COMPOUNDING IN ENGLISH AND ARABIC: 
A CONTRASTIVE STUDY

Prof. Mohammed Jasim Betti (Ph.D.)
Thi-Qar University / College of Education / Dept. of English

DOI: 10.37648/ijrssh.v10i03.001
Received: 15th May, 2020; Accepted: 08th June, 2020; Published: 18th June, 2020

1. INTRODUCTION

Linguists have shown interest in word-formation because it is one of the main elements of language. An example is the attempt to give a detailed description of Sanskrit word-formation (Bauer, 1983:2). Thus, phonological, Syntactic and semantic aspects of languages are studied, and nowadays, there has been an interest in seeing how word-formation reflects language behavior in general. One of such word-formation processes is compounding, which is the process of adding two or more bases together to form anew lexical item.

This paper seeks to shed light on Cg in Standard English and Standard Arabic. A contrastive attempt is made through describing CgE, then CgA, and it is then that a contrastive study of Cg in English and Arabic is attempted. The data of investigation consists of a hundred compound words (fifty ECs and fifty ACs).

2. CRITERIA

The six criteria coined in this study for investigating CgE and CgA are collected from a variety of references. These references include Eckersley &Eclarsly (1966: 21); Stageberg (1981: 121); Quirk &Greenbaum (1983:444); Richards, Platt and Weber (1985; 53-4) and Napoli (1996: 230 & 236); and Ibn 9aqiil (1965); Wright (1971); Hassan (1975); Nahr (1987); AI-Naajla (1988); Aziz (1989: 165) and Alkhuli (1999: 80). The criteria presented here are explained, but are exemplified in both languages in the analysis process. The criteria used for investigation are phonetic, syntactic, semantic, orthographical and historical:

A. The Phonetic Criteria

This criterion rests on the idea that Cs undergo certain phonetic and stress changes that they become different from the words they are made of.

B. The Syntactic Criteria

These criteria include;

a. Cg is based on adding two or more bases to form a new lexical item.

b. We cannot insert a unit between the two (or more) elements of the C since it behaves as one unit.

c. The elements of the C cannot be modified.

d. Cg can be endocentric in having the right member as a head or exocentric in having no head.
e. Cs can show different relationships and functions relevant to syntax, and can be classified according to the syntactic relation that holds between the parts of the C. This syntactic relation is language specific and it is the grammatical category of the head, which controls the whole C. However, there are different sentence patterns for different syntax types.

f. There is a sort of concord between the C and the other elements of the sentence.

g. Cg is related to some other word formation processes in that it works with some other morphological processes like inflection and derivation.

h. Cg is a generative word formation process in that so many words are invented by Cg.

i. Cs can be classified into N, Adj, V, Sub, Adv, P, Pro and Aux Cs with different patterns.

C. The Semantic Criterion

Cs have their meanings different from the meanings of their constituent elements.

D. The Orthographical Criterion

The C may be written as one word, and as separate or hyphenated words.

E. The Historical Criterion

Cg represents an element of the history of the language concerned since some Cs are changed throughout the history of that language into one word.

3.PROCESS

3.1. Compounding in English

As the data of the study proves, the EC word undergoes certain phonetic changes to the extent that it becomes different from the words of which it is structured. In addition, two stressed bases forming the C might be added together so that the second element is either unstressed or carries the secondary stress. If the first element of the C is an Adj, stress goes to the second syllable (or base) with secondary stress on the first, and if the first base is a N, it takes stress (Roach, 1988: 83). Some speakers, however, tend to stress the second element instead of the first. Bolinger (1975:112) and Roach (1988:83) support this idea:


2. Blackboard /ˈblaːkˌboʊd/3

3. Cowboy/ˈkoʊboʊ/3

Many grammarians agree that CgE is built on adding two or more bases together to form a new lexical item (criterion 2a). This can be testified throughout the data in this study:

4. childlike.

5. in spite of.
In addition, ECs behave as one unit in that if we try to insert a word between the two (or three) elements of the C, we will arrive at a starred C (2b), and we even cannot modify its elements (2c). This can be testified in the data:

6.a. loudspeaker  
6.b. loud nice speaker

7.a. loudspeaker  
7.b. a very loud speaker

Moreover, CgE can be classified as endocentric (2d). This is previously echoed by Bauer (1983:30) and Selkirk (1983; 91). Both types of endocentric Cs (the endocentric subordinate, which comprises a head at the right side working as the C, and the endocentric coordinate, which consists of two heads right and left) and the exocentric Cs do exist in English as clarified by the data of the study. This proves item (d) criterion (2):

8. living room (a room for living)
9. pathway (a path, which is a way or a way, which is a path)
10. a goldsmith (a smith who works in gold)

In addition, two other types of Cg do exist in English as proved by the data of this study: bahuvrihi and allomorph Cs. The bahuvrihi Cs are the Cs where the feature of a person or animal or thing refers to the referent as a whole and because the semantic head is unexpressed, it is seen as metaphorical or synecdochi. The allomorphs Cs are the ones where some root morphemes have bound allomorphs having a distinctive variant of a morpheme). Hartman & Stork (1972:10) refer to these types:

11. egghead (an intellectual and not a kind of head).
12. Franco-German

From a syntactic point of view, it is important to investigate whether there are relationships in the ECs that are relevant to syntax and whether the morphosyntactic category of the head percolates up to the word. A dike earthquake consists of N+ deberbal N, which may function as Sub + V, makes a N. There are many of such instances in the data. In considering the following two examples or some other similar instances in the data, it is the category of the head, which dominates the word:

13. head strong  N+Adj  (Adj)
14. globetrot  N+V  (V)

But at many other instances, the grammatical category of the C has nothing to do with those of its elements or it is dominated by that of the element preceding the head:

15. a. speakeasy  V + Adj  (N)
b. blackout  Adj + P  (N)
16.a. chin-up  N + P  (N)
b. underripe Adj+ N  (Adj)

ECs can be N, V, Adj, Pro, P, Adv and Aux Cs as the data shows. The syntactic analysis of CgE shows the following syntactic and grammatical relationship, which are gathered according to their types and are patterned as manifested in the data as follows:
One: NCs

These Cs can be endocentric (equative and additive) and exocentric;

a. N + deverbal N (Sub + V)  17. earthquake
b. V+N    (sub+V)  18. Hangman
c. Verbal N + N   (sub+V)        19. Washing machine
d. N + verbal N   (V + Obj)        20. brain washing
e. N + agentive or instrumental N (V + Obj) 21. Record-player
f. N + deverbal N   (N + Obj)      22. birth -control
g. N+N (V+Obj)    23. Knitwear
h. Verbal N + N   (V + Obj) 24. Spending money
j. N + verbal N   (v + Obj) 26. day dreaming.
k. N/Agentive N (Sub + Obj)   27-baby-sitter
l. N+N (Sub+Obj)  28. Windmill
m. Adj (modifier) + N (Adj+ N) 29. darkroom.
n.N+P (Sub + Adv.) 30. Chin -up
o. N + man       (Sub or Obj) 31. Fireman
p. N + V + er     (Sub or Obj) 32. Ball-player

NCs are generally pluralized by pluralizing the final element or both elements, or the noun qualified by a P phrase, Adv and Adj, or the N in the P phrase in popular speech:

33.a. schoolroom  b. school rooms
34.a. man servant b. men servants
35.a. father -in-law b. fathers-in-law
36.a. mother - in - law b. mother -in - laws

Two: Adj Cs

a. N + - ing particle  37. breath - taking
b. N + - ed participle  38. hand made
c. Adj./Adv. + - ing participle  39. easy -going
e. N + Adj   41. duty-free
f. Adj+Adj   42. deaf - mute
g. Adj+ N (P)  43. within
h. Adj + N + - ed.  44. grey - haired
i. cardinal + N  45. five - seat
This proves to us item (e) criterion (2) in that there are syntactic relations holding between, the elements of the Ecs. Nevertheless, it party achieves the idea that the grammatical category of the head determines that of the C since the EC category is determined by that of the head, or that of the pre-head elements, and it, at other times, has nothing to do with both.

The only types of concord occurring in ECs (and all grammatical categories) are Sub + V and Pro + V concord. Nevertheless, this occurs at the sentential level (f).

In English, Cg is sometimes linked to other word -formation processes. In the following two examples, it is linked to inflection and derivation:

62 -a- cowboy       cowboys
b- heat -sensitive ———^ heat -sensitiveness
This implies this inflection and derivation should follow the C. This proves item g (criterion 2).

The data also proves that CgE is a highly creative process. This result is also substantiated by Napoli (1996: 299) who invites any one of us, readers, to try to come up with a list of ECs having two or more, roots limiting himself/herself to the lexical categories (Ns, AdjS, Vs and Ps). In English, a C is formed by lexical categories rather than functional ones like determiners or conjunctions. Examples like 'nevertheless', 'in as mush as' and 'in so far as' are not called Cs, but lexicalized phrases (Ibid: 230). This proves item (6)

Semantically (criterion 2), the CE has a specialization of meaning in that the C means more than its parts. The Cs 'darkroom' and 'greenhouse' for instance, do not mean a dark room and a green house, but a room in which photographs are developed and a house in which plants are grown respectively.

Orthographically, the EC can be written as one word, or as separated or by phonated words. This is testified by the data of the study:
63, a. loudspeakers b. desk lamp c. fruit -cake

English has a great power in forming compounds from its native words. Some of these compounds are replaced by the French words as referred to in brackets (Criterion 6):
64. fore-elders (ancestors)

3.2. Compounding in Arabic
According to the phonetic criterion, the data proves that Cg is sometimes accompanied by phonetic and stress changes. In 65.a., the addition of /naj/ to /d3af/ results in /nad3af/. This process involves a deletion of /j /. This is also applicable /9abdI+ al-Hamiid/, which makes/9abdilHamiid/ in which /al/ is changed into /il/.

65.a./naj/+ /d3af/ /nad3af/: 'proper noun for a place'.
b./abd/+ alHamidi //abulHamiid/: 'proper noun',
In 66a, primary stress goes for the second element but in 66b, it goes to both elements as the data shows:
66.a. /bahaa '?'uddin/: 'proper noun'
b. / SabaHa'masaa? /: 'morning and evening'

Stress is phonemic in Arabic as this is proved by many studies including Gata (1988) in that stress placement change leads to meaning change. Nevertheless, in 66a and b, there is no change in the meaning of words after stress shift.

Syntactically, ACs are constructed by adding two (or more which rarely occurs) independent lexemes to form a new lexical item (2a). The data also proves that ACs work as one unit since they do not accept an intervening lexeme (2b) and even cannot accept an Adj or modifier of the first item (2c):
67. a. abd + allah = abdallah (proper noun)
b./surra/+/man/+ra ? aa==/surramanra?aa (proper noun)

68.a/-qawsquzah/ 'rainbow' */ qawsujamiilunquzah (an interning element)

Both endocentric and exocentric types of Cg are available in Arabic. In 69.a- /Suuq/:' market' is the head in /SuuqiSSujuux/:' proper noun ' (endocentric subordinate).
In 69.b. /kurdustaan/: 'The Kurd's region in the north of Iraq', there are two heads with no modifier/kurd/+staan/. In 69.c., /Sindibaad/: 'sindbaad ' has no head at all. The bahuvrihi type of exocentric compounds exist In 70 in which /?abbaTaJarran/: 'a proper noun' is a CN, which is seen as metaphorical since the semantic head is unexpressed :
69. a. /SuuqSSujuux/ b./kurdustaan/ c./Sindibaad/

70. /ta?abbaTaSarran/

This proves to us that the AC occurs at the right side. ACs can be divided in to /almurakkab al iTTaafi/: 'genitive compound' (a noun added to another),/ almurakkab almazd3i/: 'mixed C' (two Ns are mixed together) and /almurakkabalismaadi/: 'predication C 'subject and predicate in verbal and nominal sentences). Almurakkab al iTTaifi may consist of /muTTaaf/and /muTTaafilajh/: 'two Ns in the construct/ (See : Hassan ; 1975:23);

71 .a./nuurildiin/: 'proper noun ‘ b./muhammadali/: 'proper noun'
72 .a./faqiirulHaali/: 'poor'

AL-murakkab almazd3i consists of two mixed words with/out modification for easiness and abbreviation in forms :
73 .a. /ma9dikarb/: 'proper noun’ b./surramanra?aa/: 'place'

AL-murakkab al-isnaadi consists of a finite V +Sub (explicit or implicit) with/out an Obj, and Sub and predicate in nominal sentences /mubtada? waxabar/:
74. a. /jaziid/: 'proper N’ b. /surra man ra?aa/: 'place’ c. /Hamadallah/

The data proves for us that the genitive C, which consists of two Ns in the genitive case, depends on the right element of the C., If it is an Adj, the C is a CAdj, and if not it is a CN though it is very difficult at different situations to differentiate Adjs from NCs:
75. /qal9at SaaliH/: 'place'

Some Sub and predicate Cs in verbal sentences, nevertheless, have nothing do with their elements:
76./fjaziid/:'V + Sub' (proper N)
77. /surramanra?aa/: '/V + particle+ V/: 'proper N'
78. /ta?abbaTaSarran/: '/V+obj (Sub; implied) (proper N)

This proves that ACs carry the grammatical category of the head, or both or have nothing to do with both.

On the other hand, ACs as proved in the data can be N, V and PCs. This is the result of having only three grammatical categories in Arabic (Ns, Vs and Ps). All the other types are /tawaabi9/: 'followers of Ns'. That is why, there are three types of Cs. Accordingly, Adjs, Advs, Conjs, Pros, etc. are formed from Ns, Vs and Ps. In addition, particles in Arabic contain Ps, Conjs and interjections.

NCs are names of persons, places and times:
79. a./9abdulwaduud/ b./ba9labak/ c./SabaHamasaa?/
PCs consist of two Ps of a similar type: and they are of different types since there are many p types in Arabic (positive, negative, additive, etc.):

80.a. /Hajthu/+/maa/=/Hajthumaa/: 'wherever'

b. /kayfa/+/maa/=/kayfamaa/: 'however'

Vs of the sort of the following are Cs in that they consist of an Aux /kaana/: 'was' and a main V or two Vs. In this regard, particles in Arabic like /qad/: 'may' or 'do', and /s/ and /sawfa/ may express either probability or achievement and futurity respectively:

81./kaanajaktubu/: 'was writing'
82./kaanumazaalu ja19abuun/: 'they were still playing'

Adjs, though followers of Ns in Arabic do occur, and can be tested by using N test (plurality genitive, determiners, duality, pro, demonstrative and gender, see Alkhuli, 1999: 16-7). However, they are expressed in the genitive structure or other types of structure as the data proves:

83. /d3ajjid ilsuluuk/: 'good mannered'
   a. /d3ajjidulsuluuk/: 'plural'
   b. /kitabu d3ajjidilsuluuk/: 'genitive'
   c. /haTTa d3ajjidilsuluuk/: 'this is the good mannered'
   d. /d3ajjida alsuluuk/: 'duality'
   e. /huwa d3ajjidulsuluuk/: 'he is good mannered'
   f. /d3ajjidatul suluuk/: 'she is good mannered'

This proves to us that ACs have some syntactic relationships and consist of grammatical categories holding between the elements of the AC, and that there is no stable rule for considering the grammatical category of the C depending on its elements. Sometimes the AC category is as that of the head, at other times, as that of the pre-head. This verifies to us item (e) criterion (2). Aziz (1989:165) and Alkhuli (1999:16-7) certify this same result.

The data proves to us that ACs agree with their Vs and Adjs in number, person, gender and case (concord):

84. a. /d3aa?a sindibaad/: 'sindibaad came'
   b./d3aa?at quutilqulub/: 'quutilquluub came'

/t/ of femininity is added for concord. For the dual, the demonstrative is made as that of the dual (f):

85. a. /haTTasindibaad/: 'This is sindibaad'
   b. /haTTanisindibaadani/: 'These are two sindibads'.

In making the plurals of ACs, either the first element or the second one is pluralized:
ACg can be linked to other word-formation processes as the data shows. It is, for instance, linked to inflection and derivation. Inflection works on the head which is right word (g):

87. /9abdallah/: /9ibaadallah/: (inflection)
88. /qil9at Saalih/: /alqil9alSaa lihiya/: (derivation)

In addition ACg proves to be a very creative process since sometimes we are faced with ninety nine NCs depending upon the names of Allah. So, we put /9ab/ in front of all of these names to form a C. Furthermore, the best names in Arabic are said to be those that are coined from or start with mohammed or /9abid/.

Semantically, the two or more C words have meanings, which are different from the new C one. Hence, there is a specialization of meaning: /ta?abbaTaSarran/ consist of /Ta?abbaTa/: 'to put something under his shoulder and /Sarran/: 'evil' to become a proper noun denoting a significant item (signifier and signified). This is also applicable to:

89./quut/: 'food '+ /alquluub/: 'hearts' /quutilqluub/: 'proper noun '.

Orthographically, the data proves to us that ACs are written as one word or as two words only (criterion 4):

90- a./HaTHramawt/; 'proper noun for a place'
Moreover, spelling has no effect on the AC:

91.a. /qawsquzaah/: 'proper noun '

b. /aqwaasquzah/: 'rainbows '

Historically, no change in transforming ACs into single words is persecuted in the history of Arabic, but existed in Arabic since the very beginning in the Holy Quran, Arabic poetry and the Moslems' sayings.

4. Comparison

The analysis of the data using the coined criteria for Cg in English and Arabic makes the following explicit to us:

1. Stress change is one of the criteria to be used for describing Cg in English and Arabic. The most sound change that accompanies the formation of Cs (particularly NCs) in English and Arabic is the shift of stress to the first element of the C.

2. Syntactically, the elements of English and Arabic Cs cannot be modified in that we cannot insert any intervening element between any two of such elements. This differentiates Cs from phrases (2 a, b, c).

3. Both types of Cs (ie endocentric and exocentric) are available in English and Arabic. The head of the C is all the time the right word, which does not all the time control the whole C.

4. There are syntactic relations holding between the elements of the Cs in both languages, but the grammatical categories of the Cs are limited in Arabic to only N, V and PCs while nearly all grammatical categories are considered types of Cg in English.

5. Concord is available between the C and some other elements in the sentences because NCs agree with their Vs in English, and with their Vs and Adj's in number, pattemization and gender in Arabic.
6. There is no exact patternization of all the possible patterns of the types of Cs in both languages. This is due to the fact that patternization is a futile process of trying to put language, which is a means of communication into dogmatic rules.

7. Cg is related to some other processes particularly inflection and derivation, and it is a generative process since if creates hundreds of Cs in English and Arabic.

8. Semantically (3), the meaning of the C is different from the meaning of its elements in both English and Arabic.

9. Spelling is not, at all, a guideline for considering Cs since in both languages, the C can be written as one word, or as two words. In addition, it could be hyphenated in English, but not in Arabic. This is due to the fact that there is no hyphenation in the orthographical system of Arabic.

10. Historically, some ECs are changed to become single words. This does not occur in Arabic.

11. Whatever we try to prove the similarities of English and Arabic to prove the universal aspects of language, the differences do exist and play a vital role in our description of Cg.

BIBLIOGRAPHY


Appendix (1): A key to Abbreviations and symbols

a. Abbreviations

<table>
<thead>
<tr>
<th>Arabic Compound(s)</th>
<th>AC(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adj ective(s)</td>
<td>Adj(s)</td>
</tr>
<tr>
<td>Adverbs</td>
<td>Advs</td>
</tr>
<tr>
<td>Auxiliaries</td>
<td>Aux(s)</td>
</tr>
<tr>
<td>Compounding</td>
<td>Cg</td>
</tr>
<tr>
<td>Compounding in Arabic</td>
<td>CgA</td>
</tr>
<tr>
<td>Compounding in English</td>
<td>CgE</td>
</tr>
<tr>
<td>Compound(s)</td>
<td>C(s)</td>
</tr>
<tr>
<td>Conjunction(s)</td>
<td>Conj(s)</td>
</tr>
<tr>
<td>English Compound Adjectives</td>
<td>ECAdj(s)</td>
</tr>
<tr>
<td>English compound(s)</td>
<td>EC(s)</td>
</tr>
<tr>
<td>Nouns</td>
<td>N(s)</td>
</tr>
<tr>
<td>Object</td>
<td>Obj</td>
</tr>
<tr>
<td>Pronouns</td>
<td>Pro(s)</td>
</tr>
<tr>
<td>Prepositions</td>
<td>P(s)</td>
</tr>
<tr>
<td>Subject</td>
<td>Sub</td>
</tr>
<tr>
<td>Verb(s)</td>
<td>V(s)</td>
</tr>
</tbody>
</table>

b. Symbols

* ill- grammatical word, phrase or sentence

Appendix (2): The Data of the study

a. The Data of English compounds

Pathway, blackboard, childlike, goldsmith, Franco-German, chin-up, cowboy, in spite of, hangman, washing machine, newspaper, egghead, brain washing, record-player, loudspeaker, underripe, birth-control, knitwear, spending money, living room, swimming pool, fireman, ball-player, earthquake, baby-sitter, now that, mother-in-law, breath -taking, far - fetched, as far as, windmill, schoolroom, daydreaming, darkroom, man - servant, duty-free, deaf-mute, within, thirty-two, in that, father-in-law, handmade, stir - fry, hanger - on, grey - haired, dry - clean, easy - going, five - seat, two - thirds, lip-read, overlook, himself, bull's-eye, into, whatever, heat-sensitive, fore - elders, beeswax, headstrong, globetrot, speakeasy and blackout.

b. The Data of Arabic Compounds

/nad3af/: 'proper noun for a place', /abulHamiid/: 'proper noun', /bahaa 'uddin/: 'proper noun ', SabaHa'masaa?: / 'morning and evening', abdallah (proper noun), surramanra?aa (proper noun), /qawsquzah/ 'rainbow' , /SuuqjSSujuux/ (place) , /kurdustaan/ (place), /sindibaad/ (proper noun), /ta'?abbaTaSarran/ (proper noun), /nuurildiin/: 'proper noun ', /muhammadali/: 'proper noun', /faqiirulHaali/: 'poor', /ma9dikarb/: 'proper noun'
Appendix (3): The List of Symbols of the Arabic Segmentals

1. a. Arabic Consonants

/f/ /fill/ : 'an elephant'

/th/ /tha9lab/ : 'fox'

/Th /Thabhah/: 'he threw it'

/TH/ /THaabuT/: 'an officier'

/s/ /suug/: 'market'

/Ss/ /Ssabur/: 'patience'

/z/ /zraar/: 'button'

/t3/ /t3aali/: 'river bank'

/S/ /Smaalak/: 'What is wrong with you?'

/x/ /xubuz/: 'bread'

/G/ /Graab/: 'crow'

/h/ /hnaa/: 'here'

/H/ /Hariim/: 'women'

/b/ /baab/: 'door'

/t/timman/: 'rice'

/T/ /TamaaTa/: 'tomatoes'

/D/ /damiir/: 'conscience'
b. Arabic Vowels

/i/ /mi9d/ : 'stomach'
/i/ /bi/ : 'in us'
/a/ /saalim/ : 'safe'
/aa/ /saalim/ : 'safe'
/oo/ /tilifoon/ : 'telephone'
/u/ /ummii/ : 'my mother'
/uu/ /uuud/ : 'stick'

NOTES

1 See Appendix (1) for a list of the abbreviations and symbols used in this study.
2 See Appendix (2) for a list of the data of the study.
3 See Appendix (3) for a list of the symbols of Arabic Segmentals.
4 However, there are two points of view here. One states that ECs may consist of any part of speech (see: Eckersley & Eckersley, 1966: 21). Another ascertains that true Cs are limited to only lexical categories Ns, Vs, Adjs and Ps. The P is included because it is available in many CPs like 'love -in' and 'up -stage' (see: Napoli, 1996:230).