BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

Cues to Subject-Object Grammatical Function Assignment in Arabic as a Second Language

Abdullah Ibrahim Alsubhi

Applied College, Taibah University, Madinah 42353, Saudi Arabia

ABSTRACT

This study presents cross-sectional data on the cues learners use to identify subject/object grammatical functions in Arabic as a second language. The roles of agreement, case, and plausibility were investigated, with predictions drawn from three theoretical models: the first noun principle, the unmarked alignment hypothesis, and the shallow structure hypothesis. Twenty-two learners from diverse first-language backgrounds participated in a sentence-picture matching task. Participants were recruited via email and provided with study vocabulary sentences for clarification. The primary task involved viewing 200 sentences, presented aurally and in writing alongside two images. Participants selected the image that correctly depicted the sentence's action, with each noun phrase serving as a human referent. Sentences were modified to target specific cues and presented randomly with correct responses to prevent order effects. Results revealed that beginning learners tended to interpret noun-verb-noun sequences as SVO, even in the presence of disambiguating cues. As learners advanced in proficiency, they became more sensitive to morphosyntactic information, correctly interpreting SVO and OVS structures. The findings supported the predictions of processability theory—particularly the unmarked alignment hypothesis—and the first noun principle, while contradicting the shallow structure hypothesis, as plausibility had no observable effect on performance. These results offer insight into the developmental sequence of cue integration in Arabic as a second language and contribute to ongoing debates in second language processing and syntactic development. Future research must expand to diverse structures and L1 backgrounds and use online methods to investigate real-time processing.

Keywords: Arabic as a Second Language; Grammatical Function Assignment; Morphosyntactic Cue; Processability Theory; Second Language Acquisition

*CORRESPONDING AUTHOR:

Abdullah Ibrahim Alsubhi, Applied College, Taibah University, Madinah 42353, Saudi Arabia; Email: aisubhi@taibahu.edu.sa

ARTICLE INFO

Received: 16 August 2025 | Revised: 5 September 2025 | Accepted: 19 September 2025 | Published Online: 4 November 2025
DOI: https://doi.org/10.30564/fls.v7i12.11668

CITATION

Alsubhi, A.I., 2025. Cues to Subject-Object Grammatical Function Assignment in Arabic as a Second Language. Forum for Linguistic Studies. 7(12): 184–195. DOI: https://doi.org/10.30564/fls.v7i12.11668

1. Introduction

1.1. Background

Understanding who is doing what to whom in a sentence is a fundamental aspect of language comprehension, hinging on the ability to assign grammatical functions. The drastic variation in how languages order major sentence constituents (e.g., subject, verb, object) and how grammatical functions are marked (e.g., syntax, morphology, semantics, pragmatics) makes it a challenging area for second language (L2) learners. In response, several frameworks have been created to explain the principles L2 learners rely on to identify major sentence constituents. Some focus on the initial stages of acquisition, some on inter-learner variation, and others on cross-linguistic differences between the first language (L1) and L2.

If these principles are indeed universal, one would expect them to manifest in L2 Arabic. However, few studies have applied them to Semitic languages. This study thus extends the literature by testing the claims of three such principles—the first noun principle[1], the shallow structure hypothesis[2], and lexical mapping theory[3] —in L2 Arabic, a language with relatively free word order and a rich agreement and case-marking system. The goal is to look into the cues (morphological, syntactic, or semantic-pragmatic) that L2 learners rely on to identify grammatical functions of canonical and non-canonical sentences at different proficiency levels. Arabic offers a unique testing ground, as different cues can be manipulated to appear in isolation from one another, allowing for independent analysis of their roles in grammatical function identification.

1.2. Research Questions and Hypotheses

The research questions and corresponding hypotheses are given below.

1. What cues do beginning L2 learners of Arabic rely on to assign grammatical functions?

In accordance with the first noun principle, beginning L2 learners of Arabic will primarily interpret the first noun in a noun-verb-noun sequence as the subject/agent of the sentence. Similarly, the unmarked alignment hypothesis predicts that beginning L2 learners will interpret noun-verb-noun as SVO, as they lack the necessary processing procedure to use case marking and agreement in identifying grammatical functions. Thus, they will show high accuracy rates with SVO sentences and low accuracy rates for OVS. On the other hand, the shallow structure hypothesis predicts that learners will only rely on heuristic word order and plausibility (e.g., non-syntactic cues). This entails that beginning L2 learners of Arabic will interpret plausible SVO and implausible OVS as SVO, while implausible SVO and plausible OVS will be interpreted as OVS.

2. Does the use of such cues vary across proficiency levels (beginner, intermediate, advanced)?

3. To what extent does plausibility contribute to grammatical function assignment in SVO and OVS structures across proficiency levels?

As learners advance, the effect of the first noun principle will be minimal, and accuracy in OVS will increase, indicating that the first noun principle is not in effect anymore. The unmarked alignment hypothesis assumes that as L2 learners start to acquire the necessary processing procedures (namely, inter-phrasal, their ability to exchange information across phrases, entailing their awareness of case marking and agreement), they will use case marking, agreement, or both to identify grammatical functions. On the other hand, despite proficiency level, the shallow structure hypothesis predicts that L2 learners of Arabic will continue to rely on heuristic word order and plausibility as non-syntactic cues to identify grammatical functions, entailing a continuing struggle with implausible SVO and implausible OVS even at intermediate and advanced levels.

4. Do different cues interact with each other and influence grammatical function assignment? (e.g., Is case and agreement more effective than either alone?)

Some sentences of this study were designed to show individual cues in isolation to examine the effect of each cue on learner interpretation. Therefore, L2 learners of Arabic will be most accurate with sentences that show more than one cue (e.g., agreement vs. case and case vs. case, agreement, and word order). To address these questions, word order, case marking, agreement, and plausibility cues were manipulated to determine their effect on L2 learners' assignment of grammatical functions.

2. Literature Review

VanPatten's study was one of the earliest studies to note an interlanguage pattern in which L2 learners misidentify

OVS and OSV sequences as SVO[4]. Subsequent investigation led to different plausible proposals to explain this pattern. Other scholars such as Houston, LoCoco, and VanPatten found the same pattern and proposed the first noun principle[1,5,6]. It is argued that, at early stages of language acquisition, L2 learners interpret the first encountered noun to be the subject/agent of the sentence. Support for this principle came from native speakers of English learning Spanish as an L2[7–10]. Tight found similar results when investigating ambitransitive verbs with animate subjects in SV and VS order, despite a persistent tendency to interpret the postverbal subject as an object[11].

The shallow structure hypothesis claims that L2 learners at an initial stage of acquisition tend to be more sensitive to semantic or pragmatic cues and less to morphosyntactic cues, as the representations built are less detailed or "shallower" than those of L1 learners[2,12]. Therefore, in languages such as Korean, Arabic, and Spanish, where morphosyntactic information (e.g., case marking and agreement) is necessary to accurately interpret scrambled sentences, beginning L2 learners rely on heuristic word order or plausibility.

Neither of the above principles offers claims about interlanguage development beyond the stage where difficulty is expected due to the proposed constraints. In contrast, processability theory explicitly predicts a universal developmental sequence for the productive side of grammar acquisition[13,14]. More recently, it has been extended to address the receptive side of grammar acquisition as well[11,15–17]. Given the limited number of studies extending processability theory in this way, however, it is too early to claim its principles as universal. The present study thus contributes to ongoing efforts to empirically test such claims within the framework of lexical mapping theory.

Processability theory utilizes lexical functional grammar, which assumes three levels of grammar: argument structure, function structure, and constituent structure[18]. These structures encode information required by the predicate. Argument structure encodes information about the number and type of arguments required (e.g., agent and patient), function structure encodes information about the grammatical function of constituents (e.g., subject, object), and constituent structure encodes information about the surface syntactic structure. Different types of sentences (e.g., active or passive declarative sentences and interrogative sentences) are formed by different mappings of these three structures. The basic assumption is that linear and nonlinear mapping come at no cost for native speakers, but nonlinear mapping increases processing loads for L2 learners[14].

Therefore, processability theory claims that L2 learners will initially organize their syntax to be the least costly in terms of processing, which does not require any rearrangement of linguistic elements and no language-specific procedural skills[14]. This is explained in the unmarked alignment hypothesis:

In second language acquisition, learners will initially organize syntax by mapping the most prominent semantic role available onto the subject (i.e., the most prominent grammatical role). The structural expression of the subject, in turn, will occupy the most prominent linear position in c-structure, namely the initial position[14].

What happens after this stage can be viewed as a gradual adjustment of the interlanguage to align with the linking principles of the target language. Non-linearity starts by fronting a non-argument into the initial position, leading to a distinction between subject and initial position as well as between subject and topic. The topic position is then eligible to host core arguments (e.g., object), which means that linear mapping as a constraint on the interlanguage is no longer in effect. This process is outlined in the topic hypothesis:

In second language acquisition, learners will initially not differentiate between SUBJ [subject] and TOP [topic]. The addition of an XP to a canonical string will trigger a differentiation of TOP and SUBJ, which first extends to non-arguments and successively to arguments, thus causing further structural consequences[14].

While processability theory originally formulated two separate hypotheses, integrated them into a unified proposal, outlining a developmental sequence for the acquisition of syntactic structures[14,19]. Since the two proposals do not differ in their claims, this study refers only to the original formulation for the sake of brevity.

3. Method

3.1. Target Structures

Modern Standard Arabic exhibits relatively free word order, with SVO, VSO, VOS, OVS, and OSV all possible. However, researchers disagree about which is the most unmarked. Some posit SVO as the canonical word order[9,20],

while others argue in favor of VSO [12,21–23]. Verbal morphology and nominal case markers in noun-verb-noun sentences indicate grammatical functions. Verbs must show full agreement with their subject, except when the subject is plural in VSO order. To avoid unneeded complexities, this study investigated sentences with singular masculine and singular feminine nouns in the subject and object positions.

3.2. Participants

Twenty-two L2 learners of Arabic were recruited to participate in this cross-sectional study. All met the following inclusion criteria: 1) onset of exposure to Arabic was at 18 years or older, 2) Arabic was their L2, and 3) they had been enrolled in Arabic classes for at least six months. At the time of participation, learners were either enrolled in a four-level preparatory program in the Arabic Language Institute at the Islamic University in Medina, Saudi Arabia, or taking regular classes at the same university. L1 and their proficiency level varied to address the research questions. Learners enrolled in Levels 1 and 2 were classified as beginners, those in Levels 3 and 4 as intermediate, and those in a university-level program as advanced. **Table 1** presents participant background information.

Table 1. Participant level and L1.

Participant ID	Level	L1
P001	Intermediate	Soninke
P002	Beginner	Swahili
P003	Beginner	Hausa
P004	Beginner	Afrikaans
P005	Advanced	Swahili, English
P006	Beginner	Hausa
P007	Beginner	Bambara
P008	Intermediate	Portuguese
P009	Beginner	Swahili
P010	Beginner	Bangala
P011	Advanced	Indonesian
P012	Beginner	Yoruba
P013	Beginner	Pashto
P014	Intermediate	French
P015	Advanced	French
P016	Advanced	Portuguese
P017	Intermediate	Somali
P018	Advanced	English
P019	Advanced	Indonesian
P020	Intermediate	French
P021	Advanced	French
P022	Advanced	English

3.3. Data Collection

Participants were recruited through an email sent by their teachers with the inclusion criteria and the researcher's contact information. Individuals who contacted the researcher and met the inclusion criteria made an appointment through the Calendly application to meet with the researcher in an office. During the meeting, the researcher explained the study, after which participants signed an informed consent form. Then, they filled out a questionnaire, giving their name, age, first language, onset of exposure to Arabic, and level in the Arabic Language Institute. Participants were given a list of sentences containing the words used in the study and had the opportunity to ask questions about any vocabulary they were unfamiliar with.

The main task involved picture matching. Two hundred sentences were presented one at a time on a monitor. Each sentence was simultaneously presented in written form and as an audio recording, accompanied by two pictures. One picture showed the subject of the presented sentence performing the action, while the other picture presented the object performing the action. All noun phrases used in the construction of the sentences represented human referents. The task was to identify the picture that matched

the presented sentence. The sentences were manipulated to target the cues in **Table 2**. To minimize order effects and predictability, the sentences and their correct answers were presented in random order.

Table 2. Task structure.

Structure	Cue	Plausibility	Number of Sentences
SVO	Agreement	+	15
SVO	Agreement	−	15
SVO	Case	+	15
SVO	Case	−	15
SVO	Case and agreement	+	15
OVS	Agreement	+	15
OVS	Agreement	−	15
OVS	Case	+	15
OVS	Case	−	15
OVS	Case and agreement	+	15
Fillers			50
Total			200

3.4. Data Analysis

Each response was coded as either correct (1) or incorrect (0), and the resulting data were entered into an Excel spreadsheet. This binary coding enabled the calculation of accuracy rates for each participant across each experimental condition and word order (SVO and OVS). Using the structured dataset, descriptive and inferential statistics were conducted in SPSS to examine how accuracy varied as a function of cue, word order, and proficiency level.

The analysis began with descriptive statistics to summarize accuracy across all conditions. Mean accuracy scores and standard deviations were calculated for each proficiency group (beginner, intermediate, advanced) and each combination of cue (agreement, case, case + agreement) and word order (SVO, OVS).

To test whether accuracy varied by sentence structure and learner group, a repeated-measures ANOVA was conducted with word order (SVO vs. OVS) as a within-subjects factor and proficiency level as a between-subjects factor. To examine the combined effects of cue and word order, and their interaction with proficiency, another repeated-measures ANOVA was conducted with six within-subjects conditions (three cue types × two word orders) and proficiency level as a between-subjects factor. The dependent variable in both models was the proportion of correct responses on the picture matching task. Where significant main effects or interactions were found, Bonferroni-adjusted pairwise comparisons of estimated marginal means were used to identify differences across cues and proficiency levels. Finally, to assess the contribution of semantic plausibility, paired-samples t-tests were run separately for each proficiency group and word order to determine whether accuracy differed significantly between plausible and implausible sentences.

4. Results

With 22 participants each completing 150 experimental and 50 filler trials, there were 4400 matching trials. The mean number of correct responses to the target (non-filler) trials, broken down by proficiency level and word order, is reported in **Table 3**.

Accuracy varied substantially as a function of word order and proficiency level. Accuracy was higher for SVO sentences, with a mean of 0.97 for advanced learners, 0.95 for beginners, and 0.89 for intermediate learners. In contrast, difficulty with OVS structures differed markedly across groups. Beginners performed near floor level, with a mean accuracy of just 0.07, while intermediate learners achieved 0.30, and advanced learners reached 0.85, closer to the levels achieved for SVO sentences. This suggested a developmental trend in the ability to interpret non-canonical word order, with higher proficiency associated with greater accuracy and increased reliance on morphosyntactic cues beyond surface word order.

A repeated-measures ANOVA with word order as a within-subjects factor and proficiency level as a between-subjects factor revealed a significant main effect of proficiency level, $F(2, 19) = 140.804$, $p < 0.001$, $\eta^2 = 0.937$,

indicating that accuracy varied significantly across proficiency levels. There was also a significant main effect from word order, $F(1, 21) = 781.562$, $p < 0.001$, $\eta^2 = 0.976$, indicating that sentence structure influenced performance. The interaction between word order and proficiency level was significant, $F(2, 19) = 188.737$, $p < 0.001$, $\eta^2 = 0.952$, suggesting the impact of word order on accuracy differed across learner levels. These results are presented in **Table 4**.

Table 3. Accuracy by proficiency level and word order (M ± SD).

Proficiency Level	SVO	OVS
Beginner	0.95 ± 0.02	0.07 ± 0.04
Intermediate	0.89 ± 0.06	0.30 ± 0.06
Advanced	0.97 ± 0.03	0.85 ± 0.12

Table 4. Repeated-measures ANOVA results for accuracy by proficiency level and word order.

Effect	df_1, df_2	F	p	η^2	Effect Type
Proficiency	2, 19	140.804	< 0.001	0.937	Between Subjects
Word order	1, 21	781.562	< 0.001	0.976	Within Subjects
Word order × proficiency	2, 19	188.737	< 0.001	0.952	Mixed

The researcher examined the effects of different cues (agreement, case, agreement + case, word order) on accuracy in interpreting grammatical functions and whether these effects varied across proficiency levels. The within-subjects factor was cue type per word order, which consisted of six conditions. The between-subjects factor was proficiency, with three levels: beginner, intermediate, and advanced. **Table 5** presents learners' accuracy (M ± SD) across cue types and word orders by proficiency level.

Accuracy was consistently high in SVO structures, regardless of cue. Beginners performed at or above 0.93 in all SVO conditions, and advanced learners were near ceiling (e.g., 0.98 ± 0.03 for agreement + case). Intermediate learners showed slightly lower performance in the agreement-only condition (0.84 ± 0.14) compared to case (0.93 ± 0.03), but the difference was minimal. In OVS structures, however, performance differed markedly by proficiency. Beginners showed extremely low accuracy across all cues (e.g., 0.10 ± 0.03 for agreement, 0.02 ± 0.05 for agreement + case), indicating difficulty in interpreting non-canonical word order even when cues were available. Intermediate learners improved modestly (e.g., 0.37 ± 0.09 in agreement + case), suggesting they were developing sensitivity to converging cues. Advanced learners performed significantly better in all OVS conditions, reaching 0.87 ± 0.12 when both cues were available, compared to 0.86 ± 0.14 for agreement and 0.83 ± 0.14 for case alone. Overall, while the type of cue had little impact on SVO sentence processing, it played a more prominent role in OVS sentences, where cue integration (agreement + case) was associated with the highest accuracy, especially among advanced learners.

To investigate which cue learners relied on most when interpreting non-canonical OVS sentences, a repeated-measures ANOVA was conducted with cue (agreement, case, agreement + case) as a within-subjects factor. The analysis revealed no significant main effect of cue type, $F(2, 38) = 2.400$, $p = 0.104$, partial $\eta^2 = 0.112$. Pairwise comparisons further confirmed that none of the cues differed significantly from each other (all $p > 0.14$), with overlapping confidence intervals across all comparisons. Although mean accuracy appeared slightly higher for sentences cued by agreement and case ($M = 0.429$) and agreement alone ($M = 0.425$) compared to case alone ($M = 0.384$), these differences were not significant. Overall, learners relied on all cues to a similar degree when processing OVS sentences, with no single cue providing a consistent advantage.

Given that no clear preference emerged for a particular morphosyntactic cue, the analysis turned to the role of plausibility as a potential factor guiding learners' interpretation of grammatical functions in SVO and OVS sentences. **Table 6** displays participants' accuracy scores (M ± SD) across plausible and implausible conditions in SVO and OVS word orders, broken down by proficiency level.

In SVO sentences, accuracy was uniformly high across all groups, with beginners performing at 0.934 ± 0.040 for plausible sentences and 0.963 ± 0.020 for implausible sen-

tences. Intermediate learners scored 0.926 ± 0.041 for plausible sentences and 0.842 ± 0.110 for implausible sentences, showing a slight decrease in accuracy with implausibility. Advanced learners maintained high accuracy across both SVO conditions, with 0.975 ± 0.029 for plausible and 0.967 ± 0.037 for implausible sentences.

Table 5. Accuracy by cue, word order, and proficiency level (M ± SD).

Word Order	Cue	Beginner	Intermediate	Advanced
SVO	Agreement	0.95 ± 0.04	0.84 ± 0.14	0.98 ± 0.04
	Case	0.95 ± 0.03	0.93 ± 0.03	0.96 ± 0.04
	Agreement and case	0.93 ± 0.06	0.92 ± 0.03	0.98 ± 0.03
OVS	Agreement	0.10 ± 0.03	0.31 ± 0.10	0.86 ± 0.14
	Case	0.06 ± 0.07	0.27 ± 0.08	0.83 ± 0.14
	Agreement and case	0.02 ± 0.05	0.37 ± 0.09	0.87 ± 0.12

Table 6. Accuracy by plausibility, word order, and proficiency level (M ± SD).

Word Order	Plausibility	Beginner	Intermediate	Advanced
SVO	+	0.934 ± 0.040	0.926 ± 0.041	0.975 ± 0.029
	−	0.963 ± 0.020	0.842 ± 0.110	0.967 ± 0.037
OVS	+	0.078 ± 0.047	0.242 ± 0.088	0.848 ± 0.143
	−	0.085 ± 0.041	0.333 ± 0.072	0.841 ± 0.128

In contrast, OVS sentences revealed large proficiency-related differences. Beginners struggled in both plausible (0.078 ± 0.047) and implausible (0.085 ± 0.041) OVS sentences. Intermediate learners showed moderate accuracy, with 0.242 ± 0.088 for plausible and 0.333 ± 0.072 for implausible sentences. Advanced learners demonstrated substantial accuracy, achieving 0.848 ± 0.143 for plausible and 0.841 ± 0.128 for implausible OVS sentences, indicating robust interpretation regardless of plausibility.

To examine whether plausibility contributed to sentence interpretation, accuracy between plausible and implausible sentences within each structure was compared. Paired-samples t-tests revealed no significant difference in the accuracy of SVO sentences between plausible and implausible items for beginners (mean difference $= -0.029$, $p = 0.104$), intermediate learners (mean difference $= 0.084$, $p = 0.140$), or advanced learners (mean difference $= 0.008$, $p = 0.512$), indicating that plausibility had no consistent effect. In OVS sentences, plausibility similarly did not yield significant differences for beginners (mean difference $= -0.007$, $p = 0.512$) or advanced learners (mean difference $= 0.007$, $p = 0.777$). However, plausibility significantly reduced accuracy among intermediate learners (mean difference $= -0.092$, $p = 0.035$), suggesting a potential interfering (rather than facilitative) effect.

5. Discussion

This study investigated cues that L2 Arabic learners used to identify grammatical functions in canonical SVO and non-canonical OVS word orders across proficiency levels. The cues investigated were syntactic and semantic, namely agreement, case, and plausibility. The findings revealed clear effects of proficiency level and word order on interpretation accuracy, with advanced learners demonstrating greater sensitivity to non-canonical structures.

This finding is consistent with the studies conducted by Alkodimi et al., Alsubhi, and Kim [24–26], who demonstrated that early-stage learners exhibit an avoidance processing strategy that maps the first noun onto an agentive role, irrespective of other available morphological or semantic cues.

Additionally, accuracy patterns suggested that certain cues supported more successful interpretations than others, while performance with OVS constructions was largely proficiency-dependent. These results align with Pienemann, Kampschulte et al., and Luchkina et al. [3,27,28], who found that canonical word orders are generally processed more efficiently, as they reflect learners' expectations about basic sentence structure. The fact that only advanced learners reached target-like accuracy with OVS constructions supports research highlighting the role of increased proficiency

in developing sensitivity to morphosyntactic cues beyond surface-level word order; this idea was validated by Zrari [29]. Beginners' low accuracy across non-canonical structures suggests a reliance on default processing strategies and limited access to morphosyntactic information; this finding is consistent with Hao et al. [30].

Furthermore, the significant interaction between proficiency and word order emphasizes the dynamic relationship between structural complexity and learner development [31]. This finding goes in the same vein as Birdsong's findings that L2 processing is modulated by both linguistic input and proficiency level, with learners gradually shifting from lexical-semantic reliance towards more syntactic parsing strategies as their proficiency increases.

With regard to morphosyntactic cues (case and agreement), intermediate and advanced learners showed higher accuracy when agreement and case were both used in OVS sentences than when only case or agreement appeared in isolation in such sentences [32,33]. With regard to plausibility, accuracy was slightly higher for plausible SVO sentences among advanced and intermediate learners, while for plausible OVS sentences, only advanced learners showed a slight increase in mean accuracy. However, this effect was not significant. Below, the discussion turns to the hypotheses derived from various theoretical models, evaluating their validity in this context.

The first research question focused on the nature of the cues beginner L2 learners rely on to identify grammatical functions. In this respect, the three models had competing predictions. The first noun principle and unmarked alignment hypothesis would expect beginning learners to rely heavily on interpreting noun-verb-noun sentences as SVO despite the availability of different morphosyntactic and semantic cues. In contrast, the shallow structure hypothesis would expect reliance on heuristic word order and plausibility (non-syntactic cues). This finding was supported by Aoun et al. [21] and Henry et al. [34], who found that L2 learners are more successful in predicting thematic roles when both case and semantic cues are present, but less so when syntactic cues are isolated; Lee and Witzel demonstrated that plausibility significantly guides interpretation in ambiguous structures for non-native speakers [35]; Wang et al. observed that beginner learners lean more on semantic cues and less on syntactic structure compared to intermediate learners [36]. Thus, if plausibility guided interpretations, beginners should interpret plausible SVO and implausible OVS as SVO, while implausible SVO and plausible OVS should be interpreted as OVS.

The data supported the predictions of the first noun principle and the unmarked alignment hypothesis entailed by processability theory. Beginners consistently interpreted noun-verb-noun sentences as SVO, even when OVS sentences were marked with different morphosyntactic and semantic cues [37]. This finding contradicted the predictions of the shallow structure hypothesis, as beginning learners interpreted all sentence types as SVO, including plausible OVS and implausible SVO constructions. This pattern aligned with studies showing that beginners often default to canonical word order interpretations, even in the presence of disambiguating cues, and early learners prioritize word order over morphological or semantic information during sentence processing [1,4,6,7,9,38].

The second research question considered whether the use of such cues varied across proficiency levels, which could help identify developmental patterns in how learners interpreted canonical and non-canonical structures. It was anticipated that as learners advanced, the influence of the first noun principle would diminish, and their ability to identify grammatical functions would increasingly rely on other sources of information, such as agreement and case markers, as they acquired the necessary processing procedures entailed by the unmarked alignment hypothesis [39]. The shallow structure hypothesis predicted a persistent reliance on heuristic cues such as word order and plausibility, even as learners advanced, implying that native-like performance may not be fully attainable.

The results confirmed the predictions of the first noun principle and unmarked alignment hypothesis, as learners demonstrated a developmental shift from initially interpreting noun-verb-noun sentences as SVO to relying on morphosyntactic cues (agreement, case, and agreement and case). This shift enabled them to identify grammatical functions in both SVO and OVS sentences, as evident by the increase in accuracy in OVS sentences across proficiency levels. The analysis revealed a significant effect from proficiency level and word order, as well as a significant interaction between the two variables.

The shallow structure hypothesis claimed that L2 learn-

ers would continue to rely on heuristic word order and plausibility. However, this was contradicted by the data, as almost all advanced learners correctly identified grammatical functions in both plausible OVS and implausible OVS sentences. This finding led to the third research question regarding whether plausibility played a role in grammatical function identification. If plausibility were a guiding factor, one would expect L2 learners to show higher accuracy in plausible sentences than in implausible sentences. The strong performance of advanced learners on both plausible and implausible OVS sentences demonstrates their ability to utilize syntactic cues over heuristic strategies, directly contradicting the shallow structure hypothesis.

Furthermore, the absence of a significant accuracy advantage for plausible sentences across all proficiency levels indicates that semantic plausibility was not a primary or reliable strategy for grammatical function identification in this learner group. In some instances, mean accuracy scores were actually higher for implausible structures, for example, SVO sentences among beginners and OVS sentences among both beginner and intermediate learners. Moreover, in cases where accuracy was higher for plausible sentences, the differences were not significant.

The fourth research question explored whether different cues would interact with one another in ways that influenced grammatical function assignment in L2 Arabic. This question was relevant for understanding how L2 learners processed multiple sources of information simultaneously and whether cue convergence could facilitate more accurate interpretations. It was hypothesized that accuracy would be highest when sentences contained both agreement and case markers, compared to when only one cue was present. Regardless of the cue used in the sentence, differences in mean accuracy were not significant and therefore did not support the study's predictions. This suggested that learners relied on all cues equally, without showing a clear preference for one over the others. It was apparent that learners did not benefit from the facilitative effect of cue redundancy because there was no discernible difference in accuracy between sentences with a single cue and those with convergent cues.

Of the three theoretical models tested in this study, the unmarked alignment hypothesis was the most successful in accounting for early learner patterns and the progression observed across proficiency levels. While the first noun principle accurately captured learners' initial reliance on canonical word order, it offered little specification about subsequent stages of development. The predictions of the shallow structure hypothesis, on the other hand, were not supported by the performance of any learner group, as even advanced learners successfully interpreted non-canonical sentences using morphosyntactic cues.

In contrast, the unmarked alignment hypothesis offers more detailed predictions regarding the early stages of L2 development. It posits that learners initially rely on default mapping strategies and remain constrained by this reliance until Stage 4 of the processing hierarchy is reached, namely the inter-phrasal processing procedure. At this stage, information exchange between phrases becomes operable, allowing for the emergence of verbal agreement features and the functional use of nominal case marking, as seen in German [23,40,41], Russian [22,42], Serbian [19], and Hindi [43]. This model does not assume a fixed order of emergence for agreement and case, a position supported by the present study's findings; learners showed successful use of both cues without exhibiting a consistent preference for one over the other.

6. Study Implication

The findings of the study offer substantial pedagogical implications for teaching Arabic as a second language (L2), specifically concerning grammar, word order, and case marking. The research indicates a developmental trajectory where novice learners rely on canonical SVO word order, while advanced learners effectively incorporate morphosyntactic cues. Consequently, teaching methodologies should be customized to individual student proficiency.

For novice students, instruction should focus on reinforcing SVO processing and developing automaticity with frequent patterns through comprehensible input and production practice. Introducing complex structures like OVS or case endings prematurely is inadvisable due to novice learners' limited processing capacity to utilize them. Overburdening beginners with unabsorbable information is unproductive and potentially discouraging.

To facilitate the transition to intermediate Arabic learners from word order dependency to morphological sensitivity, educators should explicitly teach case and agreement markers in isolation, presenting the ʾIʿrāb (desinential inflection)

as an instrument for ambiguity rather than strict rules. Activities should isolate single cues, employing methods. For advanced learners, the focus should shift to integrating multiple cues and achieving automaticity in parsing complex sentences through authentic texts with altered word order and metalinguistic discussions on cue convergence.

7. Future Studies

Future research should utilize online processing methods, such as eye-tracking and self-paced reading, to examine real-time interactions in cue integration and processing ease, rather than relying solely on offline accuracy judgments. The research scope should be broadened to include a diverse array of Arabic linguistic structures, including relative clauses, VSO sentences, and passivized forms, to achieve a thorough understanding of L2 Arabic syntactic development. Comparative studies involving learners with varied L1 backgrounds, specifically those with and without case-marking systems, are essential for confirming the Unmarked Alignment Hypothesis and identifying L1 transfer influences. Finally, pedagogical implications require empirical validation through longitudinal or training studies that implement and test teaching interventions targeting morphological cues and non-canonical word orders to evaluate their effect on parsing strategies.

8. Conclusions

This study examined how L2 learners of Arabic assigned subject and object grammatical functions, focusing on the roles of the type of cue, word order, and proficiency level. Drawing on predictions from three major theoretical models (the first noun principle, shallow structure hypothesis, and unmarked alignment hypothesis), the study investigated whether learners relied more heavily on certain linguistic cues, how this reliance changed with proficiency, and whether multiple cues interacted during sentence interpretation. Beginners interpreted noun-verb-noun sentences as SVO, even when disambiguating cues were present. As learners advanced, they developed processing procedures that enabled them to use additional sources of information (e.g., morphosyntactic cues) for assigning grammatical functions. Once these procedures were in place, learners no longer showed a clear preference for one cue over another. These results supported the plausibility of processability theory, as the unmarked alignment hypothesis accounted for both the initial and developmental stages of grammatical function assignment in L2 Arabic. To confirm and elaborate on these findings, a longitudinal study tracking learners over time would be ideal, as it would allow for a more direct observation of developmental changes in cue integration and grammatical function assignment.

Funding

This research received no external funding.

Institutional Review Board Statement

This study was approved by the Institutional Review Board at Taibah University, Medina, Saudi Arabia and the Approval number is 1447183878.

Informed Consent Statement

All participants gave informed consent before participation.

Data Availability Statement

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request (aisubhi@taibahu.edu.sa).

Conflicts of Interest

The author declares no conflict of interest.

References

[1] VanPatten, B., 2004. Input Processing in Second Language Acquisition. In: VanPatten, B. (ed.). Processing Instruction: Theory, Research, and Commentary. Lawrence Erlbaum: Mahwah, NJ, USA. pp. 5–31.

[2] Clahsen, H., Felser, C., 2006. Grammatical Processing in Language Learners. Applied Psycholinguistics. 27, 3–42.

[3] Pienemann, M., Di Biase, B., Kawaguchi, S., 2005. Extending Processability Theory. In: Pienemann, M. (Ed.). Cross-Linguistic Aspects of Processability Theory. John Benjamins: Amsterdam, Netherlands. pp.

199–251.

[4] VanPatten, B., 1984. Learners' Comprehension of Clitic Pronouns: More Evidence for a Word Order Strategy. Hispanic Linguistics. 1, 57–67.

[5] Houston, T., 1997. Sentence Processing in Spanish as a Second Language: A Study of Word Order and Background Knowledge. In: Glass, W.R., Pérez-Leroux, A.T. (eds.). Contemporary Perspectives on the Acquisition of Spanish, Vol. 2: Production, Processing and Comprehension. Cascadilla Press: Somerville, MA, USA. pp. 123–134.

[6] LoCoco, V., 1987. Learner Comprehension of Oral and Written Sentences in German and Spanish: The Importance of Word Order. In: VanPatten, B., Dvorak, T.R., Lee, J.F. (Eds.). Foreign Language Learning: A Research Perspective. Newbury House: Boston, MA, USA. pp. 119–129.

[7] Isabelli, C.A., 2008. First Noun Principle or L1 Transfer Principle in SLA? Hispania. 91, 465–478.

[8] Lee, J.F., Malovrh, P.A., 2009. Linguistic and Non-Linguistic Factors Affecting OVS Processing of Accusative and Dative Case Pronouns by Advanced L2 Learners of Spanish. In: Collentine, J., García, M., Lafford, B., et al. (Eds.). Selected Proceedings of the 11th Hispanic Linguistics Symposium. Cascadilla Proceedings Project: Somerville, MA, USA. pp. 105–116.

[9] Mohammad, M., 1990. The Problem of Subject-Verb Agreement in Arabic: Towards a Solution. In: Eid, M. (ed.). Perspectives on Arabic Linguistics I: Papers from the First Annual Symposium on Arabic Linguistics. John Benjamins: Amsterdam, Netherlands. pp. 95–126.

[10] Soltan, U., 2007. On Formal Feature Licensing in Minimalism: Aspects of Standard Arabic Morphosyntax [PhD thesis]. University of Maryland: College Park, MD, USA. pp. 1–241.

[11] Tight, D.G., 2012. The First Noun Principle and Non-Canonical Structures. Studies in Hispanic and Lusophone Linguistics. 5(2), 351–368.

[12] Spinner, P., 2013. Language Production and Reception: A Processability Theory Study. Language Learning. 63, 704–739.

[13] Pienemann, M., 1998. Language Processing and Second Language Development: Processability Theory. John Benjamins: Amsterdam, Netherlands.

[14] Pienemann, M., 2005. An Introduction to PT. In: Pienemann, M. (Ed.). Cross-Linguistic Aspects of Processability Theory. John Benjamins: Amsterdam, Netherlands. pp. 1–60.

[15] Buyl, A., Housen, A., 2015. Developmental Stages in Receptive Grammar Acquisition: A Processability Theory Account. Second Language Research. 31(4), 523–550.

[16] Keatinge, D., Keßler, J.U., 2009. The Acquisition of the Passive Voice in L2 English: Perception and Production. In: Keatinge, D., Keßler, J.U. (Eds.). Research in Second Language Acquisition: Empirical Evidence Across Languages. Cambridge Scholars: Cambridge, UK. pp. 67–92.

[17] Larsen-Freeman, D., 2002. Making Sense of Frequency. Studies in Second Language Acquisition. 24(2), 275–285.

[18] Bresnan, J., 2001. Lexical Functional Syntax. Blackwell Publishers: Oxford, UK. pp. 1–460.

[19] Di Biase, B., Bettoni, C., Medojević, L., 2015. The Development of Case: A Study of Serbian in Contact with Australian English. In: Bettoni, C., Di Biase, B. (eds.). Grammatical Development in Second Languages: Exploring the Boundaries of Processability Theory. pp. 195–212.

[20] Awwad, M.A.M.B., 1973. Relativization and Related Matters in Classical, Modern Standard, and Palestinian Colloquial Arabic [PhD thesis]. Brown University: Providence, RI, USA. pp. 1–231.

[21] Aoun, J., Benmamoun, E., Choueiri, L., 2010. The Syntax of Arabic. Cambridge University Press: Cambridge, UK.

[22] Artoni, D., Magnani, M., 2015. Acquiring Case Marking in Russian as a Second Language: An Exploratory Study on Subject and Object. In: Bettoni, C., Di Biase, B. (Eds.). Grammatical Development in Second Languages: Exploring the Boundaries of Processability Theory. pp. 177–193.

[23] Bakir, M.J., 1980. Aspects of Clause Structure in Arabic: A Study in Word Order Variation in Literary Arabic [PhD thesis]. Indiana University: Bloomington, IN, USA. pp. 1–190.

[24] Alkodimi, K.A., Al-Ahdal, A.A.M.H., 2021. Strategies of Teaching Writing at Saudi Tertiary-Level Institutions: Reality and Expectations. Arab World English Journal. 12(2), 399–413.

[25] Alsubhi, A., 2021. The Development of Case Morphology and Sentential Word Order in Arabic as a Second Language: A Processability Perspective [PhD thesis]. The University of Wisconsin-Milwaukee: Milwaukee, WI, USA.

[26] Kim, H., 2025. Roles of L1 and Semantic Features in L2 Morpheme Acquisition: The Development of Morphology in L2 English for Korean Learners. Second Language Research. DOI: https://doi.org/10.1177/02676583251327099

[27] Kampschulte, Z., Braun, A., Zahner-Ritter, K., 2025. Intonational Focus Marking by Syrian Arabic Learners of German: On the Role of Cross-Linguistic Influence and Proficiency. Languages. 10(7), 155. DOI: https://doi.org/10.3390/languages10070155

[28] Luchkina, T., Ionin, T., Goldshtein, M., 2024. Acquisition of Non-Contrastive Focus in Russian by Adult English-Dominant Bilinguals. Frontiers in Psychology. 15.

[29] Zrari, A., 2025. The Impact of Moroccan Arabic on the Acquisition of French C'est-Cleft Constructions. International Journal of Linguistics, Literature and Translation. 8(8), 76–99.

[30] Hao, J., Chondrogianni, V., Sturt, P., 2024. Sources of Children's Difficulties with Non-Canonical Sentence Structures: Insights from Mandarin. Journal of Child Language. 1–28. DOI: https://doi.org/10.1017/S0305000924000424

[31] Birdsong, D., 2006. Dominance, Proficiency, and Second Language Grammatical Processing. Applied Psycholinguistics. 27(1), 46–49.

[32] Albelihi, H.H.M., Al-Ahdal, A.A.M.H., 2024. Overcoming Error Fossilization in Academic Writing: Strategies for Saudi EFL Learners to Move Beyond the Plateau. Asian-Pacific Journal of Second and Foreign Language Education. 9(1), 75.

[33] Alemán Bañón, J., Miller, D., Rothman, J., 2017. Morphological Variability in Second Language Learners: An Examination of Electrophysiological and Production Data. Journal of Experimental Psychology: Learning, Memory, and Cognition. 43(10), 1509.

[34] Henry, N., Jackson, C.N., Hopp, H., 2022. Cue Coalitions and Additivity in Predictive Processing: The Interaction Between Case and Prosody in L2 German. Second Language Research. 38(3), 397–422. DOI: https://doi.org/10.1177/0267658320963151

[35] Lee, J., Witzel, J., 2023. Plausibility and Structural Reanalysis in L1 and L2 Sentence Comprehension. Quarterly Journal of Experimental Psychology. 76(2), 319–337. DOI: https://doi.org/10.1177/17470218221092400

[36] Wang, L.F., Xie, Y., Yu, K., et al., 2021. Prosody and Sentence Processing Through Pause Insertion. International Journal of Bilingualism. 25(5), 1473–1485. DOI: https://doi.org/10.1177/13670069211018753

[37] Al-Thawahrih, J.N., 2016. Arabic Second Language Learners' Use of Word Order and Subject-Verb Agreement for Actor Role Assignment [Master's thesis]. The University of Utah: Salt Lake City, UT, USA.

[38] Kim, M.H., 2019. Processing of Canonical and Scrambled Word Orders in Native and Non-Native Korean [PhD thesis]. University of Illinois at Urbana-Champaign: Urbana, IL, USA.

[39] VanPatten, B., Collopy, E., Price, J.E., et al., 2013. Explicit Information, Grammatical Sensitivity, and the First-Noun Principle: A Cross-Linguistic Study in Processing Instruction. The Modern Language Journal. 97(2), 506–527.

[40] Baten, K., 2011. Processability Theory and German Case Acquisition. Language Learning. 61(2), 455–505.

[41] Baten, K., 2013. The Acquisition of the German Case System by Foreign Language Learners. John Benjamins: Amsterdam, Netherlands.

[42] Artoni, D., Magnani, M., Butt, M., et al., 2013. LFG Contributions in Second Language Acquisition Research: The Development of Case in Russian L2. In: Butt, M., Holloway King, T. (Eds.). Proceedings of the LFG13 Conference. CSLI Publications: Stanford, CA, USA. pp. 69–89.

[43] Baten, K., Verbeke, S., 2015. The Acquisition of the Ergative Case in Hindi as a Foreign Language. In: Baten, K., Buyl, A., Lochtman, K., et al. (Eds.). Theoretical and Methodological Developments in Processability Theory. John Benjamins: Amsterdam, Netherlands. pp. 71–104.