The Arabic Origins of Verb 'To Be' in English, German, and French: A Lexical Root Theory Approach

Zaidan Ali Jassem
Department of English Language and Translation
Qassim University
P.O.Box 6611, Buraidah, KSA
Tel. +966 6385 2338   Fax: +966 6322 0112  E-mail: zajassem@gmail.com

Received: 11-10- 2012                Accepted: 11-11- 2012                   Published: 25-11- 2012
doi:10.7575/ijalel.v.1n.7p.187          URL: http://dx.doi.org/10.7575/ijalel.v.1n.7p.187

Abstract
This paper examines the genetic relationship between verb to be forms in Arabic and English mainly as well as German, French, and Latin secondarily. It applies the principles and tools of the lexical root theory according to which be-forms are shown to be true cognates in having the same or similar forms and meanings with slight phonetic, morphological and semantic changes. Unlike traditional comparative historical linguistics views in which Arabic and English, for example, are classified as members of different language families, it shows how such verb forms are related to and derived from one another, where Arabic may be their end origin. For example, all s-based forms such as is/was in English, sein, ist in German, es, soi in French, sum in Latin are true cognates of Arabic kawan/kaan ‘be/was’ via common sound changes.

Keywords: Be-forms, Arabic, English, German, French, Latin, comparative historical linguistics, lexical root theory

1. Introduction
In comparative historical linguistics, English and Arabic are categorized as members of entirely different language families. The former is Germanic, which is affiliated to the Indo-European family which is split into five sub-families: viz., the Germanic family (e.g., English, German), the Italic (e.g., French, Italian), the Hellenic (e.g., Greek), the Slavic (e.g., Russian), and the Indic (e.g., Sanskrit, Kurdish, Persian). The latter is a member of the Semitic family, which is divided into several branches which include Arabic, Hebrew, Syriac, Aramaic, etc., with Arabic being the largest living language in the group (for a survey, see Crystal 2010: 308; Campbell 2006: 190-191; Crowley 1997: 22-25, 110-111; Pyles and Algeo 1993: 61-94; Ruhlen 1987, 1994).

However, Jassem (2012a, 2012b, 2012c, 2012d) disputed and rejected such classification where Arabic was shown to be more than genetically related to such languages. More precisely, in his (2012a) investigation of the numeral words from one to trillion (except for zero as loans are excluded here) in Arabic, English, German, French, Latin, Greek and Sanskrit, he found that all the numeral words have true Arabic cognates, considered to be their end origin. Jassem (2012b) provided further evidence by examining common religious terms like Hallelujah, God, Anno Domini, dominion, ruthless, welcome, worship, bead, solemnity, salutation, evolution, vigour, exacerbation, superiority, Christianity, Judaism, and so on, which were found to have true Arabic cognates. Interestingly enough, such expressions were presented in context in the form of phrases and sentences, every word of which had a true Arabic cognate. For example, hallelujah derives from a reversed and reduced form of the Arabic phrase la ilaha illa Allah ‘There's no god but Allah (God)' where Halle corresponds exactly to the Arabic word Allah in reverse- i.e., Allah → Halle (Halle 'God') (for further detail, see Jassem 2012b). In Jassem (2012c), personal pronouns, both independent and suffixed, in Arabic, English, German, French, Latin and related languages were shown to be true cognates, which descend from Arabic directly. Jassem (2012d) handled determiners such as the definite and indefinite articles and demonstrative pronouns in Arabic and English (and, in consequence, all European languages) which were found to be true cognates as well. This paper gives further evidence in this regard through the investigation of verb to be in Arabic, English, German, French, and so on to show not only their genetic relationship to each other but also their descent and/or derivation from Arabic cognates, which may be their end origin.
The theoretical framework in this study will be the lexical root theory, which has been proposed by Jassem (2012a, 2012b, 2012c, 2012d) to establish the genetic relationship between Arabic and English, in particular, and all other (Indo-)European languages in the field of the above-mentioned numeral words, common religious terms, personal pronouns, and determiners. It is so called because of using the lexical root of the word in examining genetic relationships between words such as the derivation of writer, underwriting, overwritten from write (or simply wrt) and kitaab 'writing', maktoba 'written', ikittaab/istiktaab 'subscription' from kataab (ktb) 'write'. Historically, this method was successfully and fruitfully used in the monumental works of the classical Arabic lexicological tradition such as Al-Khaleel and Ibn Manzour (for a survey, see Abdultawwab 1999).

The lexical root theory is simple in its outline, which comprises a construct or principle and four practical procedures for analyzing lexical roots. The theoretical principle states that Arabic and English as well as (Indo-)European languages of all branches are not only genetically related but also are directly descended from one language, which may be Arabic in the end. In fact, it claims in its strongest version that they are dialects of the same language. The applied procedures include (i) a lexicological procedure, (ii) a ‘tripartite’ linguistic procedure, (iii) a relational procedure, and (iv) a comparative historical procedure, all of which are described briefly below. At the outset, the lexicological procedure, which is dictionary-based, analyzes words by (i) deleting affixes (e.g., underwritten → write), (ii) using primarily consonantal roots (e.g., write → wrt), (iii) selecting semantic fields (verb to be in the present case), and (iv) search for correspondence in meaning. For instance, relating sixthly or sixty to their Arabic cognates must start with reducing them to the root six first, after which the search for related cognates begins on the basis of word etymologies and origins in standard works in the field such as Harper (2012) (for further detail, see Jassem 2012a).

Secondly, the linguistic procedure relates to the analysis of the phonetic, morphological, grammatical and semantic structure of lexical items or words. The main tenet of the phonetic analysis is that all ‘paired’ sounds may change within and across categories in all directions naturally and plausibly. Put more simply, consonants may change their place and manner of articulation as well as voicing. As to place, bilabial consonants → labio-dental → dental → alveolar → palatal → velar → uvular → pharyngeal → glottal (where → signals change in both directions); as for manner, stops → fricatives → affricates → nasals → laterals → approximants; and regarding voice, voiced consonants → voiceless. Likewise, vowels may change as well although they are marginal in significance and can be ignored in the analysis as shall be seen later. The basic vowels in this research are the three long vowels /aː (aa), iː (ee), uː (oo)/ and their short versions (besides the two diphthongs /ai (ay)/ and /au (aw)/ which are a kind of /iː/ and /uː/ respectively). All may change according to (i) tongue part (e.g., front → centre → back), (ii) tongue height (e.g., high → mid → low), (iii) length (e.g., long → short), and/or (iv) lip shape (e.g., round → unround). All these changes result in sound processes usually known as assimilation, dissimilation, deletion, merger, insertion, split, syllable loss, resyllabification, consonant cluster reduction or creation and so on. In addition, some changes may be more natural than others while others are plausible; for example, the change from /k/’, a voiceless velar stop, to /t/ch’, a voiceless palatal affri cate, is more natural than to /s/, a voiceless alveolar fricative, as the first two are closer by place and manner; the last is plausible.

Furthermore, sound change may proceed in three different courses (Jassem 2012a, 2012b, 2012c, 2012d). It may be multi-directional which means that a particular sound may change in different directions at the same time such as the different pronunciations of /θ/’, a voiceless interdental fricative, as in three in Arabic, English, French, Latin and so on (Jassem 1993, 1994a, 1994b, 2012a). It may be cyclic where more than one process may be involved in any given case such as the differences between the words for three in Arabic, English, German, French, etc. (see Jassem 2012a). Finally, it may be lexical or irregular where words may be affected by the change in different ways. That is, a particular sound change may (i) complete its course in some words, (ii) may vary in others, and (iii) may stall in some others. For example, in the different words for three in English, which derives from Arabic thalath ‘three’, (i) /θ/ has not changed in three, (ii) varies with /ð/ in third and /s/ in thrice, and (iii) changed to /t/ in trio-, trio, tertiary. This kind of change is known as lexical diffusion (see Jassem 1993, 1994a, 1994b, below).

Concerning the morphological and grammatical analyses, they overlap in many ways. The former concerns the inflectional and derivational aspects of the grammar such as the use of prefixes, suffixes, and infixes in general; the latter handles word classes and categories like nouns and verbs and their grammatical functions like subject and object. Because they do not influence the basic meaning of the lexical root, they may be totally ignored, therefore.
As to the semantic analysis, it focuses on meaning relationships between words, including lexical stability, multiplicity, convergence, divergence, shift, split, and change (Jassem 2012a, 2012b, 2012c, 2012d). Stability signals that word meanings have remained constant such as the numeral words for one-seven in Arabic and English (Jassem 2012a), basic religious terms (Jassem 2012b), most personal pronouns (Jassem 2012c), and most determiners (Jassem 2012d). Multiplicity indicates that words may have two or more meanings like fold as in ten-fold, folded paper (Jassem 2012a) and ship as in worship, warship, friendship (Jassem 2012b). Convergence means two or more formally and semantically similar Arabic words might have produced the same cognate in English such as the cognate words for thousand in English (Jassem 2012a); similarly, she derives from either Arabic ki 'you (f) in which /k/ became /sh/ or tha 'this' in which /th/ became /sh/ (Jassem 2012c). Divergence signals that words have become antonyms or opposites of one another such as nice in English and Arabic, which may derive from (i) na2s 'sinister' or (i) a reversed 2asan 'good' in both of which /2/ and /s/ merged into /s/, (iii) a reversed shain 'bad' or (iv) zain 'nice' where /sh & z/ merged into /s/. Shift indicates that words have switched their sense within the same field; for example, the numeral words eight and nine are reversed in Arabic, English, and all European languages. Lexical split means a word resulted in two different cognates such as Arabic hind(eed) '100' from which hundred and thousand stemmed (Jassem 2012a); Arabic Diya/Dau 'light' produced Deus, Zeus, deity, divine, day, etc. (Jassem 2012b). They, their(s), the, this and there split from tha 'this' (Jassem 2012c, 2012d). Change indicates a new meaning evolved like English four, French and Latin quatre, Russian chattiere, which all have true Arabic cognates. (For further detail, see Jassem 2012a, 2012b, 2012c, 2012d, and below.)

As regards the relational procedure, it examines the relationship between form and meaning from three perspectives (Jassem 2012a, 2012b, 2012c, 2012d). Firstly, words may be similar in form and meaning such as three, third, tertiary, thirty and Arabic thalath(ath) ‘three’, thuluth ‘third’ (pronounced talaat, tinka in Damascus Arabic) in which /th & l/ became /t & r/ each (see Jassem 2012a). Secondly, other words may be similar in form but different in meaning like eleven (elf in German) and elf ‘thousand’ in Arabic or ship and sheep (see Jassem 2012b). Finally, others may be different in form but similar in meaning such as quarter, quadrant and cadre (see Jassem 2012a). Generally all are accounted for in the lexical root theory.

Finally, the comparative historical procedure relates to the method of analyzing the genetic relationship between verb to be forms in language. It is comparative where every ‘be form’ in English in particular and German, French, and Latin in general will be phonetically, morphologically, and semantically compared with its Arabic cognates. It is historical in taking language development into consideration as all words may change, swap or reverse their forms and meanings within and across languages altogether. Overlooking word origin, history, and meaning may render the task impossible to tackle, indeed. The sources of such meanings are English etymological dictionaries and grammars (e.g., Harper 2012; Pyles and Algeo 1993) and Arabic dictionaries and grammars (e.g., Ibn Manzour 1974; Ibn Seedah 1996; Al-Ghalayeeni 2010) besides the author's knowledge of both Arabic as a mother tongue and English as a second language.

In the following analysis, all the above procedures will be utilized with different degrees of focus, though. The paper has five sections: section one is introductory, section two introduces the data, section three deals with data analysis and the results, section four provides a discussion, and section five is conclusion and recommendations.

2. The Data: The Verb To Be Forms

2.1 Be in English

In English, be has eight irregular, anomalous or unrelated variants depending on person and tense (e.g., Pyles and Algeo 1993: 126-127, 161): viz., am, is, are, was, were, been, and being. They may occur as ordinary or auxiliary verbs, e.g.,

\[
\begin{align*}
\text{I am/was a teacher.} & \quad \text{We are/were teachers.} \\
\text{You are/were a teacher.} & \quad \text{You are/were teachers.} \\
\text{He/She/It is/was a teacher.} & \quad \text{They are/were teachers.} \\
\text{He will be calling soon.} & \quad \text{She is going home now.} \\
\text{It has been raining all day.} & \quad \text{She is being called now.}
\end{align*}
\]

As an ordinary verb, be means ‘exist’ as in There is God for the universe. I am (in Syria, Britain).

In Old English (e.g., Pyles and Algeo 1993: 126-127, 161), bèon ‘to be’ had different forms as well. In the present
In the past tense, it was as follows:

- **eom/beo** '(I) am',
- **eart/bist** '(you-sg) are',
- **is/bid** '(he, she, it) is',
- **sind** '(we) are',
- **sind'**(you) are',
- **sind/bedd** '(they) are'.

In the past tense, the forms **wer** (with I, he, she, it), **wer(e/on)** (with you/we, you, they) were used (Pyles and Algeo 1993: 127); in the past participle, **gewesen** 'been' was used. Most of these forms continued into Middle English (Pyles and Algeo 1993: 161) with slightly different spellings and pronunciations, often chaotic and messy where some words had more than 500 spellings, some 60, some 20 (Viney 2008: 25; Pyles and Algeo 1993: Ch. 5-6).

### 2.2 Sein 'be' in German and Etre 'be' in French

Both words have different forms according to tense, person, and number. In German, sein 'be' has several forms in the present simple tense, which are:

- **Ich bin** 'I am'
- **du bist** 'you (guys) are'
- **Sie sind** 'you (formal) are'
- **er/sie/es ist** 'he/she/it is'
- **wir sind** 'we are'
- **ihre seid** 'you (guys) are'
- **Sie sind** 'you are'
- **sie sind** 'they are'.

In the past tense, **war** 'was' is used; in the imperative and subjunctive, **sei/seien** 'be'; in the past participle, **gewesen** 'been'. All have different conjugations except for the last (for a fuller picture, see german.about.com 2012).

In French, the verb **etre** 'to be' has the following six forms in the present simple tense:

- **je suis** 'I am'
- **tu es** 'you are'
- **il/elle est** 'he, she, it is'
- **nous sommes** 'we are'
- **vous êtes** 'you are'
- **ils/elles sont** 'they are'.

Different forms are used in the other tenses with different conjugations such as **serais, seras**, etc. in the future/conditional, **sois, soyes**, etc. in the imperative, and **etais, etait**, etc. in the imperfect (for a fuller picture, see french.about.com 2012).

### 2.3 Be in Five Indo-European Languages

**Be** has the same story in Indo-European languages. Baugh and Cable (1993: 18) give a comparative picture for it in Old English, Gothic, Latin, Greek and Sanskrit.

#### Verb To Be Forms in 5 Indo-European Languages

<table>
<thead>
<tr>
<th>Old English</th>
<th>Gothic</th>
<th>Latin</th>
<th>Greek</th>
<th>Sanskrit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eom 'am'</td>
<td>im</td>
<td>sum</td>
<td>eimi</td>
<td>asmi</td>
</tr>
<tr>
<td>Eart 'are'</td>
<td>is</td>
<td>es</td>
<td>ei</td>
<td>asi</td>
</tr>
<tr>
<td>Is 'is'</td>
<td>is</td>
<td>est</td>
<td>esti</td>
<td>asti</td>
</tr>
<tr>
<td>Sindon 'are'</td>
<td>sijum</td>
<td>sumus</td>
<td>esmen</td>
<td>smas</td>
</tr>
<tr>
<td>Sindon 'are'</td>
<td>sijuth</td>
<td>estis</td>
<td>este</td>
<td>stha</td>
</tr>
<tr>
<td>Sindon 'are'</td>
<td>sind</td>
<td>sunt</td>
<td>eisi</td>
<td>santi</td>
</tr>
</tbody>
</table>

Although the verbs look different intralingually and interlingually, there are many similarities among them. For example, in **am** forms, all share /m/ while two share /s/. In nearly all other instances, /s/ is common. Why? The answer will be attempted (in 3.) below.

### 2.4 Kaan(a) (Kawana) 'be' in Arabic

Kaan(a) 'was' comes from the main root kawana 'be', with the main consonant being /k/. In speech, kaan is variably spoken as /chaan (tshan), shaan, tsaan, or saan/ in many old and current Arabic dialects in the Arabian Peninsula or Gulf, Yemen, Iraq, Egypt, and Syria (for a survey, see Jassem 1993: Ch. 5, 1994a: Ch. 5, 1994b). It has different forms or conjugations according to person, number, gender, tense, and lexical category (noun, verb,
adjective). Its conjugations in the present and past tenses are as follows:

1. Present: Singular Plural
   \( a-k(oo/u)\)n 'I am' \( na-k(oo/u)\)n 'we are'
   \( ta-k(oo/u)\)n 'you are' \( ta-koon-oon \) 'you are'
   \( ya-koon/ta-koon \) 'he/she is' \( ya-koon-(oo/a)\)n 'they (m./f.) are'

2. Past:
   \( kun-tu \) 'I was' \( kun-na \) 'we were'
   \( kun-(a/i) \) 'you (m/f.) were' \( kun-tu(m/n) \) 'you (m./f.) were'
   \( kaan-a(t) \) 'he/she was' \( kaan-oo/kun-na \) 'they (m.f.) were'

3. Imperative:
   \( kun \) 'be (sg.)' \( koon-u \) 'be (pl.)'

As can be seen, pronominal prefixes are used in the present whereas suffixes are used in the past. Also in certain grammatical contexts, final /\( n/\) in \( kaan \) is deleted when preceded by certain negative past tense particles such as \( lam \) 'did not', for example,

\[
\begin{align*}
'&I-am = I am' &'did not' &'I-am = I was not'. \\
\end{align*}
\]

In Arabic grammar, \( kaan \) 'be' belongs to a special class of fourteen 'deficient, defective or incomplete' verbs, called \( kaana \) and its sisters or the \( kaana \) class. They may be classified into four subtypes (Al-Ghalayeeni 2010) as follows:

a) The first is specific time-restricted, including \( kaana \) 'to be/was', \( aSba2a \) 'of morning, became', \( aD2a \) 'of forenoon, became', \( umsa \) 'of evening, became', \( Dhalla \) 'remained', and \( baata \) 'of home, stayed'.

b) The second indicates continuous action, including \( ma zaala \), \( ma daama \), \( ma raama \), \( ma bari2a \), \( ma infakka \), \( ma wana \), \( ma fatia \). When preceded by \( ma \) 'not', all mean 'still, continue, last'; however, separately, each may have different senses such as \( fatia \) 'put out, calm down, forget', \( bari2a \) 'leave, go, vanish', etc.

c) The third is the verb \( Saara \) 'became', which indicates change.

d) Finally, the verb \( laisa \) 'not' indicates negation.

It has to be noted, though, that initial and final \( a- \) are not usually essential for meaning and may be dropped altogether in other derivations such as nouns, adjectives and adverbs.

They are termed 'deficient' verbs because they are usually followed by two nouns: the first in the nominative (with an \( u-\)suffix), the second in the accusative (with an \( an-\)suffix). The second noun is essential for message completion. For example,

\[
\begin{align*}
\text{kaana} & \quad \text{al-jaw-u} \quad \text{jameel-an.} \\
\text{was} & \quad \text{the-weather-nom.} \quad \text{beautiful-acc.} \\
\text{The weather was beautiful.'} \\
\end{align*}
\]

The word \( kaana \) may be replaced by any other in the group without changing the grammatical structure and/or, to a lesser extent, meaning. As in English, such a function can be termed a linking verb.

In addition, these verbs can occur as auxiliary and main verbs. The former occur before the main verb, e.g.,

\[
kaana \text{ yabkee. 'He was crying.'}
\]

\[
Saara \text{ yabkee. 'He started crying.'}
\]

Deleting the deficient verb here has no significant consequences to meaning. The latter can stand alone with their own independent meanings as in:

\[
\begin{align*}
\text{a-koon} & \quad \text{au} \quad \text{la} \quad \text{a-koon.} \\
'I-am' & \quad \text{or} \quad \text{not} \quad 'I-am = I am/exist or I am not/don't exist.' \\
\end{align*}
\]
3. Data Analysis: The Results

How can one establish the genetic links between the apparently unrelated 'be' forms or words within and among these languages, let alone Arabic? Where do these different forms come from? How do am, is/was, are/were relate to be? According to Pyle and Algeo (1993: 127), these verbs are badly mixed up and come from four different Indo-European roots. The English forms come, and sind(on) or sint, the Sanskrit forms *esmi, *esti, and *senti, and the Latin forms sum, est, and sunt are all from the root *es-. The English form eart is from the root *er- 'to arise'. The English forms with initial /b/ are from the root *bheu as in Sankrit bhavati (or bhava (Harper 2012)) 'become' and Latin fui 'have been'. The English past tense forms was/were are from the Old English infinitive verb wesan.

Although Pyle and Algeo's assertion that these verbs are unrelated and derive from different roots, their hypothetical (marked*) Indo-European origin is baseless which nobody exactly knows when, where and how it existed, if ever at all. How does eom relate to is, sind(on) or sint? Why is wesan treated separately from is, sindon or sint? Therefore, an alternative solution is called for which relates all such forms to one another intralingually and interlinguually, which lies in the Arabic origins and evolution of such verbs in English, German, French and the other European languages. More precisely, they all stem from the different forms of kaan and its sisters in Arabic. How?

i) Am (Old English eom, Greek eimi, Gothic im, Sanskrit asmi, Latin sum) derive from a reduced Arabic a-kun 'I am' via different sound changes: in the first three, /k/ turned into (i) /s/ via /ch, sh, or ts/, which merged with (ii) /n/ into /m/; in Sanskrit and Latin, /k & n/ turned into /s & m/ each besides a-loss in the latter; alternatively, Latin sum comes from Arabic kan/kun 'was/be' in which /k & n/ turned into /s & m/ each. Diagrammatically expressed, a-kun → a-(ch/ts)un → asun → asum (Sk.) (→ sum (Latin)) → am (English, Greek, & Gothic). An Arabic-based derivation of this kind neatly accounts for all the different first person singular be-forms in English, Gothic, Greek, Latin, and Sanskrit.

ii) Is (Sanskrit asi, Greek esti, Latin es(t), Gothic is, German is-, and French es(t)) all derive from Arabic 'present tense' ya-ku(n) 'he-is = he is' in which /k/ turned into /s/ via /ch/ while /n/ was dropped as happens in certain Arabic negative constructions (see 2.4 above). (A similar process affected as 'like' in English, which comes from a reversed Arabic ka- 'like' in which /k/ turned into /s/). Also the pronouns you and she in English and sie 'she, they', Sie 'you', and euch 'you-acc.' in German had an almost identical story (see Jasseen 2012c).

Old and Middle English sind, sint, and sindon are identical cognates of Arabic kun-t(a/i) 'were-you (m/f) = you were' and kun-tun 'were-you (pl. f.) = you were' in which /k/ turned into /s/ (and /t/ into /d/ (see viii below)).

German sein, sei, seien, etc., French present suis, sommes, son and imperative sois, soyes, soyons all derive from Arabic kaan/kun 'was/be' in which /k/ turned into /s/ via /ch/ and /n/ was dropped in some.

iii) Was is a short form of the Old English infinitive wesan 'be' and/or past participle gewesan (gewesen in German), both of which have Arabic cognates. Wesan is a reordered form of the Arabic root kawan(a) 'be' in which /k/ became /s/ via /ch, sh, or ts/; gewesan 'been' is a reduced and reordered form of Arabic present tense ya-koon 'he is' in which /y/ became /g/ (via /j/ perhaps). In Old and Middle English, the prefix ge- was used in forming past participles as in geholpen 'helped', gecéped 'kept' (Pyles and Algeo 1993: 120). In southern Middle English, however, the past participle prefix ge was replaced by or varied with ye-/i- as in y-ronne, geronnen 'run' (Baugh and Cable 1993: 413; Pyles and Algeo 1993: 123) and i-do 'done' (Baugh and Cable 1993: 407). Such a form is an identical cognate to what is usually known in Arabic as the five verbs as in ya-koon-oon 'they are' via reordering and grammatical shift from the present tense to the past participle. (Note that run is cognate to Arabic mar(ra) 'pass, go' via reversal and the passage of /m/ into /n/; similarly, y-ronne, geronnen 'run' and a reversed Arabic ya-murr-oon '(they) go, pass' are identical cognates where /m/ became /n/). In Modern English, ge- died out but survived into today's German, where it is very common. Formulaically written, kawan(a) → wakan → wachan → wasan (wesan); ya-koon → ya-wakan → ya-wachan → ya-wasan → gawasan (gewesan). That is the story, pure and simple.

Were is related to was in which /s/ turned into /r/ (Harper 2012). However, this might not be the case as shall be seen next.

iv) Are/were (art/wert in Old English) (German war, etc.; French future forms serais, seras, etc.) all derive from a reordered Arabic Saara 'become' (pronounced as Sora also in some Syrian Arabic accents), a related
kaan-class verb, in which /S/, an emphatic /s/, and /r/ merged. (Related conjugations such as a-Seer 'I become', Sir-tu 'became-I = I became' are also possible cognates.) The Arabic and French forms are obviously identical cognates. Diagrammatically expressed, Saara (serai in French) → aSra → are/were (and German war). This analysis receives further backing by art 'pictures and drawings' in English, which comes from Arabic Soorat, Suwar (pl.), Sawwar (v) 'picture' where /S & r/ merged into /t/r. As can be seen, both words are phonetically similar in Arabic and English but differ semantically.

v) Be (been, being) and all related Old and Middle English forms beon, beom, bion 'be, exist, come to be, happen' (Harper 2012) besides German bin '(I) am' and bist '(you) are' are traced back to Sanskrit bhava(t) 'become' and Latin fui 'have been' (Harper 2012; Pyles and Algeo 1993: 127). As such, their Arabic cognate is reordered aSaSa2a 'to become', to be' via different changes: in bhava, /S & 2/ became /h & v/ each; in fui, they merged with /b/ into /bV/; in be, /S, 2, & v/ merged into /bV/. As to inflectional /n/, it was a compulsory infinitive verb marker in Old English but still is in German; it has other morphological uses in English as well. Another likely verb is baata 'of home, to stay, remain' in which /t/ was deleted or changed to /n/. Finally, (ma) bari2a 'continue, still' is likely in which /t/ turned into /s/. Although pronunciation and meaning are almost identical in these cognates, the likeliest is the first. Schematically put, aSaSaSa2a → aSaSaSa → ahbava → ahbava → bhava (Sanskrit) → bawa (beo(n)beo(en) in Old/Modern English).

In short, all the different variants of be in English have been successfully and logically traced back to their Arabic cognates. Instead of treating them as irregularities, they are really different Arabic verbs in origin which function in grammatically the same way.

vi) The different forms of the French verb etre, which are obviously unrelated formally, have already been accounted for as being cognates of the Arabic kaan group. For reasons of clarity, here is a brief summary. The main verb etre 'be' consists of the root et- and the suffix –re, one of three such compulsory infinitive verb markers. As such, it derives from kaan 'be/he was' or ya-kun 'he is' in which /k/ changed to /t/ via /ch/ and /n/ into /t/. (Actually, inflectional /t/ in English, French, and German corresponds to inflectional /n/ in Arabic in many a similar case (Jassem MS).) Alternatively, as a whole, it may derive, though less likely, from aD2a 'become' in which /D/ became /t/ and /2/ was dropped. The present and imperative s-based forms like suis, es, sois, etc. come from Arabic ya-kaun/kun 'was/he is/be' in which /k/ turned into /s/ and /n/ was dropped; the future ones serais, sera, etc. come from Arabic Saar 'become' in which /S/, an emphatic /s/, turned into /t/; the imperfect forms etais, etait, etc. derive from etre directly, whose cognate again is kaan 'was' in which /k/ became /t/ and /n/ was dropped (or aD2a 'become', though less likely, in which /D/ became /t/ and /2/ was dropped). The present perfect fus, fui, etc. are reversed and shortened forms of (i) Arabic aSaSaSa2a 'become' in which /S & 2/ merged into /s/ and /b/ became /t/ or (ii) (ma) fatia 'continue' in which /t/ turned into /s/ or merged into /f/. In short, unlike the English and German forms, all the different French present and imperfect tense forms have one source.

vii) The German form sein and the Arabic form kaan 'to be' (pronounced kein in some 'Syrian' Arabic accents) are almost identical in both form and meaning where /k/ passed into /s/. I-initial forms like ist come from Arabic ya-kaun(n) 'he is' where /k & n/ became /s & O/ each, s-initial ones like sind from kaan, while bin and hist from aSaSa2a 'become', baata 'stay', or (ma) bari2a 'continue' as in English be above. In short, like English, the German forms come from two Arabic sources: kaan and aSaSa2a, baata, or bari2a.

viii) Morphologically, the different forms of be consist of one or two morphemes: one lexical or free and one inflectional or bound. The free morpheme is the stem, base, or root to which inflections and endings may be added. In English, be is the stem; in German, sein; in French, etre. The bound morpheme is a pronominal inflection or ending which stands for a short form of the equivalent personal pronoun such as –t/–d, –s, and –don in English, –t, –es, –ont, –ommen in French (2.1-2 above). All have Arabic cognates that can be easily gleaned from the present data (2.4 above). For example,

a) Old English –don, French –ont and Arabic –tun 'you (f. pl.) are cognates via the passage of /t/ into /d/ in one and reversal in the other;

b) –t, –d, and –s and Arabic –t you, she' are cognates where /t/ changed to /d/ or /s/;

c) French -ommen and Arabic –na 'we, us' are cognates where /n/ became /m/ while plural /s/ developed from /t/, an Arabic plural marker (for further details, see Jassem 2012c);

d) the infinitive (and past participle) verb marker /-n/ in German sein (gewesen) and Old English
the past participle in Arabic such as keena 'was' and Arabic ya- as in yakoona 'he is' are cognates where /y/ turned into /g/ via /j/.

c) finally, the past participle prefix ge- as in gewesen 'been' and Arabic ya- as in yakoona 'he is' are cognates where /y/ turned into /g/ via /j/.

4. Discussion

The above description and analysis of verb to be forms in English, German, French, Latin, Greek, Sanskrit, and Arabic has shown the adequacy of the lexical root theory for the analysis of their genetic relationships where all were found to be related. Shared vocabulary was 100% which, according to Cowley's (1997: 172-173) classification, means that they are dialects of the same language. As a consequence, the main principle that states that Arabic, English, and the others are not only genetically related but also are dialects of the same language holds true. For example, the different be forms and the suffixed pronouns or clitics were easily traced back to true Arabic cognates (also see Jassem 2012c). The minor differences between such verb forms are due to normal causes of phonetic, morphological and semantic change, especially lexical shift.

Thus, these findings are in harmony with Jassem's (2012a) description of numeral words in Arabic, English, German, French, Latin, Greek, and Sanskrit which were found to belong not only to the same family but also to be rather dialects of the same language. It also supports his investigation of common religious terms (Jassem 2012b), pronouns (Jassem 2012c), and determiners (Jassem 2012d) in such languages where the same patterns were replicated. In all, the percentage of shared vocabulary between Arabic and English, for instance, was 100%, which, according to Cowley's classification, means that they belong to the same language—i.e., dialects.

This being so, then one might ask why such languages are not mutually intelligible. Jassem (2012a, 2012b, 2012c, 2012d) discussed it at some length, to which this work lends further support. The main reasons for that were multidirectionality, cyclicity, and irregularity of sound change. Multidirectional sound change implies that, for example, the velar consonant /k/ in Arabic kaan 'be/he was' developed differently in different languages such as /s/ in English is/was and German ist 's/he/it is', and /t/ in French etre 'be', etais 's/he was', etc. (see 2.2 above) (for Arabic variants, see Jassem 1993, 1994a, 1994b). Cyclic sound change signals that was, e.g., underwent several sound changes in the course of its journey from its Arabic cognate kawan (kaan, ya-koon) 'be, ((he) was/is)' to its current state in English, including (a) reordering, (b) syllable deletion, (c) turning /k/ into /s/, and (d) /n/-deletion (see 2.1-4 above). Irregular or lexical change entails that words were affected by the change differently. For example, the different forms of Arabic kaun 'be/he was' underwent different sound changes as far as /k & n/ are concerned (see 2.1-4 above). That is, in English /k/ became /s/ in is/was and /t/ in were while /n/ was dropped. In French, /k/ became /t/ in etre, etais, etc. and /s/ in es, soi, etc. To this one can add lexical or semantic shift, a normal causes of phonetic, morphological and semantic change, especially lexical shift.

In related to the applied procedures of analysis, they operated neatly and smoothly. To start with, the lexicological procedure demonstrated that the lexical root was an adequate, analytic tool in relating their Arabic cognate is/was and related derivatives kaan/yakoon 'be' was/is' by isolating the root and ignoring the affixes –t and –don. The etymology or historical origin and meaning of lexical items was found indispensable also. For example, was/were came from the Old English infinitive form wesan 'be' and the past participle gewesen 'been' (Pyles and Algeo 1993: 127; Harper 2012), which varied with yewesent, whose Arabic cognate is a reordered kawan 'be' or ya-koon 'he is' (see 2.4 and 3. iii) above and below). Moreover, it showed the primacy of consonants and the marginality of vowels because the former are essential for meaning whereas the latter are rather phonetic and morphological in function. On the one hand, vowels link consonants to each other without which they would be impossible to pronounce; on the other, they signal grammatical categories such as tense (present and past), person (1st, 2nd, and 3rd), and so on. For example, the vowels in is and was in Modern English change to indicate tense while the consonants remain constant. The same happens in Arabic such as kaana 'he was' and kun-tu 'was-I = I was', etc. (see 2.4 above)

The phonetic analysis was extremely important in relating be forms to each other because Arabic consonants in particular underwent enormous changes in English and European languages as well as old and modern mainstream Arabic varieties themselves (e.g., Jassem 1993, 1994a, 1994b). The main sound changes that affected Arabic consonants here can be summed up as follows:

(a) /k/ in Arabic kaana 'be/he was' passed into (i) /t/ in all t-based forms in French etre, etais, etc. (ii) /s/ in all
Also there were suprasegmental changes like syllable deletion as in research (see 1. above).

Simpler, more predictable and so less significant than the consonantal ones, which are the primary focus of this section. (c) /t/ in inflectional Arabic –t(u/a/i) (see 2.2 above) and merged into /t/ in are;

(d) /D/, a voiced alveolar-pharyngealized (emphatic) stop, in aD2a 'became' passed into /t/ in French etais perhaps (see 2.2 above);

(e) /s/, a voiceless alveolar-pharyngealized (emphatic) fricative, in Saar 'became' evolved into (i) /s/ in French seras (see 2.2 above) and (ii) Ø in be (see 3.2.3 above);

(f) /2/, a voiceless pharyngeal fricative, in aSha2a 'became' developed into /v/ in bhava 'to be' and (ii) Ø in be (see 2.3.3 above);

(g) /Ø in Arabic kaan 'be/he was' passed into (i) /g/ in am, (ii) /r/ in French etre, (iii) was inserted as an infinitive and past participle verb marker as in Old English beon/wesan and Modern English been, and (iv) Ø in is, was, etc. (see 2.1-3 above);

(i) /b/ in Arabic aSha2a 'became' changed to /v/ in Latin (and French) fui (fus) (see 2.2-2.4 above).

All such sound changes affected place and manner of articulation and voice. That is, some consonants changed place, some manner, some voice while others changed two or all such features. For instance, the change from /k/ in kaan 'was' to /s/ in s-based forms like is, ist, es, etc. involved place (from velar to alveolar), manner (from stop to fricative), and voice (from voiceless to voiced) whereas its change to /t/ in t-based forms like French etais centred on place (from velar to alveolar). The change of /t/ to /d/ in English sind from Arabic –t 'you, she' centred on voice (from voiceless to voiced).

As to the vowels, all suffered different changes including fronting, backing, raising, lowering, centering, lengthening, shortening, diphthongization and smoothing. For example, the low central long vowel in Arabic kaan (also pronounced /ke:n, ke:n, ko:n/ in certain Syrian Arabic accents) 'he was' became a high back long rounded one in yakoon 'he is', a high back short rounded one in kunt (also pronounced kint) 'I was', etc. Likewise, English am, are, was, is, be, etc. have a similar story. In fact, vocalic changes are very much simpler, more predictable and so less significant than the consonantal ones, which are the primary focus of this research (see 1. above).

Also there were suprasegmental changes like syllable deletion as in was from Old English wesan 'be' or gewesen 'been' (see 2.1 above) whose Arabic cognates are a reordered kawan 'be' or yakoon 'he is' where /k/ became /s/.

The above changes resulted in different kinds of sound change like assimilation, dissimilation, deletion, insertion, reversal, reordering, merger, split, duplication, mutation, shift, and so on. In addition, the operation of all the above sound changes was multidirectional, cyclic and irregular or lexical (see 1 and this section above). All this entailed that the different forms of Arabic kaan in both classical and modern European languages such as is in English, ist 'is/he is' in German, es 'he is' in French, etc. are due to different courses of sound change. Jassem (2012a, 2012b, 2012c, 2012d) reported similar processes.

Morphologically and grammatically, although all such differences can be ignored altogether here without affecting the results of the final analysis adversely because affixes and word classes do not alter the basic meaning of the root itself, yet all inflectional and derivational morphemes (affixes) had Arabic cognates (see 3. viii) above). For example, Arabic ya-koon/ya-koonoon 'he is/they are' and English and German gewesen 'been' are identical cognates; the suffixed pronouns in all languages are identical cognates. However, the most important point was morphological and grammatical shift such as the shift from present tense Arabic ya-koon/ya-koonoon 'he is/they are' to the past participle gewesen in English and German (see 2.1-2, 2.1.4, and 3. viii) above).

Finally, semantically, certain lexical patterns were noted. Semantic stability was evident in all s-based forms
such as English is, was, were (French etre: est(i), suis, etc., German ist, sind, seid, etc.), which still retain the same or similar meanings as their Arabic cognates kaan, kunt, yakoon, etc. (2.4 above). Semantic shift was noted in gewesen 'been', whose meaning or function shifted from Arabic present tense yakoonoon 'he is/(they are)' to English and German past participle. French seraïs from Arabic past tense Saar 'became' shifted to the future tense. Lexical split took place in was/were from Old English wesan 'be'; also French etre split into s- and t-forms, both came from Arabic kawan 'be'. Lexical convergence occurred in be(on/m) which might derive from (i) Arabic aSba2a 'become' in which /S & 2/ merged into /v/ or (ii) ma hari2a 'continue' in which /r & 2/ became /v/ and Ø each, with the likeliest being the former, though. Lexical multiplicity was attested in all be-forms which may have two functions as auxiliaries and main or lexical verbs. Similarly, all the kaan-group members have two functions: one as auxiliary and one as main verb with its own meaning such as aSba2a 'it's morning time, become/be' (see 2.4a-d above); also ma-based forms in Arabic may have more than one meaning each (see 2.4 b above); for example, as compounds with ma 'not', they all mean 'continue, last'; separately, however, each has different senses. Lexical change was evident in the use of was/were in the past tense in Modern English which developed from the infinitive (wesan) and past participle in Old English and German (gewesen) from a reordered Arabic kawan 'be' in the end (see 2.2 and 3. iii above). Finally, lexical variability was manifested in the presence of variant or alternative words, which are utilized in different ways. For example, the words am/was, is/was, are/were, be/been/being in English vary formally due to their different Arabic kaana-cognates from which they came (see 2.1-2.4 above). Similar patterns were reported in Jassem (2012a, 2012b, 2012c, 2012d).

As regards the relational procedure or the relationship between form and meaning, all the cognates of the above be forms are similar in both form and meaning; i.e., true cognates. For example, Arabic kawan (kaan/yakoon) 'be (he was/is)', English is/was (wesan), German sein (sind, seid, sist, war), French etre (es, soi, etai, etc.) are all related in form and meaning, with Arabic being their main origin; some underwent lexical shift, however. French seraïs 'will be' and Arabic Saar 'became' are another instance. Some, however, are formally different but semantically similar such as was and were, both of which derive from Arabic kawan 'be' (see 2.4 and 3. iii above). Be (been, being) are another example, which derive from Arabic aSba2a 'be/became' (see 3. v) above). Some, however, are formally similar but semantically different such as as, is, Old English ge, you, German Sie/sie and euch, which derive from Arabic cognates with the same qualities (see 3. ii) above). Thus it can be seen that the formal similarities and/or differences between English words reflect those of their Arabic cognates.

In light of the above, therefore, all the foregoing verbs in Arabic, English, German, French, and related languages are true cognates in the sense of having similar forms and meanings. Arabic can be safely said to be where all words are genetically related.

As an aside, consider the following practical example You are a woman in English, French, German and Arabic where all words are genetically related.

\[
\begin{align*}
\text{Sie} & \quad \text{sad} & \quad \text{een} & \quad \text{Frau. (German)} \\
\text{Tu} & \quad \text{es} & \quad \text{une} & \quad \text{mere. (French)} \\
\text{You} & \quad \text{are} & \quad \text{a} & \quad \text{mother. (woman).} \\
\text{Thou (iyaki)} & \quad \text{kunti/Sirti} & \quad \text{--} & \quad \text{marat(t). (Arabic)} \\
\text{'you (nom./ac.)} & \quad \text{were} & \quad \text{--} & \quad \text{woman'}
\end{align*}
\]

Every single word has an identical Arabic cognate. First, the German pronoun Sie 'you (f)' and Arabic (iya)ki 'you-acc. (f.)' are identical cognates in which /k/ became /sh/ via /ch/; you (ge in Old and Middle English) comes from the same source also in which /k/ became /gh/, which later turned into /gh/ (via /j/ perhaps). The French and Old English pronouns are identical cognates of the short form of Arabic anta 'you- nom.:' i.e., -ta 'you' where /t/ changed to /th/ (Jassem 2012c). Secondly, the verbs sind 'are' and kunti (chinti, shinti, (t)sinti) are identical cognates in which /k/ turned into /s/ and /t/, a suffixed second person (feminine) singular pronoun, into /d/; French –es comes from the related derivative yakan in which /k/ became /sh/ and /n/ was lost (see 3. ii) above); and English are form Saar 'became' in which /S & r/ merged (see 3. iv) above). Thirdly, the indefinite articles eine, une, and a(n) and the Arabic ordinal numeral awvalool(a/e) 'first, one (m./f.)' are identical cognates in which /l/ turned into /n/ (Jassem 2012a). The nouns Frau and mere, mara (maraat, imraat) are identical cognates in the first of which /m/ became /f/; in mother, /t/ became /th/ besides reordering. To these can be added...
This paper is licensed under a Creative Commons Attribution 3.0 License.

5. Conclusion and Recommendations

The different be-forms in English, German, French, Latin, Greek, Sanskrit, and Arabic were found to be genetically related to one another. The main identical Arabic source cognate is kaana ‘he was’ from the root kawan ‘be’ in whose pronunciation /k/ may vary with /ch, sh, or ts/ while /n/ may be dropped in certain cases. The main points of this paper can be summed up as follows.

i) Am and all m-based forms such as Latin sum and Sanskrit asmi derive from Arabic a-kun ‘I am’ in which /n/ turned into /m/ and /k/ into /s/ which merged into /m/ later (3. i).

ii) Is and all s-based forms like German ist, French est(t) come from Arabic present tense ya-koon ’he is’ in which /k/ turned into /s/ into which /n/ merged or was dropped. The t-based French forms like etre, etais belong here in which /k & n/ passed into /t & r/ each. The s-initial forms like Old English sind(on), German sein, seid, and French suis, sois derive from Arabic past tense kaan ‘be/he was’ in which /k/ turned into /s/ (and /n/ was dropped). Was/were are short for Old English wesan/gewesen ‘be/been’ from a reordered Arabic kawan ‘be’ and ya-koon-on ‘they are’ where /k, y, & s/ became /s, g, & r/ in that order (while /n/ was dropped) (3. ii).

iii) Be and all b-initial forms in English and German bin, bist come from Arabic aSba2a(n) ‘became (they)’ in which /S & 2/ merged into /s/ and later into /b/ (while /n/ was dropped in some later) (3. iii).

iv) The French forms like seras come from Arabic Saara ‘became’ in which /S/, an emphatic /s/, turned into /s/ (3. iv and vi).

v) The suffixed pronouns like –t/-d, -s, and -ton and Arabic -t and –tun are identical cognates (3. viii); also the inflectional endings -(e)n, -ing’ in English and German and ‘-(i)r(e)’ in French stem from Arabic ‘-n’ which behaves similarly in making verbs and nouns (3. viii); Jassem MS).

To conclude, the lexical root theory has proved to be adequate for the analysis of the genetic relationship between verb to be forms in Arabic, English, German, French, Latin, Greek, Sanskrit, etc. where Arabic was found to be their origin, indeed. To further corroborate this finding, more and more research has already been and is still called for into all language levels (pronunciation, grammar, and especially vocabulary) and the application of such findings to grammar and language teaching, lexicology and lexicography, translation, cultural (including anthropological and historical) awareness and understanding (Jassem 2012a, 2012b, 2012c, 2012d). This research area is extremely interesting, immensely huge and limitless fertile that may be investigated linguistically and non-linguistically in every possible way; it is virgin indeed. Hopefully, its results will bring about boundless benefits to language teaching and intercultural relations- i.e., the promotion of cross-cultural understanding, tolerance and cooperation in espousing and fostering a culture of peace, security, stability, harmony, and unity in a turbulent world through positive diversity, of course. In the end, language learning and cultural adaptation may become a million times easier.

Acknowledgements

I wish to thank everyone who contributed directly or indirectly to this research, especially my colleagues and students in Syria, U.K., Malaysia, and KSA. My wife, Mandie Mufleh Ibrahim, deserves special thanks for her inspiration, understanding, enthusiasm, and tolerance.

References


(September 25, 2012)


(August 20, 2012)


_______. (1994a). Impact of the Arab-Israeli wars on language and social change in the Arab world: The case


_______. (2012a). The Arabic origins of numeral words in English and European languages. International
Journal of Linguistics 4 (3), 225-41. URL: http://dx.doi.org/10.5296/ijl.v4i3.1876

http://dx.doi.org/10.7575/ijalel.v.1n.6p.59


_______. (2012d). The Arabic origins of determiners in English and European languages: A lexical root theory
approach. Language in India 12 (11), 323-59.

Harcourt Brace Jovanovich.


University Press.