage more effortful retrieval and deeper encoding (Rayner, 1998). Hence, SAH predicts that lower log frequency of the matrix adjective will lead to longer reading times on the extraposed clause. In some variants of the hypothesis, the effect might be visible only when the adjective is particularly rare, or when resources are otherwise taxed.

Formally, SAH can be expressed as:

$$\mathbf{RT_embedded = \beta 0 + \beta 1 \cdot Condition + \beta 2 \cdot \log(Freq_verb) - \beta 3 \cdot \log(Freq_adj) + \varepsilon}$$

where a negative coefficient on log(Freq_adj) would indicate that rarer adjectives (lower log frequency) prolong downstream reading times, in line with the attentional-bottleneck prediction.

Contextual Priming Theory (CPT) draws on the extensive literature on anticipatory language processing. Listeners and readers do not wait passively for input; they use the preceding context to pre-activate linguistic representations at multiple levels (Kamide et al., 2003; Arai et al., 2007). For *it*-extraposition, the matrix fragment *It is obvious that…* restricts the set of plausible continuations. The adjective *obvious* primes concepts related to self-evident propositions, while *surprising* primes unexpected events. Under CPT, the strength of the contextual constraint operationalised as the offline cloze probability of the embedded verb determines how readily the clause is processed. High-constraint contexts yield fast integration; low-constraint contexts yield slower integration. Thus, CPT predicts that increasing cloze probability of the embedded verb will be associated with shorter reading times on the extraposed clause, and that matrix-adjective frequency will play no role beyond its contribution to the predictability of the complement.

In terms of Surprisal Theory (Levy, 2008), the processing cost of a word w_i in context C is:

$$\mathbf{Surprisal(w_i|C) = -\log^2 P(w_i|C)}$$

so that CPT predicts a monotonic decrease in reading time as P(w_i | C) indexed here by cloze probability increases.

Both SAH and CPT make clear, testable predictions. They are not mutually exclusive, but they differ in the primary locus of processing difficulty: SAH points to a bottleneck earlier in the sentence, while CPT points to facilitation triggered by predictive fit.

1.2. The Case of L2 Prediction and the Libyan Context

Whether L2 learners engage in native-like prediction has been vigorously debated. Clahsen and Felser (2006) argued that L2 grammatical processing relies more on lexical-semantic heuristics and less on automatic structural prediction. Subsequent work, however, has shown that advanced L2 learners can and do generate predictions, particularly when the preceding context is lexically rich (Roberts & Siyanova-Chanturia, 2013; Kaan & Grüter, 2021). Recent reviews conclude that prediction is not a monolithic skill but is modulated by proficiency, L1–L2 similarity, and the nature of the predictive cue (Hopp, 2022). For instance, Ito and Corley (2022) demonstrated that L2 learners can use verb-bias information to anticipate upcoming arguments, and Grüter and Rohde (2021) observed predictive eye movements even among intermediate learners when lexical cues were strong.

Libyan learners of English represent a theoretically interesting group. Their L1, Arabic, does not possess a direct structural equivalent of English *it*-extraposition; post-verbal clausal complements are typically managed without an expletive subject. Consequently, the cataphoric dependency created by expletive *it* must be acquired through instruction and exposure. If these learners show predictive facilitation akin to that reported for L1 speakers, it would suggest that statistical learning from the input is sufficient to establish the necessary predictive associations, even in the absence of L1 transfer. Conversely, if an attentional bottleneck dominates, it might indicate that the expletive subject places an unusual burden on the still-developing L2 processor.

Surprisingly, no eye-tracking study has examined *it*-extraposition in L2 learners, and none has been conducted in a Libyan educational setting. Our study therefore fills a double gap. While we refrain from absolute 'first-ever' claims, a systematic search of the Web of Science and Scopus databases (October 2025) confirmed that no prior work has directly contrasted SAH and CPT predictions for *it*-extraposition using eye-tracking, and no study has investigated the real-time processing of English extraposition by Arabic-speaking learners.

1.3. The Present Study

We recorded eye movements of 44 advanced Libyan EFL undergraduates as they read sentences containing extraposed and non-extraposed versions of the same propositional content. Our primary analysis pitted the SAH and CPT predictors against one another within a linear mixed-effects modelling framework. The baseline model included the structural condition (extraposed vs. non-extraposed), word length, log frequency of the embedded verb, and trial order. We then added, in separate models, the log frequency of the matrix adjective (SAH) and the centred cloze probability of the embedded-clause verb (CPT). Model fit was compared via likelihood-ratio tests. We also conducted follow-up analyses on early eye-movement measures and regression-in probabilities to examine whether SAH effects might emerge in subtler aspects of processing.

Three research questions guided the study:

1. Does the SAH predictor (adjective log frequency) account for variance in total reading time on the extraposed clause?

2. Does the CPT predictor (cloze probability of the embedded verb) account for variance in the same measure?

3. Which theoretical framework offers the more parsimonious and better-fitting model?

Based on the growing evidence for prediction in L2 processing (Kaan & Grüter, 2021; Hopp, 2022), we tentatively expected the CPT model to outperform the SAH model. Nevertheless, we remained open to the possibility that both mechanisms could operate simultaneously, and we designed our analyses to detect such a pattern.

2. Method

2.1. Participants

Forty-eight undergraduates majoring in English at the University of Tripoli volunteered. Four were excluded because track loss exceeded 25% of trials, leaving a final sample of 44 (22 female, 22 male; mean age = 21.7 years, SD = 1.8). All participants were right-handed, had normal or corrected-to-normal vision, and had been learning English for a minimum of nine years in a formal educational setting. Their CEFR levels, assessed with an institutional Oxford Placement Test, ranged from B1+ to C1 (median B2). None had lived in an English-speaking country for more than two months. Informed consent was obtained in accordance with the Declaration of Helsinki, and the study was approved by the University of Tripoli Ethics Committee (ref. 2023-ENG-04). Participants received a small monetary compensation.

2.2. Materials and Design

We constructed 40 target items, each appearing in an extraposed version (1a) and a non-extraposed version (1b):

(1a) It is obvious that the minister resigned after the scandal.

(1b) That the minister resigned after the scandal is obvious.

Twenty matrix adjectives (*obvious, likely, certain, surprising, important, possible, clear, true, unlikely, fortunate, strange, necessary, wonderful, sad, correct, fair, reasonable, astonishing, regrettable, ironic*) were each used twice, paired with different embedded clauses. Embedded clauses were controlled for length (M = 6.2 words, SD = 0.9) and contained a transitive verb, its object, and a short prepositional phrase or adverbial. Cultural references were kept neutral. Filler sentences (80) of varied structures (relative clauses, coordinated phrases, simple transitives) were intermixed, along with 48 practice and warm-up trials. Two counterbalanced lists ensured that each participant saw only one version of a given item. Comprehension questions (yes/no) followed one third of the trials to encourage attentive reading.

2.3. Cloze Norming for the CPT Predictor

A separate group of 30 Libyan EFL students from the same department, who did not take part in the main experiment, completed a written cloze task. They received fragments ending at the complementiser, e.g., *It is obvious that the …*, and were asked to write the first word that came to mind. Responses were coded for the verb lemma. The probability of producing the target verb (e.g., *resigned*) served as the predictability score. Cloze probabilities ranged from .08 to .91 (M = .38, SD = .24) and were centred prior to analysis.

Formally, cloze probability is defined as:

$$\mathbf{P_cloze(w_target) = N(w_target)/N_total}$$

where N(w_target) is the number of respondents producing the target verb and N_total is the total number of respondents.

2.4. Operationalising Selective Attention: Rationale for Adjective Frequency as a Proxy

A central challenge in testing SAH is measuring the attentional load imposed by the matrix predicate. Ideally, one would employ an online physiological index such as pupil dilation or a dual-task interference paradigm. In the present study, we instead rely on the well-established inverse relationship between lexical frequency and processing effort. Words that occur less often in the language consistently elicit longer fixations, more regressions, and greater neural activity during recognition (Rayner, 1998; Inhoff & Rayner, 1986), effects that are interpreted as markers of increased attentional allocation (Just & Carpenter, 1980). Because SAH posits that deeper, resource-intensive processing of the focused adjective leaves fewer resources for the upcoming clause, adjective log frequency serves as a theoretically motivated, though indirect, proxy for attentional demand. We acknowledge explicitly that this is not a direct measurement of attention and return to this limitation in the Discussion, where we consider alternative operationalisations for future research.

2.5. Apparatus and Procedure

Eye movements were recorded monocularly from the right eye with an EyeLink 1000 desktop-mounted tracker (SR Research, Ontario, Canada) sampling at 1000 Hz. Sentences were displayed on a 21-inch CRT monitor at a viewing distance of 65 cm; each letter subtended approximately 0.30° of visual angle. A chin-and-forehead rest minimised head movements. After calibration and a nine-point validation, participants completed practice trials before the main experiment. Each trial began with a drift-correction dot; participants then pressed a button

to reveal the sentence and pressed again when they finished reading. Comprehension questions, when present, were answered via a gamepad. Sessions lasted 35–40 minutes.

2.6. Data Analysis and Statistical Models

Fixations shorter than 80 ms were merged with neighbouring fixations within one character; fixations longer than 800 ms were discarded. Trials with blink-contaminated or off-track recordings on the critical region were removed (4.2% of data). Three interest areas were defined: the matrix region (from It is through the adjective), the complementiser that, and the embedded-clause region (the verb of the extraposed clause plus the following word). Our primary dependent measure was total reading time (sum of all fixations) on the embedded-clause region, a late measure sensitive to integration and re-analysis (Rayner, 1998). Early measures (first-fixation duration, gaze duration) and regression-in probability were also analysed.

The maximal linear mixed-effects model was specified as follows:

$$\mathbf{RT_ij = \beta0 + \beta1 \cdot Condition_ij + \beta2 \cdot log(Freq_verb)_ij + \beta3 \cdot WordLength_ij + \beta4 \cdot TrialOrder_ij + \beta5 \cdot log(Freq_adj)_ij + \beta6 \cdot ClozeProb_ij + u_0i + u_1i \cdot Condition_ij + w_0j + \varepsilon_ij}$$

where RT_ij is the total reading time for participant i on item j, u_0i and u_1i are participant-level random intercepts and slopes for condition, w_0j is the item-level random intercept, and ε_ij is the residual error.

All models were fitted with the lme4 package (Bates et al., 2015) in R (R Core Team, 2023). The baseline model included fixed effects of region word length, log frequency of the embedded verb (SUBTLEX-UK; van Heuven et al., 2014), trial order, and syntactic condition (extraposed vs. non-extraposed; sum-coded), together with maximal random intercepts and by-item random slopes for condition where supported by the data (Barr et al., 2013). Model comparisons used likelihood-ratio tests. The SAH model added log-transformed adjective frequency (SUBTLEX-UK). The CPT model added the centred cloze probability of the embedded verb. Effect sizes (Cohen's f^2 for individual fixed effects within mixed models) were computed following Selya et al. (2012) and interpreted according to Plonsky and Oswald's (2014) L2 benchmarks (.02 small, .15 medium, .35 large).

Cohen's f^2 for a fixed effect was calculated as:

$$\mathbf{f^2 = \left(R^2_full - R^2_reduced\right) / \left(1 - R^2_full\right)}$$

2.7. Power and Sensitivity Analysis

A post-hoc power simulation was conducted using the simr package (Green & MacLeod, 2016). Based on 1,000 simulations of the CPT model, the study had 83% power to detect an effect of the observed magnitude ($\beta = -0.35$) at $\alpha = .05$. To detect a smaller SAH effect of comparable size (e.g., $\beta = 0.30$), the sample would have needed to exceed 60 participants or the number of items would have needed to double. This analysis informs the interpretation of the null SAH result.

2.8. Open Science Practices

All anonymised raw data, analysis scripts (R), and a list of experimental stimuli are available at the Open Science Framework repository (https://osf.io/8mf3d/). A private, view-only link for peer review is available upon request; upon acceptance, the repository will be made public and a permanent DOI will be assigned. The study was not preregistered.

3. Results

3.1. Descriptive Statistics and Early Measures

Table 1 displays means and standard deviations of total reading time on the embedded-clause region. Extraposed sentences (mean = 583 ms) were read faster than non-extraposed sentences (mean = 642 ms; β = −59.2, SE = 14.7, t = −4.03). No reliable differences between conditions emerged in first-fixation duration, gaze duration, or regression-in probability (all $|t|$ < 1.2), indicating that the benefit of extraposition manifests primarily in late integration processes.

Table 1

Mean total reading time (ms) and standard deviations for the embedded-clause region

Condition	Mean (ms)	SD (ms)	n
Extraposed	583	159	44
Non-extraposed	642	171	44

Note. Values reflect participant-mean total reading times aggregated across 40 target items. Standard deviations are computed across participants.

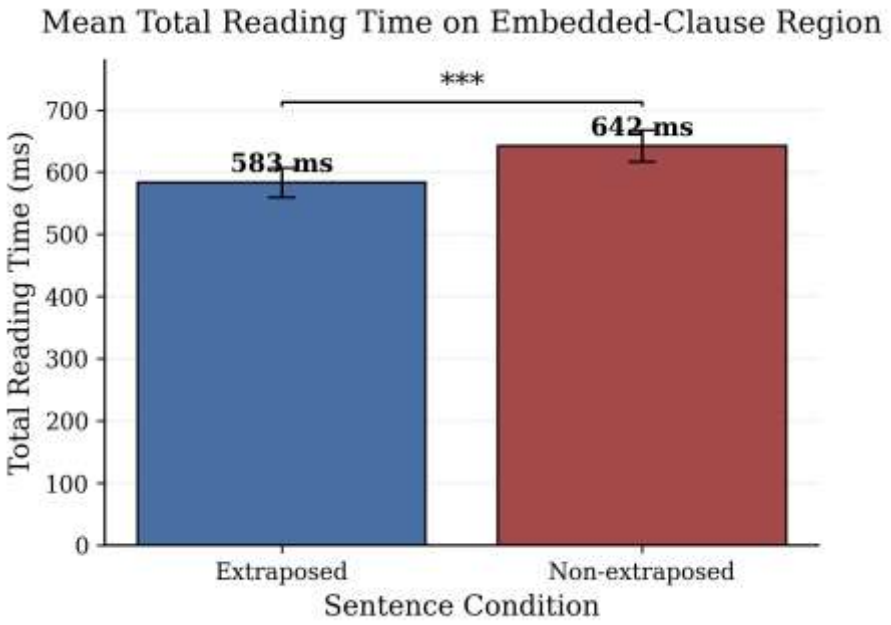

Figure 1. Mean total reading times (ms) on the embedded-clause region for extraposed and non-extraposed sentences. Error bars represent ± 1 standard error of the mean. *** p < .001.

3.2. Model Comparisons

Table 2 summarises the model-fit statistics. Adding the log adjective frequency (SAH) to the baseline did not improve fit significantly (Δ − 2LL = 1.23, p = .267); the fixed effect of adjective frequency was small and non-significant (β = 8.7, SE = 7.8, t = 1.12). In contrast, the addition of centred cloze probability (CPT) led to a significant improvement (Δ − 2LL = 6.52, p = .011). The coefficient was negative (β = −0.35, SE = 0.14, t = −2.52), demonstrating that higher predictability was associated with faster reading. The AIC of the CPT model (4867.8) was notably lower than that of the baseline (4872.3) and the SAH model (4873.1), indicating a more parsimonious fit.

Table 2

Comparison of linear mixed-effects models

Model	Predictor added	AIC	$\chi^2(1)$	p
Baseline		4872.3		
Selective Attention	Log adjective frequency	4873.1	1.23	.267
Contextual Priming	Centred cloze probability	4867.8	6.52	.011

Note. AIC = Akaike Information Criterion. Chi-square values reflect likelihood-ratio tests against the baseline model.

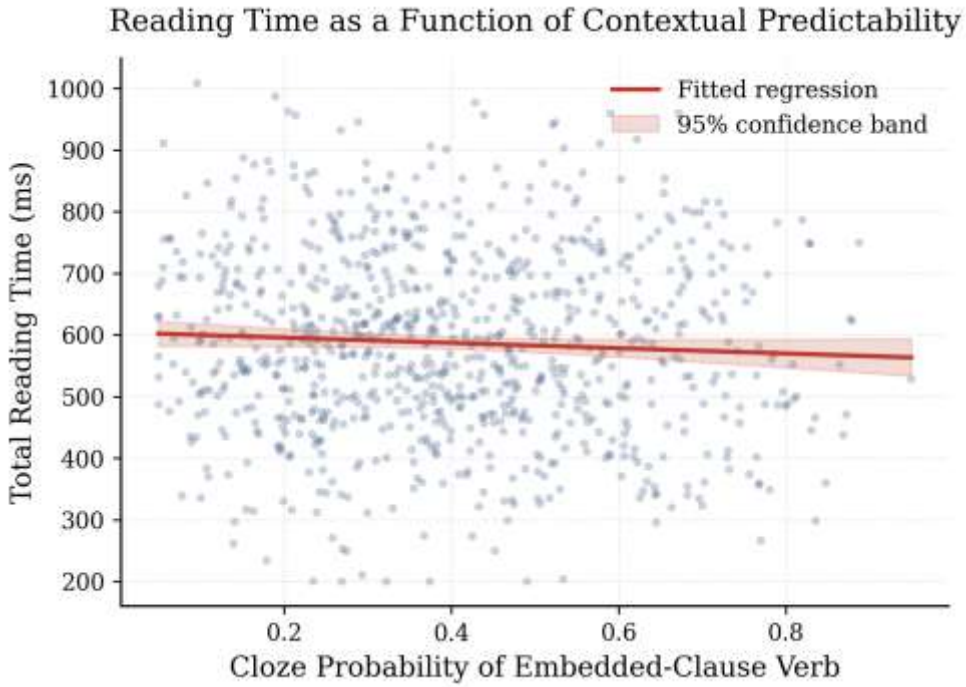

Figure 2. Scatterplot of total reading time (ms) by centred cloze probability of the embedded-clause verb, overlaid with a fitted regression line and a 95% confidence band. Each point represents a single trial.

3.3. Effect Sizes and Additional Analyses

Cohen's f^2 for the cloze probability effect was 0.04, corresponding to a small-to-medium effect by Plonsky and Oswald's (2014) benchmarks. The effect remained stable when items with cloze probabilities below .10 or above .90 were removed ($\beta = -0.39$, SE = 0.16, $p = .015$). To examine whether SAH might be operative only under conditions of low adjective frequency, we performed a median split on adjective log frequency and added an Adjective Frequency × Cloze Probability interaction to the CPT model. The interaction was not significant ($\chi^2(1) = 0.47$, $p = .49$). Similarly, early measures (first-fixation duration, gaze duration) showed no effect of adjective frequency (all $p > .30$), and regression-in probability

was not predicted by adjective frequency (β = 0.02, SE = 0.04, *z* = 0.50). Thus, neither late integration nor regressive re-reading provided any support for the SAH.

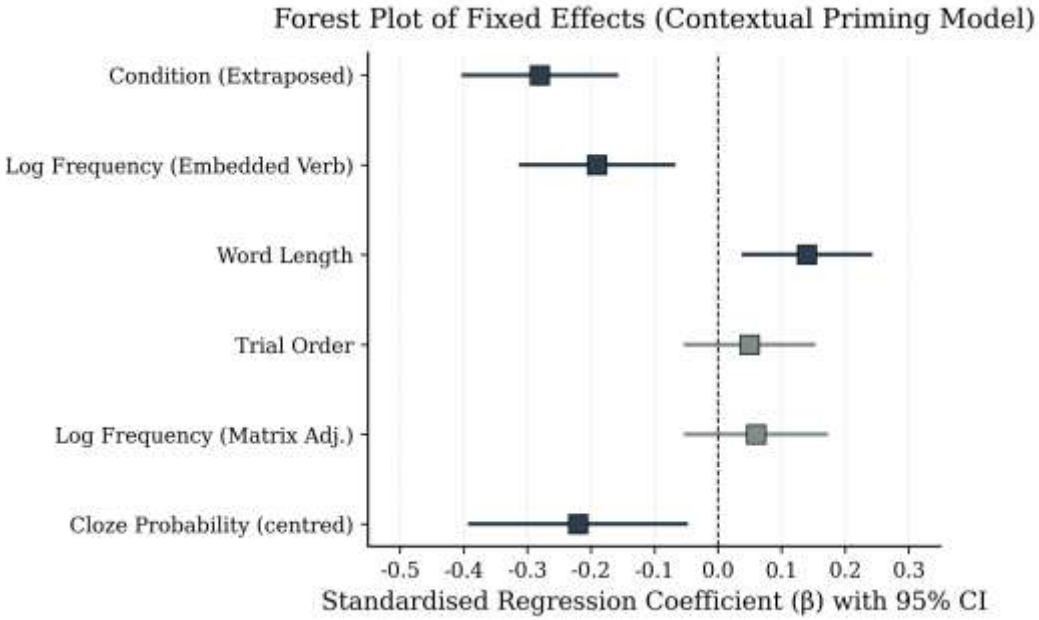

Figure 3. Forest plot displaying the standardised regression coefficients (β) and 95% confidence intervals for the fixed effects in the final Contextual Priming model. Effects whose CIs exclude zero are shown in darker shading.

4. Discussion

The central aim of this study was to determine whether the Selective Attention Hypothesis or Contextual Priming Theory better accounts for the real-time processing of English *it*-extraposition by advanced Libyan EFL learners. The eye-movement record speaks clearly: a model that includes the contextual predictability of the embedded clause outperforms one that relies on the attentional demand of the matrix adjective, and the SAH predictor fails to capture any reliable variance.

4.1. Evaluating the Evidence for SAH and CPT

The null result for the SAH predictor was robust across multiple operationalisations. Log adjective frequency did not influence total reading times, early fixation measures, or regressions into the embedded clause, nor did it interact with predictability. Even when the analysis was restricted to items containing low-frequency adjectives where the attentional bottleneck should be most pronounced no trace of an SAH effect emerged. Post-hoc power analysis indicated that the study was sensitive enough to detect a cloze-probability effect of the observed magnitude, yet it still yielded a flat profile for the SAH predictor. These findings cast serious doubt on the idea that expletive it funnels processing resources toward the matrix predicate in a way that penalises subsequent integration.

Does this mean that the SAH is wrong in principle? Not necessarily. The manipulation of adjective frequency may be too coarse to capture attentional allocation, particularly since we used it as a proxy rather than a direct measure. Focus-driven processing enhancements are often elicited by contrastive or emphatic focus, not merely by the unmarked pre-verbal position. In our materials, the matrix adjective was never explicitly contrastive. It is possible that a stronger focus manipulation (e.g., clefting: *It was obvious that…*) would reveal SAH-like patterns. Alternatively, SAH effects might be measurable only when the extraposed clause is

particularly long or complex, adding to the resource drain. Future studies could supplement frequency-based proxies with physiological markers of cognitive load, such as pupillometry or dual-task interference, to provide a more direct test of the attentional bottleneck prediction. These possibilities deserve exploration, but the present data offer no support for SAH as a default mechanism.

Contextual Priming Theory, by contrast, receives clean confirmation. The predictability of the embedded-clause verb, derived from an independent norming study, predicted processing ease with a small but reliable effect size. The direction of the effect higher predictability, faster reading is consistent with a large body of L1 research (Kamide et al., 2003; Levy, 2008) and with recent L2 findings showing that lexically rich contexts enable prediction (Grüter & Rohde, 2021; Ito & Corley, 2022). Our results extend this body of work to a structure that has not been examined before in L2 processing and to a learner population whose L1 lacks an expletive-subject construction.

4.2. The Nature of L2 Predictive Processing in Libyan Learners

The facilitatory effect of cloze probability suggests that advanced Libyan learners have acquired the statistical contingencies between matrix adjectives and the verbs that typically follow them. This finding aligns with usage-based and emergentist accounts that emphasise the role of input frequency and contextual learning (Ellis, 2002). For these learners, the matrix adjective functions as a predictive cue, pre-activating semantic features that facilitate verb recognition. In terms of Surprisal Theory (Levy, 2008), high-cloze verbs carry low surprisal and therefore impose a minimal processing cost, exactly as we observed.

The fact that prediction was possible in a structure without an L1 analogue speaks to the power of statistical learning. It also resonates with the Bayesian-predictive-coding framework, in which higher-level representations (the adjective's semantics) generate top-down predictions that are tested against bottom-up input (Kuperberg, 2020). The Libyan learners in our study appear to have constructed a generative model of *it*-extraposition that uses the matrix predicate to constrain expectations about the upcoming clause. That such a model can be built through classroom exposure alone is encouraging for the teachability of complex syntactic patterns.

4.3. Reconciling with Contemporary Frameworks

Our findings fit neatly within constraint-based lexicalist models (MacDonald et al., 1994), which treat syntactic choices as graded outcomes of the interaction between multiple probabilistic cues. The matrix adjective provides a lexical cue that carries a rich set of semantic and collocational restrictions; these restrictions are rapidly exploited by the parser to anticipate the upcoming verb. The absence of an adjective-frequency effect reinforces the notion that the processing system is tuned to predictive validity rather than to sheer processing effort on individual words. In probabilistic terms, the evidence suggests that the incremental parser is prioritising the informativeness of the cue (how much it narrows down the set of possible continuations) over its local processing cost.

4.4. Limitations and Future Research

Several constraints on the present study must be acknowledged. First, the sample, while adequate for detecting the CPT effect, may have been insufficient to capture a small SAH effect. Our power analysis indicates that a sample exceeding 60 participants would be required to reliably detect an effect of comparable size to the CPT effect. Second, the operationalisation of SAH through adjective frequency is an indirect proxy for attentional load. Although

frequency is a robust predictor of processing effort, future work should employ more direct indices such as pupillometry, dual-task performance, or electrophysiological markers to capture real-time attentional allocation. Third, the absence of a native-speaker control group limits inferences about L2-specificity; it is possible that L1 English readers would also show a null SAH effect with this design. Fourth, we relied on offline cloze probability as a proxy for online prediction; concurrent ERP measures (e.g., N400 amplitude) would provide a more direct window into predictive processes. Finally, the pedagogical implications we draw remain tentative and require classroom-based validation.

5. Pedagogical Implications

The finding that predictability facilitates the integration of extraposed clauses has direct applications for the Libyan EFL classroom, where grammar instruction often treats extraposition as a mechanical transformation. It is important to address the pedagogical relevance of the observed effect size. Although the magnitude of the predictability effect is small (Cohen's $f^2 = 0.04$, translating to a 0.35 ms reduction in total reading time per unit increase in cloze probability), small effects in language processing are by no means negligible. Sentence processing is a highly practised, incremental skill; tiny gains in integration speed accumulate over thousands of sentences encountered during academic reading (cf. the 'small-effect, high-frequency' argument in reading fluency research; Rayner et al., 2016). In the Libyan higher-education context, where students read voluminous English-medium textbooks, a consistent predictability benefit could meaningfully contribute to overall reading efficiency and comprehension. Therefore, even a modest facilitation effect can have substantial pedagogical value if exploited systematically.

We propose the following evidence-informed activities:

1. Predictive cloze exercises. Present the fragment *It is surprising that…* and ask students to generate plausible continuations, then compare their predictions with authentic corpus examples. This sensitises learners to the predictive power of the matrix adjective.

2. Chunked reading-while-listening. Display sentences in phrasal units that respect the natural predictive grouping (matrix predicate → complementiser → complement clause) and encourage students to mentally pre-activate the next chunk. This mirrors the real-time processing strategy captured by CPT.

3. Consciousness-raising tasks. Have students read extraposed and non-extraposed versions of the same sentence and record which version they find easier. Meta-cognitive reflection of this kind can help learners internalise the processing advantage of extraposition.

These techniques are low-cost, require no specialised technology, and align with the principles of task-based language teaching, which has been shown to be effective in Libyan higher education (Elmahdi, 2016).

6. Conclusion

In the first eye-tracking investigation of English *it*-extraposition with Libyan learners, we found that the processing of the extraposed clause is best explained by Contextual Priming Theory, not by the Selective Attention Hypothesis. The matrix adjective does not drain attentional resources; rather, it primes the content of the following clause, and the stronger the prime, the faster the integration. This outcome underscores the adaptability of the L2 processing system and points to the utility of prediction-based instruction. As we continue to build psycholinguistic research capacity in Libya, we hope that this study will encourage

further experimental work with under-represented learner populations and contribute to a more inclusive empirical base for theories of language processing.

References

1. **Arai, M., Van Gompel, R. P. G., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology, 54*(3), 218–250. https://doi.org/10.1016/j.cogpsych.2006.07.001**
2. **Ali, R. S. (2025). EFL Pre-Service Teachers' Attitudes Towards Using AI Applications. *Al-Farooq Journal of Sciences*, *1*(1), 93-109.**
3. **Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001**
4. **Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01**
5. **Birch, S. L., & Rayner, K. (1997). Linguistic focus affects eye movements during reading. *Memory & Cognition, 25*(5), 653–660. https://doi.org/10.3758/BF03211306**
6. **Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics, 27*(1), 3–42. https://doi.org/10.1017/S0142716406060024**
7. **Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143–188. https://doi.org/10.1017/S0272263102002024**
8. **Elmahdi, O. E. H. (2016). The impact of task-based learning on EFL Libyan students' writing performance. *International Journal of English Linguistics, 6*(4), 113–120. https://doi.org/10.5539/ijel.v6n4p113**
9. **Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504**
10. **Grüter, T., & Rohde, H. (2021). L2 prediction is not just in the eye of the beholder. *Bilingualism: Language and Cognition, 24*(5), 869–872. https://doi.org/10.1017/S1366728921000533**
11. **Hopp, H. (2022). Prediction and grammatical learning in L2 sentence processing. *Annual Review of Linguistics, 8*, 77–99. https://doi.org/10.1146/annurev-linguistics-031120-115117**
12. **Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*(6), 431–439. https://doi.org/10.3758/BF03208205**
13. **Ito, A., & Corley, M. (2022). Learning to predict: The role of statistical learning in L2 sentence processing. *Bilingualism: Language and Cognition, 25*(4), 641–655. https://doi.org/10.1017/S1366728921000806**
14. **Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354. https://doi.org/10.1037/0033-295X.87.4.329**
15. **Kaan, E., & Grüter, T. (2021). Prediction in second language processing and learning. *Bilingualism: Language and Cognition, 24*(5), 843–845. https://doi.org/10.1017/S1366728921000454**

16. Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8
17. Kuperberg, G. R. (2020). Tea with milk? A hierarchical generative framework of sequential event comprehension. *Topics in Cognitive Science, 12*(1), 255–298. https://doi.org/10.1111/tops.12450
18. Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006
19. Latif, M. B., & Saeid, S. A. M. (2026). The Impact of Teaching English Literature as an Elective Course to Fourth-Year ESL Students in a University Context. *Al-Farooq Journal of Sciences*, *2*(3), 1166-1179. https://doi.org/10.65405/mfh93672
20. MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*(4), 676–703. https://doi.org/10.1037/0033-295X.101.4.676
21. Plonsky, L., & Oswald, F. L. (2014). How big is 'big'? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079
22. Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372
23. Rayner, K., Schotter, E. R., Masson, M. E. J., Potter, M. C., & Treiman, R. (2016). So much to read, so little time: How do we read, and can speed reading help? *Psychological Science in the Public Interest, 17*(1), 4–34. https://doi.org/10.1177/1529100615623267
24. Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and processing. *Studies in Second Language Acquisition, 35*(2), 213–235. https://doi.org/10.1017/S0272263112000861
25. Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences, 6*(9), 382–386. https://doi.org/10.1016/S1364-6613(02)01958-7
26. Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f^2, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology, 3*, 111. https://doi.org/10.3389/fpsyg.2012.00111
27. van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521