Lexical Aspect in the Development of the English Progressive Among Arab EFL Speakers

Alaa AlZahrani[1] and Normala Othman[2]
King Saud University, Saudi Arabia

Abstract
Recent research calls into question the predictive power of the Aspect Hypothesis as it has been found that it is advanced learners rather than beginners who show the developmental pattern predicted by the hypothesis. Following the identification of several design issues including lack of rigorous treatment of proficiency level and reliance on controlled elicitation tasks, this cross-sectional study used an Arab learner corpus (BALC) that controls proficiency level and task. This corpus-based study analysed the development of tense-aspect among Arabic-speaking EFL learners at two levels, beginner and advanced, by comparing their performance on 200 essays. Contrary to the recent research on Aspect Hypothesis, it was found that lexical aspect influences the early acquisition of the English progressive as the lower proficiency level in the present study extensively attached the progressive on activity verbs.

Keywords
Lexical aspect, Aspect Hypothesis, English progressive, verb semantics, tense-aspect, Arab EFL speakers

1. Introduction

The development of tense-aspect morphology in Second Language Acquisition (SLA) has been a central topic in SLA research (see Bardovi-Harlig; Salaberry; Wulff et al.; Alruwaili; Al-Thubaiti; Wang and Shirai; Hackmann; Mueller, for example). One influential line of research in the Aspect Hypothesis holds that lexical aspect, i.e., verb semantics, affects the acquisition of tense-aspect morphology by Second Language (L2) learners (Andersen and Shirai). One key prediction of the Aspect Hypothesis is that the progressive *-ing* will be initially used with verbs that have certain semantic features such as dynamicity and incompleteness (e.g., play, talk) and that it will not be incorrectly overextended to stative verbs (e.g., love, live).

Recent research (see Bardovi-Harlig; McManus) calls into question the predictive power of the Aspect Hypothesis as it has been found that it is advanced

[1] **Alaa Alzahrani** is a PhD student in Applied Linguistics at King Saud University, Saudi Arabia. Email: alaa-zahrani@hotmail.com.
[2] **Normala Othman** is an Associate Professor in the Department of English Language and Literature at King Saud University, Saudi Arabia. Email: norothman@ksu.edu.sa.

learners rather than beginners who show the developmental pattern predicted by the hypothesis. These findings are difficult to generalise given the methodological differences in the body of research. The observed tendency to rely on less communicatively driven elicitation measures, whether semi- or completely controlled, might not fully represent L2 development of tense-aspect as they fail to reflect learners' performance in more authentic contexts. It is no surprise then that there have been calls to incorporate naturalistic corpus data in investigating L2 developmental stages with an aim to produce more generalisable results (see Gilquin and Gries; Granger; Lozano and Mendikoetxe, for example).

2. Review of the Literature

Vendler's well-known classification of verbs, in his article "Verbs and Times," group them into four semantic types: (1) statives, (2) activities, (3) accomplishments and (4) achievements, as illustrated in Table 1:

Table 1: Vendler's classification of verbs

Lexical aspect	Description	Examples
Statives	The event is not dynamic and can continue without additional effort or energy being applied.	seem, know, need, want, love
Activities	The event has duration but with an arbitrary endpoint, and is homogenous in its structure.	rain, play, walk, talk, sleep, snow
Accomplishments	The event has some duration as well as a single clear inherent endpoint.	run a mile, make a chair, build a house, write a letter
Achievements	The event takes place instantaneously and is reducible to a single point in time.	arrive, leave, recognize, die, reach the summit

Adapted from Shirai and Andersen (744)

Given that the progressive morpheme typically conveys meanings related to "progressivity, imperfectivity, and dynamicity" (Collins 226), it is natural that the more dynamic a certain lexical aspect category is, the greater the possibility that the verb will be marked with an *-ing*. The descriptions in Table 1 show that the most dynamic lexical aspect is activity verbs, followed by accomplishments and achievements. Meanwhile, stative verbs lack the dynamicity feature characteristic of the *-ing* verb, and this explains the relatively rare occurrence of the *-ing* markers in this category. The level of compatibility between each lexical aspect category and dynamic events corresponds to the developmental order predicted by the Aspect Hypothesis.

The Aspect Hypothesis holds that L2 learners' early acquisition of verbal morphology is influenced by the inherent semantic characteristics of the verb (Andersen, "Developmental Sequences" 305). Based on this assumption, the hypothesis predicts the developmental path followed by L2 learners in the acquisition of morphemes, mainly that the development of the progressive marking would start with activity verbs, then expand to accomplishments and progress to achievements, without extending to statives. A related prediction maintains that if L2 learners extend the progressive *-ing* to a stative verb, then it is predicted that this overextension would be correct as suggested by Andersen and Shirai ("Primacy of Aspect" 559).

Two implications can be drawn from these predictions. One is that the skewed distribution of the progressive during the initial stages of learning L2 evens out in the later stages. This means that advanced learners will attach the progressive marker on activities, accomplishments and achievements more equally than beginners. As the Aspect Hypothesis assumes that there is variation in the distribution of the progressive across proficiency levels, a description of this change would be more meaningful if more than one level is examined. A second implication of these predictions relates to the universality of the development of L2 verbal morphology. Adopting a universal perspective in the study of learner interlanguage necessitates the rejection of first language (L1) influence on the L2 acquisition process, because evidently all L2 learners irrespective of their L1 are expected to progress in a similar way (Andersen and Shirai, "Primacy of Aspect" 560). Subsequent research ruled out L1 effects, as the predictions of the Aspect Hypothesis were confirmed across various instructional settings (e.g., instructed and uninstructed learners) and L1 backgrounds (see Bardovi-Harlig; Salaberry; Salaberry and Shirai, for example).

Another defining feature of the hypothesis is the distinction between lexical and grammatical aspects. Viewing these two notions as fundamentally distinct, Andersen and Shirai ("Discourse Motivations") argued for an exclusive focus on the implicit lexical aspect of verbs instead of the explicit grammatical aspect. To illustrate, the following example from Robison show a verb in a progressive (1) and a complete (2) grammatical aspect ("The Primacy of Aspect" 316). Despite having different grammatical aspects, both verbs have the same lexical aspect, that is, each expresses an accomplishment – an event with duration and a specific endpoint.

1. I ***was making*** a pair of pants. Accomplishment verb
2. I ***made*** a pair of pants. Accomplishment verb

However, the same verb may express more than one lexical aspect. Consider the following modified example from Cowan (355):

1. He ***was walking*** in the hall. Activity verb
2. He ***was walking*** to the post office. Accomplishment verb

The example indicates that it is important to consider verb arguments (i.e., objects) in the analysis of the verb's lexical aspect, a point that should be taken into consideration in analysing the data of this study.

Related Studies on Lexical Aspect

The theoretical discussion on the Aspect Hypothesis indicates that there are two explanatory factors for the reported developmental path of the progressive among L2 learners. These are lexical aspect and input frequency. A number of studies found that advanced L2 learners rather than beginners show the prototypical association between activity verbs (e.g., play) and the progressive marker, contrary to the Aspect Hypothesis. McManus administered a spoken narrative task and a sentence interpretation task to L2 French learners from two L1 backgrounds (English and German) and at two proficiency levels (advanced and beginner). It was found that there is an increase in the number of prototypical pairings (e.g., activity-progressive) as L2 proficiency increases (McManus 315). This correlates with earlier results that the more advanced learners favour the activity-progressive association as opposed to the less proficient learners (Bardovi-Harlig; Robison, "The Aspect Hypothesis Revisited").

A similar observation was made by Upor who elicited written data from EFL instructed learners across three educational levels in Tanzania by using picture compositions. The researcher found that the developmental pattern of the progressive was not in line with the prediction of the Aspect Hypothesis, as learners added the *-ing* to activities, followed by achievements then accomplishment and overgeneralised it to stative verbs. This contrasts with the Aspect Hypothesis which states that L2 learners would attach the *-ing* in the following sequence of verbs: activity, accomplishment, achievement, and without extending to stative. Similarly, contrary to the predictions of the hypothesis, this study reported that the more advanced learners made greater use of the progressive marker with activity verbs compared to lower proficiency learners.

The role of lexical aspect in learning the progressive was to a large extent established by Bardovi-Harlig and Bergström who conducted a cross-sectional study to examine the tense-aspect acquisition in written narratives across four proficiency levels from ESL learners. Three forms of the progressive: ∅-progressive (i.e., V+ing with no auxiliary), present progressive and past progressive were analysed and coded. They found that the progressive forms were commonly attached to activity verbs. Across each proficiency level, the association between the progressive *-ing* marking and lexical aspect was in the order of: (1) activity, (2) accomplishment, (3) achievement, with a few instances of overextension to stative verbs. This order of development confirmed the

prediction of the Aspect Hypothesis. However, support was not shown for the claim that there is a strong association at the initial stage of acquisition, as the percentages of the progressive forms used with activity verbs from level 1 to level 4 learners were 63%, 84%, 100% and 52%, respectively (Bardovi-Harlig and Bergström 325). The variance in percentages indicates that certain proficiency levels are responsive to the prediction of the Aspect Hypothesis, and that the beginners' use of the progressive is not largely influenced by the verb's lexical aspect.

These deviations from the Aspect Hypothesis can be explained in light of their research design. The stated studies did not control L2 proficiency as participants were placed in groups based on their instructional level instead of their performance on a proficiency test. As the researchers did not appropriately operationalise the distinction between advanced and beginner levels of L2 proficiency, the cited evidence against the predictions of the Aspect Hypothesis is questionable.

Even when the proficiency level is objectively determined, the Aspect Hypothesis may fail to accurately describe the initial acquisition of the progressive. In a study by Mueller, forty English native speakers were taught a miniature artificial language and were exposed to an equal number of activity and accomplishment verbs attached to progressive and past tense morphemes. Speaking and listening tests were administered to the participants, and a two-way factor ANOVA analysis found that both lexical aspect and morphological marking had a significant effect on the two tests scores. However, further analyses showed no interaction between lexical aspect and morphological marking. This suggests that the expected relationship between lexical aspect (e.g., activities) and verb morphology (e.g., the progressive) as predicted by the Aspect Hypothesis was not found. This was attributed to the fact that the participants were in the initial very short stage of learning the artificial language. The results might not be surprising given that a minimum proficiency level is essential to assess the predictions of the Aspect Hypothesis. A cross-sectional study design examining learners at different proficiency levels at the same time or a longitudinal one following the same group of learners over a period of time would be more suitable to trace L2 development of tense aspect.

Sugaya and Shirai attempted to explain why the progressive marker is mostly associated with activity verbs, while rarely extended to stative verbs by examining the effects of task type, L2 proficiency and L1 transfer on the L2 acquisition of Japanese progressive marker. Sixty-one advanced and beginner learners of Japanese from progressive and non-progressive languages participated in an oral task and an acceptability judgment test. While L1 effects were limited to beginners, task type and L2 proficiency had a greater impact on the exclusive use of the Japanese progressive marker "-teiru" with activity verbs (Sugaya and Shirai

30). Specifically, the progressive marker was strongly associated with activity verbs for lower proficiency learners.

Shami administered a fill-in-the-gap test and a two-option multiple choice task to EFL Saudi learners in order to examine the development of the English past progressive. A support for the Aspect Hypothesis was reported as learners mostly applied the progressive *-ing* to activity verbs followed by accomplishment. Farag adopted a more open-ended method by collecting thirty-seven essays from ESL Arab learners to examine the development of the progressive on the verbs. To do so, the study calculated the correct and incorrect uses of the progressive, arguing that an inappropriate use reveals that the learners relied on the lexical aspect of the verb to add an *-ing* rather than on interpreting the context. Using the linguistic tests for lexical aspect categories used by Shirai and Andersen, Farag found that the progressive was commonly used with activity verbs followed by stative verbs, indicating a partial support for the Aspect Hypothesis prediction (22).

Based on the reviewed studies so far, it is likely that the lack of support for the Aspect Hypothesis depends on three design issues. These issues arise if learners' proficiency was defined in terms of naturally occurring classes or groups, the data was inappropriate for capturing development and the sample size was too small for generalisations. This implies that the predictions of the Aspect Hypothesis can be adequately tested if proficiency is defined by a standardised scale, the data is cross-sectional or longitudinal and the sample size is representative.

3. The Study

The purpose of this study is to test the Aspect Hypothesis by examining the role of lexical aspect on the development of the English progressive in the writings of beginner and advanced EFL Emirati learners, using an Arab learner corpus that controls L1, proficiency level, age and task. The BUiD Arab Learner Corpus (BALC) is sourced for a sample of 200 essays composed by EFL learners at two proficiency levels. In order to investigate whether Emirati learners' use of the English progressive confirms the predictions of the Aspect Hypothesis, this study addresses the following question: What is the distribution of the progressive form used by beginner and advanced students in terms of the following lexical aspect categories: activities, accomplishments, achievements and statives?

4. Methodology

The BALC corpus

The BUiD Arab Learner Corpus (BALC) (Randall and Groom) contains 1,531 texts written by Grade 12 Emirati nationals for a large-scale and high-stakes language placement test called Common Educational Proficiency Assessment (CEPA) (Coombe and Davidson). The CEPA-writing section required the

students to write on a specific narrative and/or descriptive topic within 30 minutes. This corpus was deemed appropriate as it contains texts collected at one point in time from learners of different proficiency levels, allowing a valid comparison between lower and higher-level learners. A sample of 100 texts from the two levels was extracted and analysed in the study.

Learners' proficiency was estimated by their performance on the essay writing test scored by trained raters from the UAE's Ministry of Higher Education and Scientific Research and evaluated according to a 6-level scoring guide, ranging from 1 (low proficiency) to 6 (high proficiency), based on writing fluency and coherence, grammar, vocabulary, spelling, punctuation and content (Daiban). Essays receiving the same score were grouped together, resulting in 6 levels of language proficiency. They were further grouped into two proficiency levels, beginner (A1-A2) and advanced (C1-C2) levels. Other than identifying the learners' proficiency in the examined corpus, the corpus was cleaned and corrected, tagged and normalised to further ensure the reliability of analysis.

Tools and Procedure

Two corpus analysis tools were utilised in this study. Rayson's Wmatrix was used in the first stage of analysis as this web-based software offers a reliable automatic tagging system called CLAWS7 which assigns each word to a specific grammatical category. The other tool that enabled the analysis of corpus data is Anthony's AntConc which highlighted trends in the data through features such as concordances and frequency lists.

As BALC was primarily designed to investigate Arab learners' spelling mistakes, it contained a disproportionate amount of misspelled words, specifically in the lower levels, preventing automatic word retrieval and accuracy of word tagging. Given the large number of errors, the spelling error correction process was found to be inconsistent as sometimes grammatical mistakes were adjusted along with misspelled words. To ensure the consistency of error correction, a doctoral candidate who published a paper analysing the essays in BALC was contacted to provide an error-corrected copy of BALC (Elturki and Salsbury).

The corpus was then tagged for grammatical categories on Part-of-Speech (PoS) by CLAWS7 which is available on a web-based software tool for corpus processing called Wmatrix (Rayson). According to the data available on Wmatrix, CLAWS7 assigns a "VVG" tag to any *-ing* form of lexical verbs. This means that the "VVG" tag is assigned to *-ing* forms that act as a verb, a noun (e.g., enjoy *dancing* on loud music) and a modifier (e.g., the award-*winning* movie). This is exemplified in the sample below from the tagged advanced learners' corpus showing the "VVG" tag attached to verb as well as noun *-ing* forms:

> there I was ***siting*** VVG on the side walk outside alone ***saying*** VVG again and again if only I thought ahead.

As a result, the automatic grammatical tagging of the progressive needed manual editing. Out of the 1031 "VVG" tags, the manual check excluded 515 (53%) of the tags that were not verbs as identified mainly by the lack of a preceding form of "be" (e.g., is, are, was, were). Table 2 presents the total number of the automatically retrieved and retained "VVG" tags for the lower and advanced levels.

Table 2: Total instances of automatically identified and retained VVG tags (-ing forms)

	Identified VVG tags	Retained VVG tags
Level 2	248	205
Level 6	783	185

Table 2 shows that 83% of the automatically identified tags for level 2 were retained as opposed to 24% for level 6. This difference in the exclusion rate resulted from the use of two exclusion criteria to meet the developmental constraints of the beginner level and the rich vocabulary of the advanced level. It was observed that most beginners produced bare progressive forms with no preceding "be" forms (e.g., writing, studying) compared to the advanced learners. As grammatically incorrect forms were not excluded from the analysis of the beginners' essays, the exclusion rate for the beginner level was low.

The next step following the cleaning and tagging of the corpus was normalising the number of words in the corpus. It is important that comparative corpus studies normalise word frequencies to ensure a reliable comparison (see Gries; McEnery and Hardie, for example). Computing a normalised count was necessary as this study compared between two proficiency levels and because there was a huge difference between the number of words in each level: 16827 words for the lower level and 42698 words for the higher level.

A normalised number of words can be calculated when the raw frequencies are divided by the corpus size and then multiplied by any base number, which is commonly reported to be one million or ten thousand (Baroni and Evert 780). However, it is often noted that the choice of the base number should be estimated according to the size of the compared corpora (see Baroni and Evert; McEnery and Hardie, for example). Considering the number of words in the two examined levels, the most suitable base number was seen to be ten thousand, as it can minimise data distortion and prevent the inflation of rare categories. A web-based normalising calculator was used to enhance accuracy. To illustrate, the identified tokens of *-ing* verbs in level 2 were divided by 16827, which is the total number of words in level 2, and the resulting value was multiplied by 10,000. Similarly, level 6 was normalised following the same method.

Whereas the raw count shows that the advanced learners use more accomplishments (40) than beginners (24), the normalised count demonstrates

the opposite, as more accomplishments were used by beginners (14) than the more proficient learners (9), as shown in Table 3.

Table 3: Raw and normalised count of *-ing* verbs across the four lexical aspect categories

		Activities	Accomplishments	Achievements	States
Beginner	Raw count (Normalised)	169 (100)	*24 (14)*	9 (5)	3 (2)
Advanced		95 (22)	*40 (9)*	29 (7)	21 (5)

Italics = a difference between the raw and normalised count.

Data Coding

The progressive *-ing* was coded according to the following points. The main unit of analysis is the progressive verb which is "expressed by (a form of) 'be' in conjunction with a following (though not necessarily directly following) -ing participle" (Collins 226). As the Aspect Hypothesis is interested in the developmental stages of acquiring morphological markers, researchers within this framework have investigated learners' interlanguage expressions and in turn considered deviations from the target-like norms (Shirai, "Defining and Coding Data"). This deviation usually consisted of the progressive verb with no helping verb, such as "He ***driving*** fasted and he is crazy man…."

Due to the prevalence of the Ø-progressive construction in the lower level data, this study adapted the above definition to include this construction. Thus, a progressive verb is defined as an *-ing* form that can be expressed as: (1) the bare progressive (e.g., *writing*), (2) the present progressive (e.g., *is/are writing*) and (3) the past progressive (e.g., *was/were writing*). All correct and incorrect forms of *-ing* were coded in the study. It was useful to count atypical uses because limiting the analysis to well-formed verbs can run the risk of ignoring interlanguage forms that are more likely to reflect the initial association between verb types and the progressive marker. The decision to analyse non-target like forms is further justified by the literature (see Bardovi-Harlig and Bergström; Upor; Hackmann, for example).

Another coding decision concerns the coding of verb tokens rather than verb types since the verb's lexical aspect can only be understood in light of a specific context, as the literature review demonstrated. The analytic framework adopted in this study is Vendler's four-way classification of verbs, as given in Table 1 earlier. This operational test determined which category a verb belongs and enabled the results to be compared to previous works.

The coding procedure began with the 200 texts uploaded to Wmatrix and automatically PoS tagged by CLAWS7. After assigning a grammatical category for each word, Wmatrix generated PoS tagged files in TXT format. These generated text files were downloaded and explored on AntConc first by calculating the frequencies of the "VVG" tag in the 200 texts. Next, concordance

lines of each *-ing* word were extracted using AntConc and copied and pasted on a separate word-processing document for further analysis. Then, *-ing* tagged words were inspected and excluded if they did not meet the predetermined inclusion criteria. Following the tagging and cleaning process, each verb tagged as "VVG" in the data was coded for two variables: lexical aspect and proficiency level.

Coding Reliability

An inter-rater reliability test was conducted to ensure the reliability of the qualitative analysis. To guide the inter-rater procedure, coding instructions were adapted from Hackmann and presented to the second rater. The second rater recoded 10% of the data independently. Double coding a minimum of 10% of the data was recommended by Mackey and Gass (78) to ensure that the coding method is reliable. Cohen's kappa ($\varkappa$) was run to determine if there was agreement between the researcher's and the second rater's judgements on whether the thirty-nine *-ing* verbs encoded activity, accomplishment, achievement or stative features. There was fair agreement between the raters' judgements, $\varkappa = .341$.

Statistical Analyses

As the collected data were categorical rather than numerical, and since this study investigated the association of two categories (i.e., the progressive morphology and language proficiency), a chi-square test was employed, as it was deemed appropriate given that it can analyse nominal as well as ordinal data (Mitchell and Jolley 120), and because it tests the likelihood of association between two variables (Saunders, Lewis and Thornhill 78).

While the chi-square test allows the analysis of the categorical data in this study, it requires a sample greater than five. This condition was violated in some lexical aspect categories as they occurred less than five times. To solve this problem, the four lexical aspect categories were regrouped into two categories, e.g., category A and non-category A (i.e., the remaining categories were B, C and D). This resulted in conducting four separate chi-square tests for the following: activities and non-activities, accomplishments and non-accomplishments, achievements and non-achievements, and statives and non-statives.

After establishing binary variables, the association between the two learner groups' use of the progressive and lexical aspect category A was addressed in a chi-square test and adjusted for each situation. To determine the significance of association when expected frequencies are less than five, the Fisher's exact test results were reported instead of a chi-square.

5. Results

The distribution of the progressive *-ing* used by beginner and advanced students

in terms of activities, accomplishments, achievements and statives are given in the Table 4.

Table 4: Frequency of occurrence of the progressive morphology by proficiency level

Lexical Aspect/Level	Beginner		Advanced	
	#	%	#	%
Activities	100	82.6	22	51.2
Accomplishments	14	11.6	9	20.9
Achievements	5	4.1	7	16.3
Statives	2	1.7	5	11.6
TOTAL	121	100	43	100

Table 4 shows the higher frequency of the progressive on activities at the beginner level (100) relative to the advanced level (22). Similarly, more accomplishments inflected with the *-ing* appeared at the beginner level (14) than the advanced level (9). However, the advanced learner group used more achievements (7) and statives (5) in an *-ing* form, while the beginner group produced fewer achievements (5) and statives (2).

Both groups used considerably more activity verbs than non-activity ones with the *-ing* progressive marker; constituting 82.6% in the beginners' essays and 51.2% in the advanced learners' essays. Meanwhile, fewer *-ing* verbs were found in the other lexical aspect categories. Whereas the beginner learner infrequently attached the *-ing* marker to accomplishments (11.6%), achievements (4.1%) as well as statives (1.7%), advanced learners displayed a relatively frequent use of the *-ing* maker with accomplishments (20.9%), achievements (16.3%) and statives (11.6%).

Table 5. Progressive verbs across lexical aspect categories for both levels

	Frequency	Percentage
Activities	122	74.4
Accomplishments	23	14.0
Achievements	12	7.3
Statives	7	4.3
Total	164	100

The results as shown in Table 5 support the Aspect Hypothesis in that the progressive was more associated with activities, as the progressive was found in 122 instances (74.4%) to activity verbs, followed by accomplishment (14%), then achievements (7.3%) and, finally, statives (4.3%). A chi-square test further indicated the significance of the association between activity verbs in the *-ing* form and the proficiency of learners (χ 2(df 1) = 16.505, $p < 0.001$, $\varphi = .317$). To

evaluate the strength of the association between the two binary variables, the effect size was measured by a phi coefficient. The phi value (φ = .317) indicated that there was a weak positive association between the number of activity verbs marked with an *-ing* and language proficiency.

Accomplishment verbs are the second most frequent category to be inflected in the progressive (23), as shown in Table 5. However, no statistically significant relationship was found between learners' level and accomplishment verbs in the progressive ($p > 0.05$). On the other hand, despite the small number of *-ing* verbs categorised as achievements (12) and statives (7), they were significantly correlated with the progressive in the two proficiency levels (Fisher's exact test, $p < 0.015$ for achievements, $p < 0.014$ for statives).

The above results show evidence of a prototypical association between the progressive *-ing* and activities, in line with the Prototype Theory (Andersen and Shirai, "Primacy of Aspect"), but it remains unclear how this prototypical association took shape. In an attempt to understand the emergence of such reported association, this study investigated the role of frequency as a possible factor for learners' performance. The role of frequency was measured in terms of the distribution of verbs typically associated with the progressive in the native speech. This distribution was calculated by Wulff et al. based on data from the British National Corpus (BNC spoken) and Michigan Corpus of Academic Spoken English (MICASE). It was argued by Wulff et al. that these two corpora fairly represent the type of language adult L2 learners are usually exposed to, justifying this by their selection of those corpora (362). The same argument can be advanced for examining these two native speakers' corpora in this study. This suggests that it is assumed that the learners involved in this study are exposed to a variety of native speech similar to the one captured in BNC spoken and MICASE.

Based on the assumption that L2 learners are sensitive to the input they receive, the role of input frequency in shaping the prototypical progressive-activity association among the examined learners was investigated. This was done by comparing the ten most distinctive and the ten most frequent progressive verbs in native speech and the learners' essays, respectively. Rather than being simply frequent, distinctive progressive verbs are statistically much more positively associated with the progressive than any other verb occurring with the progressive (Wulff et al. 362). Table 6 shows the ten most frequent verbs produced in the *-ing* progressive by the examined learners and native English speakers.

Table 6: The 10 most distinctive and frequent progressive verbs among native speakers and the examined L2 learners

	Native speakers		L2 learners	
	BNC spoken	**MICASE**	**Beginners**	**Advanced**
1.	look*	try	go**	go**
2.	come*	look*	play**	wait*
3.	sit*	work	swim	get
4.	play**	go**	read	head
5.	wait*	move	travel	rain
6.	walk*	sit*	watch*	play**
7.	joke	wonder	talk	look*
8.	run	deal	eat	sit*
9.	watch*	play**	walk*	talk
10.	deal	miss	race	come*

Native speakers' data are adopted from Wulff et al. (362). * occurred in native speech and one learner group's data. ** occurred in native speech and the two learner groups' data.

Table 6 shows that learners sometimes associate the *-ing* marker with verbs that are also frequent in the progressive in the input. Out of the 16 verbs distinctly associated with an *-ing* by natives, six were found at the advanced level as opposed to four at the beginner level. Only *go* and *play* occurred in both of the native and learners' data. Other verbs such as *watch, walk, wait, look, sit* and *come* were commonly formed in the progressive within one learner group as well as in native speech. Compared to beginners, the advanced learner column has more verbs that tend to be produced in the progressive by native speakers.

6. Discussion

The study found that the distribution of the English progressive among beginner and advanced learners supports to some extent the Aspect Hypothesis as a chi-square test revealed a significant association between activity verbs with the progressive *-ing* marker across the two proficiency levels. Learners tended to supply the progressive morpheme on activity verbs (74.4%) more than accomplishment, achievement and stative verbs. Contrary to what was reported in recent studies on the Aspect Hypothesis (see Upor; McManus; Hackmann, for example), it was found that the predicted link between verb semantics and morphology is strongest at the early stages of L2 development. This suggests that the progressive-activity association guides the L2 acquisition of tense-aspect, giving support to the earlier studies (see Chan et al.; Housen, for example).

This finding is consistent with the Aspect Hypothesis as the impact of lexical aspect was more visible in the beginners' production of the progressive than the advanced group. Token frequencies highlighted a strong preference for attaching

the *-ing* on activity verbs in the early stages (A1-A2 CEFR). It also showed that the bias of the *-ing* towards activity verbs spread out as learners became more proficient users (C1-C2 CEFR) and extended the application of the progressive across the lexical aspect categories. While a similar developmental path was noted by Sugaya and Shirai (33), different developmental stages were reported by earlier studies. Some concluded that the prototypical association between activities and the *-ing* was strongest in the more advanced stage (Upor 151) or the intermediate level (Bardovi-Harlig and Bergström 325). Yet, Mueller (18) found no lexical aspect effects on the initial acquisition of the progressive.

Lack of sensitivity to lexical aspect in the early stages as suggested by Shirai (22) and Sugaya and Shirai (30) may be due to the observed task effects in tense-aspect acquisition studies. Additionally, participants in studies examining the Aspect Hypothesis are not always assigned to identifiable proficiency levels, resulting in inconsistent reports on the most sensitive developmental stage to lexical aspect effects. Another explanation for the difference between the results of this study and previous works is that the beginner learners in this study might have had restricted form-meaning mappings; in that they expect each progressive marker to denote "action-in-progress" (Andersen and Shirai, "Discourse Motivations" 143) as opposed to the more advanced learners who would form multiple form-meanings for the progressive marker. Learners who have multiple form-meanings of the progressive marker can use it across all verb types. This gradual spread can be seen at the advanced level, where less typical categories such as achievements and statives were increasingly supplied with an *-ing*, indicating that the distribution of the progressive morpheme would become less skewed over time. The examples in Table 7 illustrate this pattern in the data.

Table 7: Samples from the advanced learners' essays

1. The movie revolved around a cop, who declined the position to be a CIA in part one and ended up a protector of the helpless, who is also ***expecting*** a baby from his newly married wife 2. You know I was ***thinking*** of my dad.	Stative
3. It's pretty amazing, I bet your [you're] ***getting*** an idea of what I am saying.	Achievement

This extension to the non-prototypical categories is done by advanced learners to "impose their own perspective on situations that can be viewed from different perspectives" (Andersen and Shirai, "Discourse Motivations" 148). Both *expect* and *think* are non-dynamic events which are largely considered to be stative verbs, yet examples (1) and (2) show that they were written in the progressive form. It

can be seen the progressive in example (1) is used to show the reader that there are more than one event happening at the same time, in that the cop was *expecting* a baby when he *declined* a job. As for example (2), *think* was used in the progressive to highlight the duration of the thinking process the learner experienced as he thought of his *dad*. Rather than treating *thinking* as a completed event that took time in the past, the learner wanted to highlight the durative aspect of the thinking event by using the progressive, triggering an interpretation of a continuous and time-consuming event. A similar reading can be noticed in example (3) which includes an achievement verb with an inherent end-point rather than dynamicity in a progressive form. All of these examples suggest that the more advanced learners can extend the progressive aspect to verbs that do not have a durative meaning.

Despite the general support for the Aspect Hypothesis, a deviation from the expected distribution was found in the essays. The distributional sequence of the progressive marking for beginner learners should be: (1) activities, (2) accomplishment, (3) achievement (4) but without extending to statives. In other words, the Aspect Hypothesis predicted that there will be no overextension to stative verbs. However, the beginner learners in this study added the *-ing* to stative verbs (1.7%). Overextending the *-ing* to statives in the early stages of acquisition was reported across different instructional contexts (e.g., instructed and uninstructed) as well as age groups (see Bardovi-Harlig and Bergström; Housen; Chan et al., for example).

The frequency analysis revealed a possible role for frequency effect on the development of the progressive. This supports the work of Wulff et al. on the role of frequency on the acquisition of tense-aspect morphology. The current study showed that nearly half of the most frequent progressive verbs produced by the advanced level learners were also typically used in the progressive form by native speakers. Beginners, on the other hand, used a small number of verbs that are strongly associated with the progressive in native talk. This is in line with the results found in McManus (315) who attributed the impact of input frequency on L2 French advanced learners to their exposure to French in a naturalistic setting. This explanation can be applied to the results of the present study as well. However, given that BALC does not give information on whether learners spent time in an English-speaking country, it is hard to attest such a conclusion.

Another explanation that can account for the reported differences in frequency effects in relation to proficiency level in this study is the difference in type-token ratio. In the present study, the beginners applied the progressive marker to a limited type of frequently used activity verbs compared to the advanced learners who produced various verb types in their essays. The difference in type-token ratio in the two learner groups can then explain the variance in their use of verbs that are highly correlated with the progressive aspect in spoken English.

7. Conclusion

This study examined 200 texts by advanced and beginner Arab EFL learners to find if the distribution of the English progressive in those texts followed the Aspect Hypothesis. Results showed that the predicted association between the progressive and activity verbs was statistically significant, indicating that both learner groups mainly used the progressive with activity verbs and the two learner groups showed a preference for associating the progressive with activity verbs first, before the other verbs.

Apart from lexical aspect, frequency of the progressive in native talk was investigated as an explanation for the observed acquisition pattern and more specifically the progressive-activity association. While an overlap between the learners' and native speakers' use of progressive verbs was found, this interaction was not further tested for statistical significance given the small data set. Thus, it was not possible for this study to suggest that input frequency has a role in the L2 acquisition of tense-aspect.

Generally, the findings of this study are in line with the Aspect Hypothesis. To place this study within the larger body of research on the Aspect Hypothesis, theoretical implications of these findings should be discussed. One implication is that lexical aspect can account for the early acquisition and distribution of the English progressive as the lower proficiency level extensively attached the progressive on activity verbs. This is contrary to the recent research on the Aspect Hypothesis who found no impact of lexical aspect on the emergence of the progressive among L2 beginners. A second implication is that the effects of lexical aspect on the distribution of the progressive continue even at an advanced proficiency level. A further point of interest emerging from this study is that learner corpus data support the prototypical association between the progressive marker and activity verbs, suggesting that corpus-based data can be valuable in the study of interlanguage stages.

Works Cited

Al-Thubaiti, Kholoud. A. "L2 Acquisition of English Aspect by L1 Arabic Speakers: The Role of Interpretable Features at the Syntax-semantics Interface." *The Acquisition of the Present*. Ed. Dalila Ayoun. Amsterdam: Benjamin, 2015. 185-214.

Alruwaili, Ruwayshid. "Second Language Acquisition of Aspect and Tense by Saudi-Arabic Learners of English." Doctoral Dissertation, University of York, UK, 2014 (Unpublished). http://etheses.whiterose.ac.uk/7203/. 16 February 2018.

Andersen, Roger. W. "Developmental Sequences: The Emergence of Aspect Marking in Second Language Acquisition." *Crosscurrents in Second Language*

Acquisition and Linguistic Theories. Eds. Thom Huebner and Charles A. Ferguson: John Benjamins, 1991. 305-24.

Andersen, Roger. W. and Yasuhiro Shirai. "Discourse Motivations for Some Cognitive Acquisition Principles." *Studies in Second Language Acquisition* 16.2 (1994): 133-56.

---. "Primacy of Aspect in First and Second Language Acquisition: The Pidgin/Creole Connection." *Handbook of Second Language Acquisition*. Eds. William C. Ritchie and Tej K. Bhatia. San Diego: Academic Press, 1996. 527-70.

Anthony, Laurence. "Laurence Anthony's Website: AntConc." Tokyo, Japan: Waseda University, 2010. http://www.antlab.sci.waseda.ac.jp/, 2010. 20 February, 2018.

Bardovi-Harlig, Kathleen. *Tense and Aspect in Second Language Acquisition: Form, Meaning and Use*. Oxford: Blackwell, 2000.

Bardovi-Harlig, Kathleen and Anna Bergström. "The Acquisition of Tense and Aspect in SLA and FLL: A Study of Learner Narratives in English (SL) and French (FL)." *Canadian Modern Language Review* 52 (1996): 308-30.

Baroni, Marco and Stefan Evert. "Statistical Methods for Corpus Exploitation." *Corpus Linguistics: An International Handbook*. Eds. Anke Lüdeling and Merja Kytö. Berlin: De Gruyter, 2009. 777-803.

Chan, Ho. L., Jessica Finberg, Willie Costello and Yasuhiro Shirai. "L2 Acquisition of Tense-aspect Morphology: Lexical Aspect, Morphological Regularity, and Transfer." *Space and Time in Languages and Cultures: Linguistic Diversity*. Eds. Luna Filipović and Kataryzna M. Jaszczolt. Amsterdam: John Benjamins, 2012. 181-204. https://doi.org/10.1075/hcp.36.12chan. 20 February 2018.

Collins, Peter. "The Progressive Aspect in World Englishes: A Corpus-based Study." *Australian Journal of Linguistics* 28.2 (2008): 225-49.

Coombe, Christine and Peter Davidson. "Common Educational Proficiency Assessment (CEPA) in English." *Language Testing* 31.2 (2014): 269-76.

Council of Europe. *Common European Framework of Reference for Language Learning and Teaching*. Cambridge: Cambridge UP, 2001.

Cowan, Ron. *The Teacher's Grammar of English*. Cambridge: Cambridge UP, 2008.

Daiban, Salma. "The Examination of the Psychometric Quality of the Common Educational Proficiency Assessment (CEPA)-English Test." Doctoral Dissertation. University of Pittsburgh, 2009 (Unpublished). http://d-scholarship.pitt.edu/id/eprint/10300, 2009. 25 February 2018.

Elturki, Eman and Tom Salsbury. "A Cross-sectional Investigation of the Development of Modality in English Language Learners' Writing: A Corpus-driven Study." *Issues in Applied Linguistics* 20 (2016): 51-72.

Farag, Islam. M. "The Influence of the Lexical Aspect Hypothesis on Arabic-speaking ESL Students' Acquisition of the Progressive Aspect." *MIDTESOL*

2015. 116-28.

Gilquin, Gaetanelle Stefan T. Gries. "Corpora and Experimental Methods: A State-of-the-art Review." *Corpus Linguistics and Linguistic Theory* 5.1 (2009): 1-26.

Granger, Sylviane. "How to Use Foreign and Second Language Learner Corpora." *Research Methods in Second Language Acquisition: A Practical Guide*. 1st ed. Eds. A. Mackey and S.M. Gass. Oxford: Blackwell, 2012. 7-29.

Gries, Stefan, T. "Useful Statistics for Corpus Linguistics." *A Mosaic of Corpus Linguistics: Selected Approaches*. Eds. A. Sanchez and M. Almela. Frankfurt: Peter Lang, 2010. 269-91.

Hackmann, Emily, E. "Story Time: The Expression of Temporal Events in Narration by L2 Learners of German." Doctoral Dissertation, University of Kansas, 2016 (Unpublished). 27 February 2018.

Housen, Alex. "A Corpus-based Study of the L2 Acquisition of the English Verb System." *Computer Learner Corpora, Second Language Acquisition and Foreign Language Learning*. Eds. Stefan Granger, J. Hung and S. Petch-Tyson. Amsterdam: John Benjamins, 2002. 77-116)

Lozano, Cristobal and Amaya Mendikoetxea, "Learner Corpora and Second Language Acquisition: The Design and Collection of CEDEL2." *Automatic Treatment and Analysis of Learner Corpus Data*. Eds. Ana Díaz-Negrillo, Nicolas Ballier and Paul Thompson. Amsterdam: John Benjamins, 2013. 65-100.

Mackey, Alison and Susan M. Gass. *Second Language Research: Methodology and Design*. Mahwah, NJ: Lawrence Erlbaum, 2005.

McEnery, Tony Andrew Hardie. *Corpus Linguistics: Method, Theory and Practice*. New York: Cambridge UP, 2011.

McManus, Kevin. "Prototypical Influence in Second Language Acquisition: What Now for the Aspect Hypothesis." *International Review of Applied Linguistics in Language Teaching* 51.3 (2013): 299-322.

Mitchell, Mark L. and Janina M. Jolley. *Research Design Explained*. 7th ed. Belmont, United States: Wadsworth, 2010.

Mueller, Charles, M. "Initial Acquisition of Tense-aspect Morphology in an Artificial Language." *Second Language Research* (2018): 1-22.

"Normalizing Word Counts." http://www.thegrammarlab. com/?p=160. 16 February 2018.

Randall, Mick and Nicholas Groom. "Introducing the BUiD Arab Learner Corpus: A Resource for Studying the Acquisition of L2 English Spelling." *Proceedings of the Corpus Linguistics Conference 2009*. Eds. Michaela Mahlberg, Victoria G. Díaz and Catherine Smith. Liverpool: University of Liverpool, 2009.

Rayson, Paul. "From Key Words to Key Semantic Domains." *International Journal of Corpus Linguistics* 13.4 (2008): 519-49. https://doi.org/10.1075/ijcl.13.4.06ray, 2008. 15 February 2019.

Robison, Richard E. "The Primacy of Aspect: Aspectual Marking in English Interlanguage." *Studies in Second Language Acquisition* 12.3 (1990), 315-30. https://doi.org/https://doi.org /10.1017/S0272263100009190, 1990. 15 February 2019.

Robison, Richard E. "The Aspect Hypothesis Revisited: A Cross-sectional Study of Tense and Aspect Marking in Interlanguage." *Applied Linguistics* 16.3 (1995): 344-370.

Salaberry, Maximo. *Marking Past Tense in Second Language Acquisition: A Theoretical Model*. London: Continuum, 2008.

Salaberry, Maximo and Yasuhiro Shirai, eds. *The L2 Acquisition of Tense-aspect Morphology*. Amsterdam: John Benjamins, 2002.

Saunders, Mark, Philip Lewis and Adrian Thornhill. *Research Methods for Business Students*. 5th ed. Harlow, England: Prentice Hall, 2008.

Shami, Joman H. *The Role of Lexical Aspect in the Acquisition of English Past Tense by Saudi EFL Learners*. Doctoral Dissertation, Lancaster University. Retrieved from EThOS. (556640), 2010. 10 February 2018.

Shirai, Yasuhiro. "A Multiple-factor Account to Form-meaning Connection in the Acquisition of Tense-aspect Morphology." *Form-Meaning Connections in Second Language Acquisition*. Eds. Bill VanPatten, Jessica Williams, Sussane Rott and Mark Overstreet. NJ, US: Lawrence Erlbaum Associates Publishers, 2004. 1-26.

Shirai, Yasuhiro. "Defining and Coding Data: Lexical Aspect in L2 Studies." *Research Design and Methodology in Studies on L2 Tense and Aspect*. Eds. Maximo Salaberry, Llorenc Comajoan and Paz Gonzalez. Boston, USA: De Gruyter Mouton, 2013. 271-308.

Shirai, Yasuhiro and Roger W. Andersen. "The Acquisition of Tense-aspect Morphology: A Prototype Account." *Language* 71.4 (1995): 743-762. Retrieved from http://www.jstor.org/stable/415743. 10 February 2018.

Sugaya, Natsue and Yasuhiro Shirai. "The Acquisition of Progressive and Resultative Meanings of the Imperfective Aspect Marker by L2 Learners of Japanese: Transfer, Universals, or Multiple Factors?" *Studies in Second Language Acquisition* 29.1 (2007): 1-38.

UAE Ministry of Higher Education and Scientific Research. *Interpreting CEPA Scores: FAQ*. Retrieved from http://ws1.mohesr.ae/cepa/Files/Public_ CEPA_ English_Specifications.pdf, 2014. 25 February 2018.

Upor, Rose A. *The Acquisition of Tense-aspect Morphology Among Tanzanian EFL Learners*. Doctoral Dissertation. Retrieved from University of Georgia Electronic Theses and Dissertations. (25687), 2009. February 4, 2018.

Vendler, Zeno. "Verbs and Times." *Linguistics and Philosophy* 66.2 (1957): 97-121.

Wang, Jing and Yasuhiro Shirai. "The Development of Third Person Singular Present Form -s: Verb Semantics or Input Frequency?" *The Acquisition of the Present*. Ed. Dalila Ayoun. Amesterdam: Benjamin, 2015. 1-20.

Wulff, Stefanie, Nick C.Ellis, Ute Römer, Kathleen Bardovi-Harlig and Chelsea J. Leblanc, "The Acquisition of Tense-aspect: Converging Evidence from Corpora and Telicity Ratings." *The Modern Language Journal* 93.3 (2009): 354-69.