

COLLOCATIONAL STRATEGIES OF ARAB LEARNERS OF ENGLISH: A STUDY IN LEXICAL SEMANTICS

Muhammad Raji Zughoul

Hussein S. Abdul-Fattah

Department of English, Yarmouk University, Irbid, Jordan

1. Introduction and Review of the Literature

Researchers (e.g. Zughoul, 1991; 1995; Hussein, 1998) reiterated their awareness of the fact that the study of the lexicon which can be classified under interlanguage semantics has not received as much emphasis as the other two components of interlanguage, viz., phonology and syntax in language learning/ teaching research. No doubt both phonology and syntax lend themselves more easily to rigorous analysis within the different frames of linguistic analysis. The lack of emphasis on lexis in traditional pedagogy and instructional materials was influenced by the tenets of audiolingualism which generally relegate lexical learning to secondary status in contrast with phonology and syntax (cf. Ramsey, 1981).

The development of word lists for English based on frequency distribution (Fries and Traver, 1940; Thorndike and Lorge, 1944; West, 1959; Kucera and Francis, 1967; Carrol, et al, 1972), directed the attention of teachers, ELT practitioners and curriculum specialists alike to the upper ends of these lists, to the neglect of the lower ends on the premise that mastering the upper ends would be adequate for the development of a good measure of proficiency in English. Besides, the upper ends would be

efficiently exploited for learning English phonology and syntax. This restricted view of vocabulary, beside its notional falseness from a statistical point of view has caused learner lexical deficiency and incapability to function adequately in real life situations. Work in the area of corpus linguistics (e.g. Twadell, 1973; Kucera and Francis, 1967) has convincingly triggered the urge for a reconsideration of the role of vocabulary in FL instruction. It has indicated beyond doubt that language pedagogues have been on the wrong track in their assumptions about the role of vocabulary frequency distribution. The counter argument has led to the recognition of a fundamental role for lexical learning. Judd (1978) suggests that more emphasis should be given to direct vocabulary teaching. Wilkins (1972) holds the same view. Quite recently, Verstraten (1992), even more, proposes the learning of a great quantity of lexical elements and fixed phrases. She believes that the ability to use such lexical constituents is a good indicator of language proficiency.

On another plane, interlanguage lexical studies which have been conducted on learners of various language backgrounds reveal roughly similar trends. In her classic study, Duskova (1969) concludes that lexical errors form less homogeneous material for study than syntactic errors. She established a typology of lexical errors deriving from formal similarity, semantic relatedness, assumed equivalence, and or distortion. Her limited taxonomy has motivated several other similar studies including those of Ghadessy (1976) on Persian learners of English, and Al-Ani (1979) on Arab students. Later on, Arabski (1979) conducted a large-scale quantitative and qualitative analysis of the lexical errors of Polish learners of English based on a large corpus. The errors were conceived of as indicators of the development of the learners' interlanguage. Based on Blum and Levenston's (1976) universal principles of simplification, Arabski provides a typology of eight categories of lexical error, including the use of Polish words, morphological similarity, graphic similarity, hyponymy, primary counterpart coinage, semantic similarity, and others.

COLLOCATIONAL STRATEGIES

To narrow the spectrum of discussion, the area of collocation within the realm of lexis is of prime importance and forms a serious problem for language learners. Wardell (1991) points out that one peculiarity of the learners of English is their failure to produce collocations in the proper order. These linguistic sequences do not follow a prescribed pattern as they are not rule-governed. While NL speakers acquire them throughout the natural acquisition process, FL learners need to be instructed and trained in producing them in the proper context. Fluency in the FL, as implied by Kjellemer (1992), is enhanced by 'automation of collocation'. The more the learner is capable of producing the correct collocations, the fewer hesitations or pauses he makes in long chunks of discourse and consequently, the more competent in the FL he becomes.

2. What is Collocation?

The definition of a collocation is not much a matter of controversy among linguists. The British linguist, Firth is often quoted as one of the earliest who tackled the collocational phenomenon. According to Palmer (1981: 94), Firth argues that "You shall know a word by the company it keeps." He exemplifies this by the English word *ass* which occurs in a limited set of contexts (You silly _____; Don't be such an _____) and in the company of a limited set of adjectives (*silly, obstinate, stupid, and awful*). In an article on modes of meaning published back in 1951, Firth introduces his classic definition of collocation as "the company words keep together." He also maintains that the meaning of collocation is an abstraction at the syntagmatic level and is not directly associated with the conceptual or ideational approach to lexical meaning. He gives an illustrative example of the word *night* where one of its meanings allows its collocability with *dark*. Leech (1974: 20) discusses seven categories of meaning including what he terms "collocative meaning" which consists of "the associations a word acquires on account of the meaning of words which tend to occur in its environment." This conceptualization of collocation is almost a replica of

Firth's aforementioned definition, but instead of *ass*, Leech exemplifies *pretty* and *handsome* and the collocates of each. These words share the common grounds of 'good looking', but they are differentiated by the range of nouns with which they are likely to co-occur. Leech also gives examples of 'quasi-synonymous' verbs like (*wander / stroll* and *tremble / quiver*), where each keeps a different company from the other. More recently, Ilson (1986) has developed criteria for defining these linguistic sequences including the notions of relative fixedness and non-idiomaticalness; that is, recurrent combination and fixed combination.

Cruse (1986) offers a more comprehensive, more exclusive, and more formal definition. He has developed criteria to distinguish collocations from idioms. According to him, collocations refer to "sequences of lexical items which habitually co-occur, but which are nonetheless fully transparent in the sense that each lexical constituent is also a semantic constituent." Moreover, collocations have a kind of 'semantic cohesion' where the constituent elements are mutually selective. For example, the word *heavy* in *heavy drinker*, *heavy smoker*, and *heavy buyer* has a defined contextual environment which requires the selection of the notion of 'consumption' in the immediate environment. Accordingly, such expressions as *fine weather*, *torrential rain*, *light drizzle*, *high winds* are also examples of collocation. An idiom, on the other hand, is defined in terms of two requirements: lexical complexity, i. e. it consists of more than one single lexical constituent, and single minimal signification, i.e., it has a single minimal semantic constituent. A minimal semantic constituent is indivisible into semantic constituents. Thus, expressions like *to pull someone's leg*, *to have a bee in one's bonnet*, *to kick the bucket*, *to cook someone's goose*, *to be off one's rocker*, *round the bend*, and *up the creek* are all examples of idioms. Traditionally, the main criterion for defining an idiom is the fact that its meaning cannot be inferred from the meanings of its parts (cf. Fromkin and Fromkin, 1989). There are 'transitional areas'; however, where collocations border idioms. Bound (inseparable constituents) collocations such as *foot*

the bill and *curry favour* are a case in point. These expressions are semantically transparent, but idiom-like, because *bill* is freely modifiable in expressions like the *electricity bill* and the *bloody bills*. A distinct, idiom-like characteristic of the collocational sequence is that *foot* requires the presence of a specific lexical partner and resists interruption.

It is noteworthy here that different languages have different collocational modes: what collocates in one language does not necessarily collocate in another (Zughoul, 1991). Thus, some collocations may sound odd and out of place when translated. Consider, for example, the translation into Arabic of *false teeth* or *false beard* where the collocant *false* implies 'untrue' and relates to 'lying'.

3. Objectives and Methodology

The objective of this paper is tri-dimensional. It investigates the use of collocates as an indicator of language proficiency. It also investigates the strategies Arab learners of English at different levels of competence use in their attempts to come up with the proper collocation and it demonstrates how equivalent English and Arabic words combine differently.

For these purposes, the writers have chosen to test the learners' proficiency in rendering into English a range of the collocations of the Arabic verb *kasara* 'broke'. Obviously, words manifest different paradigms of collocability, depending on their syntagmatic distribution. It may not be feasible to identify the whole range of the collocability of a certain lexical item as these are open-ended. Nonetheless, the writers attempted to pursue the collocational range of *kasara*, but, admittedly, the list is not exhaustive since many collocations emerge as idioms and, thus, fall beyond the scope of the present study. Sixteen such lexical sequences with *kasara* have been cited and administered in a two-form test to elicit the receptive and productive competence of Arab learners of English in the collocations of this verb. The assumption is that despite the familiarity of the constituents of

those collocations to the learners in both NL and FL, their cross-linguistic translatability is primarily a function of the learners' awareness of their collocability in the FL.

The first form of the test is a multiple-choice task, a translation of the sixteen *kasara*-collocations given to the study sample subjects to recognize the appropriate collocant from plausible alternatives. The other form of the test is a free translation task of the same 16 Arabic collocations (see appendix) administered for the same subjects to explore their proficiency in this linguistic area. The second task was performed prior to the first to avoid any potential post-test effect on the free translation choices of the subjects. The target items are familiar in Jordanian Arabic. Nevertheless, they, along with their translation, were validated by a jury of specialists to ensure their content and face validity. Their reliability co-efficient was established by a test-retest procedure. Pearson's correlation co-efficient was computed for the compatibility of the correct responses on both tasks and t-tests were applied to assess the significance of the difference between the means of the study groups' correct responses on both tasks.

The study sample consisted of two groups of EFL university students (38 graduates and 32 third-year undergraduates) from the Department of English at Yarmouk University, Jordan. The responses were tallied and analysed, and the deviant responses were analysed in terms of hypothetical hunches made by the writers leading to the characterization of the learners' communicative strategies in their endeavour to convey the target collocations from Arabic into English. (The numbers opposite to the cited examples refer to the original items in the appendix). The overall frequency and percentages for the undergraduates', graduates', and all subjects' responses were computed out of 512 (32 x 16), 608 (38 x 16) and 1120 (70 x 16), respectively.

4. Findings and Discussion

4.1 Collocational Proficiency

Task1, a controlled multiple-choice lexical translation, was meant to elicit the learners' receptive competence in recognizing the correct English collocations equivalent to those of the Arabic verb *kasara* 'broke'. Task 2, a free translation of the same Arabic collocations, was intended to explore the same subjects' productive proficiency in this linguistic phenomenon. Both tasks are viewed as complementary indicators of the learners' competence in this TL area. The validation process showed that the Arabic lexical sequences meet Cruse's (1986) definition of collocation (see Introduction above). Furthermore, Pearson's correlation co-efficient was high (.0.91), which suggests a strong link between the subjects' competence on both tasks (see Table 2 below).

Data analysis reveals that the overall performance of the subjects in the target collocations is far from being satisfactory, bearing in mind that even the undergraduate students are Junior English language majors. As indicated in Table I, only 50.27% of the total attempts of all subjects on the objective recognition task were correct, with less correct performance by the undergraduates (42.77% vs. 56.58%, respectively). The results on the productive, free translation task were even worse only 38.75% of the subjects' total attempts were correct, with less correct performance by the undergraduate group (36.52% vs. 43.76%). All the subjects succeeded in 44.50% of all the responses on both tasks. T-tests in Table 2 indicate that the differences between the groups' performance on both tasks are statistically significant in favour of the graduates (t-values for the paired differences are 3.76, 4.73 and 5.47 for undergraduates, graduates and all subjects, respectively). This disparity is understood in terms of the difference between the two groups' exposure to the TL. Among the collocations which obtained the highest number of correct response on both tasks are items (1, 6, 10, 11, 13 and 15):

1. *She broke her husband's oath.*

2. He *broke her heart*.
3. Some workers *broke the strike* and returned to work.
4. The police *broke the law* when they arrested the MPs.
5. The *waves broke* on the shore rocks.
6. The *enemy was defeated* in the battle.

It appears that these collocations were the easiest for the subjects on both tasks because they seem to be more transparent than the others and recur more frequently in their literature readings.

Table 1: Frequency and Percentage of Subject Correct Responses

Item	Undergraduates (N = 32)				Graduates (N = 38)				All (N = 70)				Grand Total	
	Rec. Task		Prod. Task		Rec. Task		Prod. Task		Rec. Task		Prod. Task			
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
1	21	65.63	18	56.25	24	63.16	20	52.63	45	64.29	38	54.29	83	3.67
2	7	21.88	8	25.00	29	76.32	20	52.63	36	51.43	28	40.00	64	2.86
3	12	37.50	8	25.00	14	36.84	10	26.32	26	37.14	18	25.71	44	1.96
4	15	46.88	10	31.25	20	52.63	6	15.79	35	50.00	16	22.86	51	2.28
5	14	43.76	10	31.25	21	55.26	16	42.11	35	50.00	26	37.14	61	2.72
6	14	43.76	12	37.50	35	92.11	28	73.68	49	70.00	41	57.14	90	3.97
7	6	18.75	3	9.38	14	36.84	6	15.79	20	28.57	9	12.86	29	1.29
8	12	37.50	8	25.00	18	47.37	7	18.42	30	42.86	15	21.43	45	2.01
9	5	15.62	6	18.75	9	23.68	6	15.79	14	20.00	12	17.14	26	1.16
10	21	65.63	18	56.25	26	68.42	22	57.90	47	67.14	39	54.29	86	3.67
11	22	68.75	21	65.63	28	73.68	24	63.16	50	71.29	45	64.29	95	4.21
12	7	21.88	6	18.75	18	47.37	6	15.79	25	35.71	12	17.14	37	1.61
13	22	68.75	22	68.75	33	86.84	20	52.63	55	78.57	42	60.00	97	4.33
14	9	28.13	7	21.88	6	15.79	12	31.58	15	21.29	19	27.14	34	1.52
15	26	81.25	22	68.75	31	81.58	26	68.42	57	81.43	48	68.57	105	4.69
16	6	18.75	8	25.00	18	47.37	18	47.37	24	34.29	26	37.14	50	2.23
Total	219	42.77	187	36.52	344	56.58	247	43.76	563	50.27	434	38.75	997	44.50

Surprisingly in addition, a good number (1.52%) of the subjects' total responses on the free translation task consisted of irrelevant utterances or message substitution which may have been triggered by an unconscious or wrong interpretation of the target collocational expression due to lack of

COLLOCATIONAL STRATEGIES

concentration, fatigue or carelessness while performing the task. The following utterances are examples of irrelevant responses or message change of the target test items:

7. She broke her husband's arm.
8. He ate all the food.
9. He opened all the buffet a while ago.
10. He didn't have juice in the coffee shop.
11. He cleaned his illness by water.
12. He drank the bitterness of the medicine.

Furthermore, a cursory look at Table 1 reveals a disparity between the students' receptive and productive competence in collocation, a finding supported by other researchers (cf, Shakir and Shdeifat, 1996). Both groups manifested lower competence in the production task on most of the target collocations; only four items (2, 9, 14 and 16, see appendix) reflect an adverse tendency. Obviously, production tasks provide more challenge than objective, multiple choice tasks since in the latter subjects have options at their disposal and hunching is always possible. Besides, it is plausible that the testees were bewildered by the semantically-related options given for those items in the recognition task, thus, narrowing their intuitive, selectional ability and distracting their recognition of the appropriate collocates.

Table 2 Correlation, t-values for Means of Graduates and Undergraduates on Both Tasks

Group	Task						
	rec. task		prod. task		t-value	2-tail sig	correlation
	\bar{x}	s.d	\bar{x}	s.d			
Undergraduates	13.69	6.88	11.69	6.33	3.76	0.002	0.95
Graduates	21.50	8.48	15.44	7.85	4.73	0.000	0.81
All.	35.188	14.006	27.13	13.26	5.47	0.000	0.91

On another level, this paper is concerned with the characterization of the learners' strategies in producing the Arabic *break*-collocations in English. Compared to their errors on the production task, the subjects' errors on the recognition task provide little speculation about their sources or about the conceptualization the learners might have had when communicating in the TL as the former task would reflect their unrestricted intuitions about their communication. However, a look at the frequency of the deviant options on the recognition task (see Table 3) may help to support the analysis below.

4.2 Learners' Collocational Strategies

It is not precisely clear how the subjects made their choices. It is likely that they used various test-performing strategies in order to accomplish the tasks at hand. It could have been that they relied on their NL or made a conjecture, or lost patience or avoided rendering a response on purpose, especially when the task was rather difficult. They might have focused on extraneous factors - say whether the entire sentence was semantically anomalous (cf. Ellis, 1991:164-5). Besides, the testees may have given careless, rather irrelevant responses, sort of 'ad hoc forms' created merely to fill in a perceived gap in their IL vocabulary (Claus and Kasper, 1983). Briefly, one cannot claim a single distinctive factor for the subjects' choices: several stimuli may work simultaneously to motivate their options. In all cases, nonetheless, the subjects must have relied on their implicit and explicit, 'immature' knowledge when generating their communicative utterances in the TL. Stated otherwise, they must have drawn on their existing IL, or on what they assumed to be the TL norms.

Data analysis has displayed twelve distinct communicative strategies manipulated by the subjects when rendering into English the collocational sequences of the Arabic verb *kasara*. They are characterized as avoidance, literal translation, substitution, overgeneralization, quasi- morphological similarity, assumed synonymity, derivativeness, imitation of literary style, idiomaticalness, paraphrase and circumlocution, graphic ambiguity, and

finally, false IL assumptions. These will be briefly explicated and exemplified below.

4.2.1 Avoidance

It is a common observation of researchers that testees often refrain from carrying out a certain task on the grounds that it is perceived to be difficult or time-consuming or when they just feel no guts to be tested. In our data, it is not clear whether the subjects' avoidance was due to their IL level or lack of determination or concentration. As evident in Table 3, the subjects avoided giving response to the objective task 53 times out of 1120 possibilities, i.e. (4.73%) while they avoided responding to the production task 130 times (see Table 4) i.e. (11.61%) of the total attempts, with the undergraduates having done that more often than the graduates on the first task and less often on the second. To account for this high percentage of avoidance, the writers preponderate the hunch that the subjects were impatient and plausibly unwilling to continue a voluntary test for about an hour or so.

Table 3: Response Frequency on the Recognition Task

Group/Option	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Tot
Grad. (N=38)	A	24*	6	6	8	21	1	14	2	15	2	28	18	2	4	1	18
	B	24	1	7	20	8	-	14	2	8	3	2	8	-	6	31	3
	C	8	29	10	5	3	35	4	18	3	4	2	4	33	2	1	6
	D	4	1	14	3	5	1	5	15	9	26	1	5	1	23	1	6
	Av**	2	1	1	2	1	1	1	1	3	3	5	3	2	3	4	5
U.grad (N=32)	A	21	13	5	7	14	5	6	5	18	3	22	7	3	2	3	6
	B	3	3	8	15	10	1	13	2	6	4	4	10	4	9	26	11
	C	3	7	5	4	4	14	4	12	3	3	2	4	22	10	2	2
	D	4	8	12	5	3	12	7	12	5	21	4	10	2	10	1	11
	Av.	1	1	2	1	1	-	2	1	-	1	-	1	1	1	-	2
Total correct responses (N=70)	45	36	26	35	35	49	20	30	14	47	50	25	55	15	57	24	53

*The bold figure represents the frequency of the correct option.

**Av.= avoidance

4.2.2 Literal Translation

This strategy accounts for (7.59%) of the subjects' potential attempts. Research suggests unequivocal evidence for the role of the NL in the acquisition of the TL. Learners tend to manipulate their NL in their TL production whenever they do not have the necessary knowledge of the relevant TL form to be communicated. However, such manipulation may not necessarily lead to error. Contrary to the claims of the developmental 'creative constructive hypothesis' (cf. Dulay and Burt, 1973, and others), interference is a well-attested phenomenon at all levels of language proficiency (Kellerman, 1979:38). Apparently, not all NL forms will be susceptible to transfer; some features will, others will not. To be sure, transfer from the NL to the TL is motivated by three broad factors, viz., the learner's psychological structure, his perception of NL-TL distance, and his actual knowledge of the TL (ibid: p. 53). Kellerman maintains that interference is a product of the learner's cognitive system, and conceives of it as creative, operating at varying levels of consciousness and emanating from a decision-making mechanism. Kellerman suggests that the notion of markedness of a polysemous word can be a factor in interference. Homonyms or polysemous words display an array of multifarious collocational paradigms manifesting varying semantic functions. They may form a continuum of semantic differentiations departing from a core meaning to a metaphorical or idiomatic intention. For instance, *blue* basically refers to colour before *depression*, *jazz*, or pornographic material, but all these homonymous uses have reference to the core, unmarked use. It is along this line, that Kellerman ascertains that the intuitions of native speakers about 'semantic space' in NL can be utilized to predict their judgment of the translatability of morpho-syntactic forms from the NL to the TL. Stated differently, a polysemous word like Arabic *kasara* is more likely to be transferred when its meaning is closer to the core 'unmarked, neutral' meaning which is the most common one.

Table 4: Frequency of Learner Collocational Strategies

COLLOCATIONAL STRATEGIES

Strategy	Undergrad. (N=32)		Grad. (N=38)		All (N=70)	
	No.	%	No.	%	No.	%
Avoidance	70	13.67	60	9.87	130	11.61
Lit. translation	50	9.76	35	5.76	85	7.59
Substitution	20	3.91	12	1.97	32	2.86
Overgeneralizations	10	1.95	5	0.08	15	1.34
Quasi-morph. similarities	2	0.04	2	0.03	4	0.04
Assumed synonymy	4	0.08	17	2.80	21	1.88
Derivativeness	1	0.02	1	0.02	2	0.02
Verbosity	4	0.08	15	2.47	19	1.70
Idiomativeness	7	1.37	12	1.97	19	1.70
Paraphrase and circum.	110	21.48	65	10.69	175	15.63
Graphic Ambiguity	7	1.37	5	0.82	12	1.07
False IL assumptions	30	5.86	25	4.11	55	4.91

Furthermore, Kellerman proposes an influential role for the learner's perception of NL-TL distance in interference. In this respect, he hypothesizes:

In two perceptually closely-related languages, 'non-core' meanings may be transferred, while in two distantly-related languages perhaps only the 'core' meanings will be transferred. (p. 52).

The findings of the present study endorse, though partially, some of the aforementioned views on the role of the NL. Many of the *kasara*-collocations have equivalents in English. Surprisingly, however, many of our subjects failed to produce the equivalent forms, perhaps conceiving them as Arabic-specific. Examples of literal translation cited from both tasks are:

13. He *broke his opponent's nose*.
14. The *enemy's thorn* was crushed.
15. The *enemy broke* in the battle.
16. He *broke his thirst* with cold juice.
17. I *broke the dining table* a moment ago.
18. He *broke the fork* of the enemy.
19. He *broke the wall of silence*.

20. *Sun rays break in water*

A scrutiny of (13-20) clearly shows literal translation of *kasara* 'broke'. However, 'broke' in these utterances combines opaquely with the following word to denote meanings different from those intended by the original Arabic collocates, namely 'humiliated his opponent'; 'eliminated the power / was defeated'; 'quenched thirst'; 'had breakfast'; 'broke silence', and 'sun rays refract', respectively. In (13), the lexical sequence *kasara ?anf (xaSmih)* is viewed as a transparent collocation since *?anf* 'nose' is taken to be the symbol of pride and dignity in Arabic. Thus, the collocational sequence means 'broke his opponent's pride/dignity', i.e. 'humiliated his opponent'. Similarly in (14 and 15) some subjects translated the second element in *kusirat shawkat al.9aduww* by the word 'thorn' and 'fork', respectively--a synonym that is unlikely in the given English context though in Arabic the collocation provides a metaphorical sense to the effect that *shawkah* symbolizes 'power'. It seems, however, as though those subjects who opted for this literal translation were motivated by the collocant *shawkah* in isolation from its sequential context.

4.2.3 Substitution

Collocational meaning is conceived of as part of the lexical meaning; thus a deviant collocation turns the meaning anomalous and non-felicitous. Failing to produce the proper lexical item, the learners resort to using a substitute term that shares with it certain semantic properties. In the data, many learner incorrect substitutions are traceable to NL transfer or paraphrase and circumlocution. The data exhibit an appreciable ratio of the subjects' errors (2.86%) incurred by inappropriate substitution. Illustrative examples are:

21. He *lightened / lessened / crushed / diluted* the bitterness of the medicine.
22. The waves *shattered* on the sea shore.

23. He *violated / cut / interrupted / distracted the prevailing silence.*
24. He *cracked her heart.*
25. His death *cut down / cancelled the army's spirits.*
26. She *flouted / put down her husband's oath.*
27. The police *penetrated / violated the law when . . .*

Unaware of the correct English verbs which collocate with the following object nouns in (21 - 27), it is likely that the subjects resorted to their NL collocational norms and searched for an Arabic word which would roughly convey the intended message, then translated it into English. Some choices succeeded in conveying the message, some failed, and some added extra connotations. Thus, whereas sentences (21, 22 and 24) are honest renditions of the intended messages, they lack proper collocation. However, in (26), the substitute verb 'flout' colours the intended message (i.e. broke his oath). In (22), apparently, the subjects substituted *kasara* by a 'remote synonym', viz., *baddada* and, erroneously, used its English equivalent 'shattered' to collocate with 'the waves'. Likewise, the options in (23 and 27) are substitutes for the literal translation of the subjects' perception of the verb *kasara* in the given sequence. The subjects must have interpreted it in the original statements as *intahaka / qaTa9a / qa:Ta9a / baddada / ixtaraqa* - all sharing certain common semantic features of *kasara* 'broke'.

4.2.4 Overgeneralization and Analogy

Overgeneralization and analogy are communication strategies in which a certain TL feature or form is expanded to a different contextual use in the TL. It is thus characteristic of learner TL. This strategy forms a small ratio (1.34%) of the subjects' overall responses. Illustrative examples are:

28. The *waves refracted* on the shore rock.
29. The irrevocable *debts defeated* him, so. . .
30. He *wetted/extinguished his thirst* with cold juice.

31. The police *ashamed the law* when they arrested the MPs.

In (28), the subjects extended by false analogy the refraction of light to the breach of waves on the shore rock. Both meanings are inherent in the Arabic verb *kasara*. Likewise, in (29) the subjects extended the meaning of *kasara* in *inkasara al.9aduww* 'The enemy was defeated' to *kasarathu ad.diyu:n* 'The debts broke him', then substituted *defeated* for 'broke' by analogy. In (30), moreover, the respondents paraphrased *kasara 9aTashahu* simply as *balla ri:gahu* or *?aTafa?a Thama?ahu* and then extended the translation of this interpretation into English, viz *wetted/extinguished* his thirst. In (31), furthermore, the subjects seem to have confused the verb *shamed* and the adjective *ashamed* in their interpretation of *kasarat ash.shurTatu al.qa:nu:n* 'The police broke the law' and overgeneralized *ashamed* to replace *broke*.

4.2.5 Quasi-morphological Similarity

Words with quasi-morphological similarity could tempt learners to make deviant lexical choices in the TL. Duskova (1969) and other studies of learners of different language backgrounds have identified this strategy. However, it was attested on a very limited scale in the data of the present study (0.04%). Only two instances are cited in the data:

32. *The enemy was retreated in the battle.*

33. *Sun rays retracted / fractured on the shore rock.*

What is intended by *retreated* and *retracted / fractured* is *defeated* and *refracted*, respectively. In both instances, the subjects' errors seem to have risen from the apparent similarity of the morphological construction of these verbs.

4.2.6 Assumed Synonymity

Recourse to semantic affinity in the translation of a given collocation is a common lexical simplification strategy (c.f. Blum and Levenston, 1978). University FL students know many synonymous words, but not their complete range of synonymity in all contexts. In other words, they may not be fully conscious of the complete selectional restrictions imposed on the use of synonymous words. Their knowledge is restricted by the type and amount of instructional input they have received and by the impact of bilingual dictionaries which gloss some words as synonymous without much detailed contextual distinction. Thus, when short of the appropriate collocant, FL students look for a synonym or a near-synonym: the result is a non-idiomatic, non-felicitous IL utterance.

Assumed synonymity comprised only (1.88%) in our data, where the chosen words do not collocate well. Examples from the data are:

34. *The enemy was failed in the battle.*
35. *The irrevocable debts failed/defeated him,.....*
36. *Some workers interrupted / violated the strike.*
37. *He cracked / shattered her heart.*

Obviously, the subjects used the main verbs in the above sentences as synonyms for *defeated* (34) and *broke* in (35 - 37) respectively, on the assumption that they are semantically equivalent. Consequently, they yielded such bewildering or rather amusing utterances. Generally, false synonymity is the product of immature interlanguage.

4.2.7 Derivativeness

This strategy has been identified by Zughoul (1991) and is attested in the present study. Here, learners may take a TL adjective, for example, and derive a verb form from its NL counterpart and then overgeneralize this derived form to the TL form, or they may paraphrase the NL collocant word

and derive a verb form from its TL equivalent. This strategy, however, is cited only twice (0.02%), viz.,:

38. The irrevocable *debts caused him to broke*.
39. He *lighted the bitterness of the medicine*.

It is likely that in (38) the subjects took the Arabic adjective *muflis* 'penniless' to be the counterpart for the English adjective *broke*, then derived the infinitive verbal form from *muflis*, i.e. *yuflis* 'to cause to be penniless', then extended this by analogy to the English adjective *broke*; thus producing 'caused him to broke'. Similarly, in (39) the subject who used 'lighted' must have interpreted the target Arabic collocation idiomatically as 'lightened the bitterness'. He knows the equivalent Arabic adjective *xafi:f* 'light', he then derived the English past verb *lighted* from *light* 'not heavy' and used the derived form erroneously in the given context to replace the semantically correct, though collocationally wrong, form 'lightened'.

4.2.8 Imitation of literary style

Literary imitation refers to the effect of literary styles on the learners' choices. In this strategy, called verbosity by Zughoul (1991), the learners' errors may initiate from their recourse to their literary knowledge, from their tendency to select big sounding words under the impression that such words make their TL statement more impressive and literary-like (Zughoul, 1991:54). Moreover, some subjects, influenced by the Arabic literary styles, used certain English literary words to the effect that the outcome altered the intended, original message. This strategy was attested in (1.7%) of all the testees' responses. Illustrative examples are:

40. She *broke her husband's vow*.
41. She *flouted her husband's oath*.
42. He *picked at his food*.
43. When he spoke, he *shattered silence*.

COLLOCATIONAL STRATEGIES

44. He couldn't *redeem his debts*, so he declared himself bankrupt
(4)
45. He *vanished his thirst* with cold juice.
46. He *shunned her* and broke her heart.

Plausibly, the subjects opted for the bold-type words in (40-46) because of their literary ring, influenced by their literary readings which abound with such terms. The underlying conjecture here is that the subjects are not aware of the full range of the collocational distribution of these words, assuming they are precise counterparts of the target Arabic words. In (40-41) the respondents perhaps assumed full semantic equivalence between *yami:n* 'oath' and *vow*, and between *kasara* 'broke' and *flouted*. Apparently, there is some semantic overlapping between these pairs in the TL, but it does not warrant complete synonymity in the given contexts. Likewise, in (42-43), the verbs '*pecked at*' and '*shattered*' must have been taken by some subjects as a valid paraphrase of Arabic *kasara* in *kasara as.sufrah*, 'broke fasting' and in *kasara aS.Samt* 'broke silence'. It is likely that in (42) the subjects used '*pecked at*' metaphorically, meaning '*ate a little, like birds*'. Plausibly, those subjects interpreted the target Arabic form as 'just ate a little'- an implied meaning, though vaguely- and then extended this metaphorically into English. In the same manner, the subjects who used *shattered* (*baddada*) instead of *broke* seem to have been influenced by the former's literary ring although it does not meet the collocational constraint of the given context. The same conjecture, furthermore, could account for the learners' choice of *redeem*, *vanished* and *shunned* in (44-46), respectively, where these literary choices were thought to make the subjects' paraphrase of the target message more appealing and more literary-like.

4.2.9 Idiomaticalness

Idioms are not easy targets for translation as learners realize their marked systems within the NL and TL. Their meaning may not be decoded by

syntactic analysis. The strategy of idiomaticity is rooted in the learner's translation endeavour. Some of the target tasks in the present study border idioms. Unaware of the NL-TL equivalent expressions, our subjects tended to contrive idiomatic forms in the TL parallel to those of the NL. The resultant utterance is often deviant or anomalous. This tendency towards idiomaticity accounts for (1.7%) of the subjects' total responses. Examples of which are:

47. He *brought him down* to his knees.
48. He *pinned him down*.
49. He *brought down* his opponent.
50. The workers *turned down* the strike and returned to work.

It is quite evident from these examples that the subjects tended, unsuccessfully, to invent idiomatic expressions in English parallel to the target Arabic expressions. In (47-49), the subjects interpreted *kasara ?anfahu* idiomatically as 'humiliated him' and hence searched in their TL lexicon for an equivalent idiomatic expression. Nonetheless, they missed the intended message and produced anomalous utterances. Likewise, in (50) the subjects, inspired by their strive, for idiomaticity, contrived the collocation 'turn down the strike' in parallel with the Arabic idiomatic expression *kasara (x) al. ?iDra:b*.

4.2.10 Paraphrase and Circumlocution

When learners fail to translate idiomatically due to a deficiency in their lexical knowledge, they often resort to paraphrase and circumlocution to convey the intended message. Some paraphrase is economical, and some is circumlocutive and verbiage. More often than not, learners miss certain essential semantic aspects of the message or produce loose sentences which sound non-felicitous by the TL norms. Paraphrase was the most frequently employed in our data (15.63%)- almost twice as many by the undergraduates as by the graduates (see Table 4). Illustrative examples are:

COLLOCATIONAL STRATEGIES

51. She disobeyed her husband's order.
52. He stopped his thirst by juice.
53. He ate a minute ago.
54. He stopped the prevailing silence.
55. His death caused the army to lose their morality.
56. The irrevocable debts made him lose his money.
57. He made the taste of the medicine bitter.

4.2.11 Graphic Ambiguity.

As Table 4 indicates, the subjects' responses demonstrate a very small, though interesting, ratio (1.07%) of a special type of error which might have been initiated by the ambiguity of certain words as inspired by a biased reading of the Arabic script. Here, ambiguity might have arisen from the Arabic orthographical script. Such errors may be characterized as graphic. Examples.

58. I *cancelled the trip* before a while.
59. I *stopped travelling* before a while.
60. She *broke her husband's right arm*.

Some subjects seem to have read the second collocant *sufrah* meaning 'dining / dining-table' in the Arabic sentence (see 7 in appendix) as *safrāh* meaning 'trip / travel' and consequently produced (58-59 above). Sentence (60), moreover, manifests a partial reading of the intended Arabic collocation (see 1 in appendix). Apparently, some subjects read *yami:n* 'oath' which is intended by the Arabic sentence as 'right arm' --a far-fetched meaning implied in the Arabic core word *yami:n*, but the context makes this reading unlikely.

4.2.12 False IL Assumptions

Both undergraduates and graduates still make lexical and syntactic errors in the FL. As evident from the figures in Table 4, such errors are not likely to be persistent: graduates demonstrated far better progress on the same translation task (4.11% vs. 5.86%). One would hypothesize that such errors as (*The enemy defeated in the battle* or *she broken his oath*, or *the general is managed to take over the blockage and goes beyond the enemy*) were triggered out of carelessness had it not been the case that a good number of subjects iterated similar error types. Such lexical and syntactic errors which are characteristic of learner interlanguage could be ascribed to lack of learning at both receptive and productive levels - at school and at university. Quite a high ratio of the subjects' deviant collocations in the data (8.95%) may have stemmed from false IL assumptions about the meaning, form and distribution of certain lexical items in the TL, such as:

61. The enemy *defeated* (*was failed*) in the battle.
62. The debts made him *breaking*.
63. He *tolerated / breaked* the unsweet taste of the medicine
64. The workers *cut off / split* the strike and went back to work.
65. His death *lost* the army's spirits.

Conclusion

Arab learners of English, even at advanced levels of proficiency, still have difficulty with English collocational sequences. The competence of seventy undergraduate and graduate English major students was explored on 16 Arabic *kasara* - collocations. The students' overall proficiency in this linguistic area was found to be inadequate. The study aimed at probing any discrepancy in the learners' competence on the set tasks attributed to task type or to their academic levels. As expected, students' proficiency in the recognition, receptive task was significantly better than in the productive

COLLOCATIONAL STRATEGIES

task, and graduates surpassed the undergraduates quite remarkably on both the receptive and productive tasks.

The study also investigated the communicative collocational strategies employed by the learners when translating from the NL to the TL. Twelve such strategies have been identified, exemplified, and described. On the whole, the findings of the study have subscribed to the role of the NL in FL acquisition, suggesting that NL transfer is a creative cognitive process. And finally, the findings raise the need for a more constructive instructional focus on the phenomenon of collocation in English at both school and university levels.

Références

Al-Ani, Lamia. (1979). *Iraqi Students' Errors in Composition Writing*, Unpublished Manuscript, Iraqi Development of English Language Teaching Institute.

Arabski, Janus. (1979). *Errors as Indicators of the Development of Interlanguage*. University Slaski Katowice. In Arnaud, P. and H. Bejoint (eds.) (1992, *Vocabulary and Applied Linguistics*. Hampshire: Macmillan Academic and Professional.

Blum, S. and Levenston, E. (1978). "Universals of Lexical Simplification." *Language Learning*, 28, 2.

Carroll, J. B., et al. (1972). *The American Heritage Word Frequency Book* New York: American Heritage Publishing Company.

Claus, Faerch and Kasper, G. 1983. "On Identifying Communication Strategies in Interlanguage Production" in Faerch, C. and G. Kasper (eds) *Strategies in Interlanguage Communication*, CUP.

Cruse, D. A. (1986) *Lexical Semantics* Cambridge: CUP.

Dulay, H. & Burt, M. (1973). "Should We Teach Children Syntax?" *Language Learning* . (23), 235-252.

Duskova, L.(1969). "On Sources of Error in Foreign Language Learning," *IRAL*, 7, 1.

3. Ellis, Rod. (1991) "Grammaticality Judgements and Second Language Acquisition" *SSLA*, 13, 161-186.

Firth, J. R. (1951) "Modes of Meaning" In *Papers in Linguistics*, London: OUP

Fries, C. and Traver, A. (1940). *English Word Lists*. Washington: American Council on Education.

COLLOCATIONAL STRATEGIES

Fromkin and fromkin (1989). *The Study of Language* Cambridge

Ghadessy, M. (1976). "An Error Analysis of the Writings of Iranian Freshman Students Learning English: A Pilot Study," *IRAL*, 14, 1.

Hussein, Riyad F. (1998) "Collocations Re-visited" *Journal , of King Saud University: Languages and Translation*, 10, 39-46.

Ilsou, Robert. (1986). "Etymological Information: Can It Help Our Students?" *ELT Journal*, 37, 1

Judd, E. (1978) "Vocabulary Teaching and TESOL: A Need for Re-evaluation of Existing Assumption." *TESOL Quarterly*, 1, 71-76.

Kellerman, Eric. (1979). "Transfer and Non-Transfer: Where we are Now?" *SSLA*, 2, pp. 37-57.

Kjellemer, Goran. (1991) "A Mint of Phrases" In K. Aijmer and B. Altenbert (eds) (1991) *English Corpus Linguistics*.

Kucera, H. and W. N. Francis. (1967). *Computational analysis of Present Day American English*. Providence, R. I.: Brown University Press.

Leech, G. (1974). *Semantics*. Harmondsworth: Penguin Books.

Palmer, F.R. (1976). *Semantics: A New Outline*. London: Cambridge University Press, 1976.

Ramsey, Robert. (1981). "A Technique for Interlingual Lexico-Semantic Comparison: The Lexigram." *TESOL Quarterly*, 15, 1.

Seidle, J. and McMordie, W. 1988. *English Idioms* 5th ed. London: OUP

Shakir, A , and O. Shdeifat (1996). *The Translation of Collocations as an Indicator of Development of FL Competence*. *Al. Manarah* 1: 3- 9-27.

Thorndike, E. L. and I Lorge. (1944) *The Teacher's Book of 30,000 Words*. New York: Teachers' College, Columbia University.

Twaddell, Freeman. (1983). "Vocabulary Expansion in the TESOL Classroom," *TESOL Quarterly*, 7, 1.

West, Michael. (1959) *A General Service List of English Words*. London: William Clowes & Sons.

Verstraten, Linda. (1992) "Fixed Phrases in Monolingual Learners" *Dictionaries*" In Arnaud, P. and H. Bejoint (eds.) *Vocabulary and Applied Linguistics*. Hampshire: Macmillan Academic and Professional.

Wardell, David (1991) "Collocations: Teaching word pairs in EFL. Classes" *English Teaching Forum*, 29, 2.

Wilkins, D.A. (1972). *Nstional Syllabuses*. OUP.

Zughoul, M. R. (1991) "Lexical Choice: Towards Writing Problematic Word Lists", *IRAL*, 19, 1, 45-60.