Title: Designing an Artificial Language: Arabic Morphology

Author: Rick Morneau

MS Date: 07-14-1994

FL Date: 01-01-2019

FL Number: FL-000058-00


Copyright: © 1994 Rick Morneau. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.

http://creativecommons.org/licenses/by-nc-nd/3.0/
Here's an idea for those among you looking for an unusual morphology for your artificial language (henceforth AL). How about basing it on the tri-consonantal system of Arabic (and of other Semitic languages such as Hebrew and Syriac)?

Rather than start with a formal description of Arabic morphology, let me give you a general definition and several simplified examples. Later, I'll discuss how all this stuff can be applied to an AL design.

Most Arabic morphemes are defined by three consonants, to which various affixes (prefixes, suffixes and infixes) can be attached to create a word. For example, the tri-consonant "ktb" represents the concept of writing. Here are some of the ways in which "ktb" is turned into real Arabic words:

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
<th>Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>kataba</td>
<td>to write</td>
<td>CaCaCa</td>
</tr>
<tr>
<td>?aktaba</td>
<td>to cause to write</td>
<td>?aCCaCa</td>
</tr>
<tr>
<td></td>
<td>(where '?' is a glottal stop)</td>
<td></td>
</tr>
<tr>
<td>kaatib</td>
<td>writing</td>
<td>CaaCiC</td>
</tr>
<tr>
<td>kitaab</td>
<td>a book</td>
<td>CiCaaC</td>
</tr>
<tr>
<td>kutub</td>
<td>books</td>
<td>CuCuC</td>
</tr>
<tr>
<td>kitaabah</td>
<td>writing profession</td>
<td>CiCaaCah</td>
</tr>
<tr>
<td>kattaab</td>
<td>author</td>
<td>CaCCaaC</td>
</tr>
<tr>
<td></td>
<td>(note doubled middle consonant)</td>
<td></td>
</tr>
<tr>
<td>miktaab</td>
<td>writing instrument</td>
<td>miCCaaC</td>
</tr>
</tbody>
</table>

In addition, verbs can be inflected to indicate person, number, gender and tense:

<table>
<thead>
<tr>
<th>Word</th>
<th>Subject</th>
<th>Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>kataba</td>
<td>he wrote</td>
<td>CaCaCa</td>
</tr>
<tr>
<td>katabna</td>
<td>we wrote</td>
<td>CaCaCna</td>
</tr>
<tr>
<td>katabuu</td>
<td>they wrote</td>
<td>CaCaCuu</td>
</tr>
<tr>
<td>yaktubu</td>
<td>he writes</td>
<td>yaCCuCu</td>
</tr>
</tbody>
</table>
naktubu - we write pattern: naCCuCu
yaktabuuna - they write pattern: yaCCaCuuna
sayaktubu - he will write pattern: sayaCCuCu
sanaktubu - we will write pattern: sanaCCuCu
sayaktabuuna - they will write pattern: sayaCCaCuuna

And there are many, many other derivations. Note, however, that all of them have "k", "t", and "b" in common (in exactly that order - a different order will indicate a different basic concept). Thus, the three consonants define the basic concept, while the affixes (i.e., patterns) define the way that the concept is applied. In actual Arabic, most of the affixes are applied in a regular and predictable way. Some, however, are irregular, with basic nouns and their plural forms being the biggest offenders.

The use of affixes inside a word, called "infixation", may seem strange at first, but even English has constructions that are reminiscent of infixation, such as "mouse/mice", "goose/geese", "take/took" and "sing/sang/sung". Note that these are not true examples of infixation, which does not exist in English, since they are irregular and unproductive. I use them only as an illustration. However, generalizing and regularizing the process could result in an extremely productive system.

Incidentally, if you're interested in just how productive this system is, consider that Arabic has approximately 31 consonants (the actual number depends on dialect and on who's counting). Thus, the number of possible tri-consonantal permutations is $31 \times 31 \times 31 = 29,791$. Not all of them are used, of course, but when you multiply that large number by the number of possible patterns (about four hundred), the result is an extremely rich and flexible lexicon, something for which Arabic is justifiably famous.

Now, if you want to create an AL with this type of morphology, all you've got to do is define a concept for each tri-consonant and for each affixation pattern. Here are some examples that I made up:

\[
\text{mbs} \Rightarrow \text{the concept of "beauty".}
\]
\[
\text{CaCoC} \Rightarrow \text{an adjective describing a state.}
\text{Thus, "mabos" = "beautiful".}
\]
\[
\text{CuCiC} \Rightarrow \text{a noun describing a state.}
\text{Thus, "mubis" = "beauty".}
\]
\[
\text{aCCaC} \Rightarrow \text{infinitive form of a transitive, causative verb.}
\text{Thus, "ambas" = "to make beautiful, to beautify".}
\text{Also, if you wish to inflect your verbs for person, number, gender and tense, you could, for example, use suffixes:}
\]

\[
\text{}}\]
"ambasufda" = "we beautify"
"ambasufti" = "we beautified"
"ambasufku" = "we will beautify"
"ambasizda" = "he beautifies"

and so on.

puCCaC => an infinitive describing a state.
Thus, "pumbas" = "to be beautiful".
Again, suffixes can be used for inflections:
"pumbasanda" = "she is beautiful"
"pumbasanti" = "she was beautiful"
"pumbasanku" = "she will be beautiful"
"pumbasulda" = "they are beautiful"

and so on.

diCCaC => an infinitive meaning "to become" a state.
Thus, "dimbas" = "to become beautiful".
Again, suffixes can be used for inflections:
"dimbasosda" = "it is becoming beautiful"
"dimbasosti" = "it became beautiful"
"dimbasosku" = "it will become beautiful"
"dimbasufda" = "we are becoming beautiful"

and so on.

To create compound words, convert one of the components to a pattern and apply it to the tri-consonant of the head word of the compound. As an example, let's create a compound for the word mansion from the roots for wealthy and house. Assume that the tri-consonant for wealth is "wlt", and the tri-consonant for house is "hns". Thus, using the patterns I listed above, the adjective wealthy would be "walot", and the noun house would be "hunis". Furthermore, we will create rules that convert a tri-consonant to a pattern for exclusive use in forming compounds. I won't elaborate on what these rules could be, but will simply provide an example. Let's say that the adjective/noun compounding pattern of "wlt" is "waltaCoCCi". Thus, we can create the word "waltahonsi" to mean mansion.

I leave the rest up to the reader's fertile imagination.

Please keep in mind that I'm not trying to convince anyone to actually design an AL with this kind of morphology. I'm simply providing an illustration of how things could be done differently. However, if I should ever come across an AL designed by a native speaker of Arabic, I would not be at all surprised if it had a morphology similar to what I've described here.

End of Essay
[Postscript: I'd like to thank Ken Beesley for correcting my English examples of infixation, and for providing me with a more accurate value for the number of infixation patterns in Arabic.]