

ARABIC COMPOUNDING AS A CASE OF ROOT CLASSIFICATION

Ladislav DROZDÍK

Institute of Oriental and African Studies, Slovak Academy of Sciences,
Klemensova 19, 813 64 Bratislava, Slovakia

The study tries to define the structural background of incompatibility of compounding with the root-and-pattern system of Arabic (Semitic). The most hazardous issue of the evidence collected consists in assigning Arabic compounds the status of a mere theoretical construct, deprived of actual word value, that can only be achieved at the level of monomorphemic root -words or at that of structurally fully assimilated lexical units related to them.

1. The close interaction of roots (R) and patterns (P) in the domain of inflection and word-formation, particularly in derivation, is one of the salient characteristics of the linguistic structure of Arabic (allowedly Semitic). The root-and-pattern system (henceforth, RP system) of Arabic, resulting from this close morphemic interaction which is, in turn, made possible by the incoherent nature of the Arabic root, is prohibitive to and overtly incompatible with the compound-generating processes currently operating in numerous IE languages.

The present inquiry starts from the assumption that, in Arabic, there are no exclusively consonantal roots beyond what we call the quadriliteral limit (q-limit), that is roots consisting of more than four root consonants (radicals). Quinqueliteral root-said structures (Fleisch 1968: 75; Grande 1963: 63) will invariably be relegated to the class of nonsegmentable root-words. The q-limit, marking the highest point of morphemic segmentability along structural lines of the RP system, indicates at the same time that of internal inflection and derivation. In delimiting the interval of segmentability at its minimum pole, a number of controversial approaches has to be taken into account. While D. Cohen (1963: 174, n.1), for instance, does not hesitate to speak about what he calls *racine monolitère*, as in the Ḥassānīya Arabic *ḥ* in *ḥət* “sister”, Father Fleisch moves the dividing line between what is segmentable and what is not considerably higher when qualifying as root-words even nouns derived from biradical roots, such as *yad*, *dam*, or the like (1961: 248, 252; 1968: 32: *mots qui sont à eux-mêmes leur racine*). In accordance with the most currently accepted classifica-

tion, compatible with the basic assumptions of the present paper, the smallest segmentable units will be identified with biradical structures. Within this interval of morphemic segmentability, delimited by biliterals, at the minimum pole, and quadrilaterals, at its maximum, lies the field of trilateral roots most widely represented in Semitic languages.

Since the interval of segmentability is, simultaneously, that of internal inflection and derivation, all morphological structures exceeding the q-limit are deprived of their ability for intra-root pattern-marked (henceforth, intra-root or, simply, internal) procedures. The formal and functional characteristics of the RP system, delimited by the interval of segmentability, stretch over one and only one word, and over each word separately, as a mark of word integrity. Compound units, nouns or verbs, consisting of several words, represented by their roots, within one word (Greenberg 1960: 178-194), are overtly incompatible with the constraints imposed by the RP system. The very existence of the RP system, underlying the morphology of Semitic languages is, then, sufficient to account for the awkward status of compounding, or what is presented as such, as a word-formational procedure. In opposition to the linearly structured IE root, clearly separated from accidental thematic formatives that have no access to the space it occupies, the Arabic consonantal root, as a structural and semantic abstraction, is in constant intra-root interaction with the stem-formative pattern morpheme.

Another assumption, associated with the previous one, assigns the word one and only one root, as part of the RP thematic complex. Any instance of asymmetry between the number of roots (RPs) and that of words will be treated as a theoretical construct, marked by hyphenated writing, as distinct from real words, appearing in the form of nonsegmentable root-words, graphically represented by unhyphenated writing. Unless subject to quadrilateral restructuring (as most root-words of compound origin consist of more than four consonants), root-words of the latter background represent the final stage of the compound-generating process.

No attention will be paid, in this context, to historical aspects of the root evolution, notably to the time-honoured theories trying to prove the biliteral origin of *all* trilateral roots (Mayer-Lambert 1897: 354, in: Brockelmann 1908: 285). Neither will the notion of *base*, adopted by some Semitologists, be found relevant to the matter.¹ The base, as a morphemic segment incorporating a con-

¹Brockelmann's preference accorded to the concept of base over that of root is unambiguous: "... der Begriff der Wurzel /ist/ für die Formenlehre unbrauchbar . . . Die Analyse der nominalen und der verbalen Ausdrucksmittel führt uns schließlich auf gewisse einfache Grundformen, die wir nach dem Vorgang der Indogermanisten /Basen' nennen. Wie im Idg. solche Basen (wie *pede* 'Fuß', *eie* 'gehn') an sich sowohl als Nomina wie als Verba fungieren können" (1908: 287).

Grande's *baza*, such as *ktab*, for instance, is a stem of the 1st order that, at the thematic level, cannot be further analysed into smaller units. (1963: 10)

In modern descriptions, *ktab*-like segments are regarded as smallest thematic units mostly called *simple stems* (e.g. Georgetown University dialectological series).

sonantal root and a vocalic pattern, created under the influence of Indo-European linguistics, was in fact minimum root-substituting thematic unit co-extensive with the bimorphemic RP segment. Kuryłowicz's root-said units of the type *qtul*, *qtil*, *qtal* evidently follow the IE type of a linearly structured root (1961: 174).

1.1. The basic difference between compounds in IE languages and those in Arabic (Semitic) consists in the structure of the morphemic segments subject to compounding. While, in IE languages, these segments are represented by roots (R), possibly operating as words (W),² in Arabic (Semitic), they are represented by roots integrated in stems (S),³ possibly identified with words, too:

English: *mile + stone > milestone*;
R/W R/W

German: *Auto+Bahn > Autobahn* "highway, autobahn";
R/W R/W

As against Arabic, the number of constituents is here limited only by pragmatic factors:

Auto+Bahn+Brücke > Autobahnbrücke "highway bridge";

Auto+Bahn+Kontrolle /allomorph: *kontroll/+ Punkt > Autobahnkontrollpunkt* "highway checking point", etc.

Arabic: ^ʿ*irq + sūs > ʿirq-sūs, ʿirqsūs* (see below);
S/W S/W

Since the Arabic root is inherent in the RP complex which is the minimum thematic unit, there is no need to restate Greenberg's criterial formula R>W, with the understanding, however, that R is a part of RP.

1.2. Let us examine some compound constructions over several procedural steps:

(a) input (i)(unorganized): *ra's* 'head' + *māl* 'wealth, money' (**m-w-l*): five root consonants (radicals) + one consonant-substituting long vowel: summarily, six radicals; alternatively:

input (ii): syntactic structure: *ra's (al-) māl* 'capital', plur. *ru'ūs al- amwāl*;⁴

(b) output (i): compound (one-word-two-root/stem) construction: *ra's-māl* 'capital';

one-word structure (distributional evidence): *ar-ra's-māl, hādā r-ra's-māl; ra's-māl-ī, -īya*.

² subject to constraints imposed by their status of compound constituents; the same holds for Arabic;

³ though, for the most part, S is represented by a root-and-pattern (RP) segment, viz. simple stem or the archaic *base*, in some single cases, S may equal a complex stem (RPA), by involving affixation (A), as in *maqām* in *qā'im-maqām*: *ma-* (prefix) *qām* /**qwam* /, (see in the text).

⁴ In spite of the assumption that most compound and compound-said constructions evolved from or are outright identical with a syntactic pattern, the latter will be quoted only if really recorded in textual or lexicographic sources.

Structural properties:

compound (one-word-two-root/stem) construction: three + three radicals; conflicting opposition between one-wordness and two-rootness (or rather two-stemness) persists; internal procedures in both inflection and derivation are inactivated (for particulars see below);

output (ii) reinterpreted:

In virtue of the premises adopted for the present inquiry, morphological structures extending beyond the q-limit can no longer be analysed into morphemic constituents and have to be treated as monomorphemic units (root-words): *ra'smāl* 'id.': five consonants (no longer radicals) + one consonant-substituting long vowel.

The latter categoric postulate is further supported by the fact that a word (here compound word) cannot involve components whose word-integrity is confirmed not merely at the root level (R), as in IE languages, but at the hierarchically higher thematic level (RP).

Structural properties:

noncompound (root-word), unsegmentable: no intra-root procedures are possible; the monomorphemic nature of *ra'smāl* is attested by the lack of a structurally compatible plural that had to be borrowed from *rasmāl*, viz., *rasāmīl*, as indicated in (c) below. Obviously, the difference between *ra's-māl* and *ra'smāl* is merely descriptive and interpretative.

Summary:

Relying on the assumptions adopted, lexical units of the (b-i) type, that is compound words (*ra's-māl*), represent mere theoretical constructs which make it possible to describe and specify structural characteristics of the resultant (b-ii) type root-words (*ra'smāl*). The undue asymmetry between the number of words and that of roots (RPs), still persisting in (b-i) units, is here eliminated.

- (c) Quadrilateralized, structurally assimilated, one-word-one-root unit: *rasmāl*; the quadrilateral root /*r-s-m-l*/ was obtained by deletion of the glottal stop /ʔ/ of *ra's* and the long vowel /*ā*/ of *māl* (**m-w-l*) and, with the inserted (Holes / 1995: 81/ speaks about 'interdigitated') stem-constituting vowel pattern /*-a-∅-ā-*/ the procedure has been completed.

As in *barr-mā'-ī* > *barmā'ī* below, quadrilateral reconstruction is supported by a parallel process of consonant cluster-prevention /*'-s + m/* > /*s + m/* : *ra's-māl* > *ra'smāl* > *rasmāl*; the allomorphic root (*r-s*), resulting from this reduction, is paralleled by the requalification of the long-vowel substitute for the etymological weak radical in */*m-w-l/* > /*m-ā-l/*, as an element of the pattern morpheme.

Structural properties:

noncompound, segmentable (quadrilateral) noun, fully operative in internal procedures:

inflection: *rasmāl*, plur. *rasāmīl*;

derivation: *rasmāl* "to capitalize", *rasmala* "capitalization", *murasmīl* "capitalizing", *tarasmāl* "to be capitalized", etc.

Similarly:

(It should be noted that not all these structural states are necessarily present with each particular compound(-said) item; in accordance with our convention, the (b-i) examples are written with a hyphen in contrast to the unhyphenated writing of their (b-ii) counterparts; for the sake of economy, unless necessary for descriptive purposes, the full-scale representation of the (b)-featured examples will be reduced to the (b-ii)-type; derivational affixes in final compound (b-i) or root-word (b-ii) forms will not be hyphenated save for a few illustrative examples).

- (a) *ʿirq* 'root' + *sūs* 'licorice (bot.)';
(b) *ʿirqsūs* "licorice root": (*al-*) *ʿirqsūs*; *ʿirqsūs-ī* "relat. adj. to *ʿirqsūs*"; cf. E.W. Lane's record of the Cairo Arabic *ʿir* 'isūs (*erq-soos*) "id." or *ʿir*' *sūs-ī* (*erq-soosee*) "producer and seller of a sort of sweet drink made from licorice root" (Lane 1860/1954: 331; on the other hand, the compound nature of *ʿir*' *sūs* is not explicitly marked in Badawi-Hinds 1986, as against, say, *bani ʿādam*, fem. *bani ʿadma*, plur. *bani ʿadmīn* "person, human being"; epenthetic vowels are written in superscript);
(a) *qā'im* 'holding' + *maqām* 'position, rank';
(b) *qā'immaqām*, *qā'imaqām* "district president"; *qā'immaqāmīya*, *qā'imaqāmīya* "administrative district headed by a *qā'immaqām*" (see note 3);
(a) *barr-ī* relative to mainland /*barr*/, as opp. to sea, water' + *mā'-ī* 'pertaining to water /*mā'*/; or syntactic sequence of two coordinated attributes: *barr-ī mā'-ī* "amphibious", as in *sayyāra barrīya mā'īya* "amphibious vehicle";
(b-i) *barr* + *mā'-ī* > *bar-mā'ī*, allomorph /*b-r*/ represents here the root /*b-r-r*/; the impact of the q-limit is here intensified by that of consonant cluster-prevention /*r-r+m*/ > /*r+m*/;
(b-ii) *barmā'ī* "amphibious", as in *dabbāba barmā'īya* "amphibious tank"; *barmā'īya* "amphibian" (Wehr);⁵
(a) *šibh* "similar to, quasi-, semi-" + *ġirā'ī* "glutinous, viscous"; backed by the annexion-type syntactic pattern: *šibh(w/i/a) ġirā'ī(yin)* "colloidal";
(b) *šib-ġirā'ī* (Monteil 1960: 131-152) > *šibġirā'ī* "id.";
(a) *ʿarḍ* 'presentation, demonstration' + *ḥāl* 'condition, state'; syntactic *ʿarḍ ḥāl* "application, petition";

⁵ The structural status of inflectional and/or derivational affixes (e.g. the *nisba*-suffix *-ī*, *-īya*, or their inflected variants *-īyūn*, *-īyāt*, or the *nisba*-abstract suffix *-īya*) is not quite easy to describe. It seems that there are two distinct types of junctures between root-words and affixes appended to them: in units like *ra'smāl-ī*, the *nisba*-formative *-ī* is joined to already existing well-established root-word in contrast to, say, *aš-šarq'awsaḥ-ī*, derived from *aš-šarq al-awsaḥ* whose root-word nature is due to the affix appended. The same holds for *-āt* in *māġarayāt* and similar constructions. While, in the former case, the affix maintains its structural autonomy, in the latter case it is not quite clear whether the affix has to be integrated in the root-word or classified apart (see 1.3.2). As for writing, apart from a few illustrative examples, affixes will be joined to root-words unhyphenated.

- (b-i) *ʿarḍ-u-ḥāl*, the cluster-preventive elision is here replaced by the insertion of an epenthetic vowel;
 (b-ii) *ʿarḍuḥāl* “id.”; the affix-marked plural *ʿarḍuḥālāt* confirms its inability to intra-root inflection.

- (a) *lā* (negative particle)⁶ + *šūʿūr* ‘consciousness, awareness’;
 (b) (*al-*)*lāšūʿūr* “the unconscious, unconsciousness”, also adj. *lāšūʿūrī* “unconscious, unaware”; as against other compound types, the present one is extraordinarily productive:, (quoted in their compound shape):

lā-ʾahlāqī “amoral”;
lā-dīnī “antireligious, irreligious”; *lā-dīniya* “irrelegion, godlessness”;
lā-sāmī “anti-Semitic; anti-Semite”; *lā-sāmīya* “anti-Semitism”;
lā-silkī “wireless; radio”; *ʾiṣāra lā-silkīya* “radio message”;
lā-nihāʾī “infinite”; *lā-nihāʾīya* “infinity”, etc. (Wehr)

Nevertheless, in accordance with the requalification proposed and the writing convention adopted, the examples quoted have to be represented as *lāʾahlāqī*, *lādīnī*, and so on.

- (a) *lā* (see above) + *šayʾ* ‘thing’; synt. *lā šayʾa* “(there is) nothing”;
 (b-i) *lā-šayʾ* “nothing, nonentity”; *lā-šayʾīya* “nonexistence, nothingness” (Wehr); in view of one-consonant structure of the first constituent (*lā*), no further reduction was needed to meet the q-limit.
 (b-ii) *lāšayʾ*, *lāšayʾīya*, etc.
 (c) *lāšā*, imperf. *yulāšī* “to suppress, destroy, annihilate”; *talāšā* “to be suppressed, etc.”; *mulāšāh* “annihilation, destruction”; *talāšīn* “annihilation, disappearance, vanishing” (Wehr), etc.

Consonantal reduction, however, did not stop at the q-limit. By eliding the final radical of the second constituent, the glottal stop, a biliteral allomorph has been formed: /š-y-ʾ/ > /š-y/ that made it possible to create a new trilateral root /l-š-y/. Quadrilaterality of the (b)-type units contrasts here with trilaterality of the (c)-type ones.

- (a) *mā* (relative particle: ‘that which, what’ + *ġarā* ‘to happen’, *mā ġarā* lexicalized relative clause, lit. ‘that which happened’ / ‘what happened’;
 (b) *māġarayāt*, *māġariyāt*, plural of *mā ġarā* (as against plural, this singular does not seem to display a well-established noun identity) “events, happenings” (Wehr);
 (a) *yā* (exclamatory particle) + *našīb* ‘chance; fate’; exclamatory phrase: *yā našīb* “O, chance!”;
 (b) *yānašīb* “lottery”; etc.

1.2.1. This listing seems to imply that morphological structures in the theoretical state of compoundness may either undergo two opposite processes, quite

⁶ Its relation to the *root*, as a criterial element in compound morphology, evidently depends on the descriptive philosophy of the analyst and the technique adopted. For purposes of our compound study we find it sufficient to identify *lā* with a root-substituting element. The same applies to the relative particle *mā*, vocative and addressing particle *yā*, and similar monomorphemic units.

arbitrarily designed as *de-* and *re-segmentation*, or may not. On this basis, the following dichotomous division may be established: (1) one-state structures, occurring only in *desegmented* state (its segmented, or perhaps better, segmentable opposites may be found exclusively at the input level with only a few exceptions, like *lā*, *mā* and some other monomorphemic units), such as *ʿirqsūs*, *ʿarḍuḥāl*, *lādīnī*; (2) two-state structures, alternatively occurring in any of the two states, like *raʿsmāl*, *desegmented* root-word, contrasting with *rasmāl*, *resegmented*, structurally fully assimilated quadriliteral.

1.3. Quadriliteral reduction may be, in some single cases, applied to relatively complex syntactic phrases by selection of four representative radicals to this purpose. The procedure was already described by the early Arab grammarians and is echoed, in unaltered form, in very recent efforts of the Arab terminologists and norm-giving linguistic authorities to fill frustrating gaps in the technical lexicon:

mediaeval coinages:

- (a) *bi-smi llāhi (r-raḥmāni r-raḥīm)* formulaic expression: 'in the name of God, (the Beneficent, the Merciful)';
- (c) *basmal*, imperf. *yubasmil* 'to utter the above formula'; *basmala* 'the act of uttering the latter formula' (Ibn Fāris /d. 1005/ in: as-Suyūṭī /d. 1505/ 1958-i: 484-5); similarly:
- (a) *lā ḥawla wa-lā quwata ʿillā bi-llāh* 'there is no power and no strength save in God';
- (c) *ḥawqal* 'to utter the formula'; *ḥawqala* 'the act of uttering it' (ibid.);

modern lexicon:

- (a) *taḥlīl bi-l-mā* lit. 'analysis by means of water', 'hydrolysis';
- (c) *ḥalmaʿa* "id." (both variants in Khatib 1971);⁷
- (a) *ittiḥād bi-l-haydrūḡīn* lit. 'combination /of a substance/ with hydrogen';
- (c) *hadraḡa* "hydrogenation" (ibid.).

Terminological units of the latter type frequently evoke critical debates even in academic circles throughout the Arab world. Thus, for instance, Cairo Academy of the Arabic Language, proposes the term *lamaʿ* < (*lā māʿ* 'no water') "to dehydrate" instead of the Syrian *balmah* < (*bi-lā māʿ* 'without water') "idem" (*Maḡalla*, vii, 1953: 64).

Some modern terminologists push the latter procedure to the extreme by coining neologisms that are generally rejected and find no access to the Arabic lexicon. The technique may be illustrated on the following term proposed by Ismāʿīl Maḡhar (*Taḡdīd* n.d.: 73). The term consists in selecting radicals of the Arabic equivalents to Greek components of the reatomized term, for example *Hypsiprimnodontinae* (zool.): *al-ʿawsanīyāt* < *hupsos* 'high': *ʿālin* /ʿ/; *primn-* 'stern': *kawṭal* /w/; *odous*, *odont-* 'tooth': *sinn* /sn/.

Coinages of the similar type were criticized by aš-Šihābī in the Damascene *RAAD* (xxxiv, 1959: 545-554): "La langue arabe risque de devenir une nouvelle

⁷ No page indications are given with alphabetically arranged sources.

langue nabatéenne qui remplacera la langue arabe claire (*al-lisān al-‘arabī al-mubīn*). Despite the fact that similar neologisms may involve longer constituents of a thematic structure, and exhibit a relatively higher level of self-explicativeness, they are found inferior to the full-length syntactically organized constructions, as in:

- *ġamġanāhīyāt* (*ġimd* ‘sheath’ + *ġanāh* ‘wing’), syntactic: *ġamdiyāt al-‘aġniha* ‘sheathwinged’, “Coleoptera (zool., entom.)”;
- *ġišġanāhīyāt* (*ġišā* ‘membrane, hymen’ + *ġanāh*), synt. *ġaša’iyāt al-‘aġniha* “Hymenoptera”;
- *mišġanāhīyāt* (*mustaqīm* ‘straight, upright’ + *ġanāh*), *mustaqīmāt al-‘aġniha* “Orthoptera”, etc. (Hamzaoui 1965: 33-34).

Despite a rather general defiance of similar coinages, a number of them found their way into the modern Arabic lexicon, especially those of them which exhibit an acceptable degree of self-explicativeness (for more transparency, the compound-like representation is here maintained), such as:

- *kahra-miġnaṭīsī*, *kahra-tīsī* (< *kahrabā*, *kahrubā* ‘amber, electricity’ + *miġnaṭīs*, *maġnaṭīs* ‘magnet’) “electromagnetic”, as in *‘iš‘ā^c kahra-miġnaṭīsī* “electromagnetic radiation” (Khatib);
- *kahra-mā’ī* “hydroelectric” or a hybrid *haydru-kahrabā’ī* “id.” (ibid.);
- *kahra-ḏaw’ī* “photoelectric”, as in *ḥaliya kahra-ḏaw’iya* “photoelectric cell” (ibid), etc.
- (*aš-*) *šarq-‘awsaṭī* (*šarq* ‘east’ + ‘*awsaṭ*’ middle’) “Middle East (adj.), Middle Eastern”.

In spite of the acceptance of some of these and similar terms, various types of syntactic constructions clearly prevail:⁸

- *‘ifriqī ‘āsiyawī* “Afro-Asiatic”, co-occurring with *‘afro-‘āsiyawī* > *‘afro ‘āsiyawī*,
- *nawawī ḥarārī* “thermonuclear”;
- *bawli tanāsuli* “urogenital”; or even:
- *ḏaw’ikahrabī* “photoelectric”.

1.3.1. Recently, some of these constructions, particularly compound adjectives, tend to be accepted as compounds / root-words which is externally visualized by a one-word representation in the Arabic script, in contrast to substantives, which maintain their unaltered syntactic structure, as in:

- substantive (syntactic structure): *tibb an-nafs*, *aṭ-ṭibb al-‘aqlī* “psychiatry”;

⁸ Syntactic constructions played an important part in the history of compound description. Misleading classification of lexicalized syntactic phrases, shared by several influential authors, proved to be one of the most disturbing factors in this field of morphological analysis. Bielawski’s compounds (*noms composés*) involve cases, like *‘ašḥāb ar-rivāq*, *‘ašḥāb al-‘uṣṭuwān* “Stoics”; *fā‘iliyat al-‘ašī‘a* “radioactivity”; *‘āla ġahannamiya* “infernal machine”; *mu‘ālaġa bil-‘ilqāh* “vaccinotherapy” or even *ġayr māddī* “immaterial” and *‘adam at-tadaḥḥul* “noninterference” (1956: 263-320). Monteil (1960: 131-152) considerably restricts the domain of compounding; a number of evidently syntactic constructions, however, are still qualified as compounds, namely negative constructions with *ġayr*, *‘adam* (see Bielawski above), *sū* ‘mis-’, e.g. *sū at-tafāḥum* “misunderstanding”, and the like.

- adjective (root-word): *ṭibbīnafsi*, *ṭibbīʿaqlī* “psychiatric”; similarly:
- substantive: *at-tahlīl an-nafsi* “psychoanalysis”; *at-tahlīq (at-tarkīb) aḍ-ḍawʿī* “photosynthesis”;
- adjective: *tahlīlīnafsi* “psychoanalytic”; *tahlīqīḍawʿī* “photosynthetic” (*al-Mawrid* 1998).

1.3.2. A high degree of coherence of some lexicalized, nominatively highly relevant and discriminative multiword units, may lead to deviations from accepted, grammatically canonized syntactic patterns that underlie them. Terms of personal reference are particularly sensitive to the lexical dominance. The comparison of the following two syntactic structures will attest to the dominance of the lexical factor over the established syntactic pattern (the segments compared are marked by underlining):

- (1) lexical priority: *ʿamīn ʿāmm al-ʿumam al-muttaḥida* “Secretary General of the United Nations”;
- (2) syntactic priority: *ʿamīn al-ʿumam al-muttaḥida al-ʿāmm* “id.”

The last variant, though thoroughly correct, is clearly unacceptable for practical communicative purposes. Nevertheless, the lexically controlled deviation from the ‘correct’ (that is, accepted) syntactic pattern does not seem to justify a root-word classification of this attributive phrase, perhaps not yet so. Perhaps exactly because of this still existing lexical/syntactic bipartition.

The same syntactic behaviour is characteristic of a number of similar multi-componental units of syntactic background, such as *mudīr ʿāmm* “general director, manager”; as well as those denoting several high-ranking military officers, such as *ʿaqīd rukn*, *rāʾid rukn*, *fariq ʿawwal*, and the like (*Syntax* 2000: 49). The one-word (compound) nature of these units, in contrast to syntactic structures that underlie them is attested distributionally (*ibid.*).

In some other cases, however, the root-word qualification seems to be possible:

- (a) *šarq ʿeastʿ* (*aš-šarq ʿEastʿ*) + *ʿawsaṭ ʿmiddleʿ*; *aš-šarq al-ʿawsaṭ* “the Middle East”;
- (b) (*aš-šarq ʿawsaṭī* “Middle East (adj.), Middle Eastern” (for the status of the nisba suffix *-ī*, see note 5).

2. Bipartite classification of morphological constructions of compound origin, proposed in 1.2.1 above, may be integrated in a more widely conceived class of morphological constructions classified according to their ability for morphemic segmentation. In view of the more general frame, involving constructions which can be reduced to compounds side-by-side with those that cannot, the over-all arrangement will slightly differ from that used in 1.2.1). Morphological constructions displaying only one structural state, either monomorphemic, nonsegmentable (root-words), or multimorphemic, segmentable, that is structurally assimilated lexical units, typically quadrilaterals, will be qualified as *one-state* or *stable constructions*, as against those alternatively displaying both structural states (either themselves or some of their inflectional and/or derivational formatives), that will be referred to as *two-state* or *unstable constructions*. For comparative purposes, the following classification will involve two classes of morphological constructions: (a) constructions reducible back to

compounds, like *‘irqsūs, ra’smāl; rasmāl, lāšā* < *‘irq-sūs, ra’s-māl; lā-šay’*, etc.(see above), and (b) constructions irreducible to compounds, such as *tranzistar* “transistor”, *kardīnāl* “cardinal” or *munāwara* “maneuver”. In root-words of both these types only extra-root procedures are possible, mostly affix-marked plurals and relative adjectives, such as *ra’smāl* “capital”, *ra’smāl-ī* “capital (adj.); capitalistic; capitalist”; *tiliskūb, teliskūb* “telescope”, *tiliskūb-āt* “telescopes” or *tiliskūb-ī* “telescopic”. In their structurally assimilated counterparts both intra-root (i) and extra-root (e) procedures may be used, most currently affix-marked plurals and relative-adjectives, as well as the whole assortment of pattern-marked units of both inflectional and derivational nature (see further on).

(1) morphological constructions reducible to compounds:

(1.1) stable constructions:

monomorphemic (root-words):

‘irqsūs, ‘arḍuḥāl

e: *‘irqsūs-ī, ‘arḍuḥāl-āt, etc.*

māğaray-āt

barmā’-ī, -īya

bimorphemic (assimilated units):⁹

lāšā

i: *yulāšī; mulāšāh, talāšīn, etc.;*

(1.2) unstable constructions:

monomorphemic:

ra’smāl

e: *ra’smāl-ī, -īyūn, -īya, etc.;*

bimorphemic:

rasmāl

e: *rasmāl-ī, etc.;*

i: *rasāmīl; rasmāl, rasmala, tarasmāl,*

etc. (see 1.2 above);

(2) morphological constructions irreducible to compounds:

(2.1) stable constructions:

monomorphemic:

termūitr “thermometer”¹⁰

bimorphemic:

⁹ Terms *monomorphemic / bimorphemic* refer here exclusively to the morphemic structure of root-words and that of the structurally assimilated segmentables corresponding to them, with no account to affixes possibly co-occurring with them. As already indicated, affixes will be marked by hyphenation only exceptionally, for illustrative purposes. To maintain maximum of simplicity, affixes, possibly co-occurring with the intra-root procedures, will summarily be quoted under (i), as in *rasmāl* “capital” (i: *rasmāl-a* “capitalization”, or *ta-rasmāl* “to be capitalized”), instead of introducing a separate subclass of (i/e). Inflectional and derivational units, unless their respective identity is established in terms of the (e)/(i) distinction, will only be separated from each other by a semicolon, instead of introducing autonomous subclasses for them.

¹⁰ Unless otherwise specified, the examples are quoted after Wehr in a slightly modified writing and considerably shortened series of English equivalents; lexical borrowings are presented in the Wehr-reading, as well.

e: <i>tirmūmitr-āt</i> , plur.;	
' <i>aspirīn</i> "aspirin" (Sa'id 1967)	---
<i>tranzistar</i> , <i>tranzistor</i> "transistor"	---
e: <i>tranzistar-āt</i> , plur.	
<i>rūmatizm</i> "rheumatism"	---
e: <i>rūmatizm-ī</i> , rel. adj.;	
---	(<i>al-</i>) <i>kalfāna</i> , (<i>al-</i>) <i>ġalwana</i> "galvanization"; ¹¹
	i: <i>kalfān</i> , <i>yukalfīn</i> "to galvanize",
	<i>kalfānī</i> "galvanic", (<i>al-</i>) <i>kalfāniya</i>
	"galvanism", etc. (<i>al-Mawrid</i> , 1998);
	<i>munāwara</i> "maneuver; shunting (railroad)"; ¹²
	i: <i>nāwar</i> "to maneuver"; <i>munāwir</i>
	"maneuverer, manipulator");
(2.2) unstable constructions:	
monomorphemic:	bimorphemic: ¹³
(12345):	(1234):
<i>tarġumān</i> , singular (see n. 13)	i: <i>tarāġīm</i> , <i>tarāġīma</i> , plural (see <i>ibid.</i>);
	<i>tarġama</i> "translation", <i>mutarġīm</i>
	"translator, interpreter";
	i: <i>anākīb</i> , plural;
<i>anākībūt</i> "spider"	(1245):
(12345):	i: <i>barāmiġ</i> , plural;
<i>barnāmaġ</i> "program, plan";	<i>barmaġ</i> "to program", <i>barmaġa</i>
	"programming", <i>mubarmiġ</i> "programmer",
	etc.

¹¹ *kalfāna*, *ġalwana* are coextensive with the interval defined by the q-limit, with no intervening consonantal reduction; analysed into a root (*k-l-f-n*) and a pattern (*-a-Ø-a*) + affix (A) (*-a*), conveniently expressed in a complex RPA symbol *CaCCaCa* henceforth used to represent morphological structures quoted; the same holds for the lexical variant *ġalwana*. The derivational pattern *CaCCaCa* is profusely used to form native quadriliteral verbal nouns of various structural types, like *tarġama* "translation, translating", *awlama* "globalization", *haṣḥaṣa* "privatization", and the like.

¹² The derivational pattern *muCāCaCa*, also productive in modelling native verbal nouns of the same derivational class (e.g. *muṣā'ada* "help, assistance", *mu'araḍa* "opposition, resistance", etc.), backing up the borrowed *munāwara*, may be analysed as consisting of a trilateral root * *n-w-r* combined with a native pattern (cf. Sa'id 1967: 86).

¹³ Internal (broken) plurals represent the most common case of intra-root procedures and the transition from monomorphemic singulars (root-words) to bimorphemic plurals (see note 9 above) may involve various types of quadriliteral reduction, such as (in numerically indexed root symbols) *C1C2C3C4C5* (singular) - *C1C2C3C4* (plural) *CaCCuCāC tarġumān* "translator, interpreter", plur. *CaCāCiC*, *CaCāCiCa*: *tarāġīm*, *tarāġīma*. Besides 12345-1234, other types of reduction are equally frequent, e.g.: 12345-1245, 1235, 1345 or even 123456 - 1235/1236 (see 2 (2.2)).

(12345):
taliḡrāf, taliḡrāf “telegraph”;
e: *taliḡrāfi* “telegraphic”;
kardīnāl “cardinal”

(12345):
faylasūf “philosopher”

(123456):
tīlivizyōn, televizyōn, talavizyōn

e: *tīlivizyōnī* “TV-, television
(rel. adj.)”

(123456):
baṭriyark “Patriarch (eccles. title)”

e: *baṭriyarkī* (rel. adj.)

e: *baṭriyarkīya* “patriarchat (Chr.)”

(123456):
baṭlaymūs (Wehr, Schr.) “Ptolemy”;

e: *baṭlaymūsī* “Ptolemaic”

(123456):
baṭlamiyūs (Wehr; see above)

e: *baṭlamiyūsī* “Ptolemaic”.

(1235):
i: *talḡafa* “to telegraph” (Baranov);

i: *karādila*, plural;

(1345):

i: *falāsifa*, plural;

falsaf “to philosophize”, *falsafa*
“philosophy”, *falsafi* “philosophical”,
mufalsif “philosopher”;

(1234):

i: *talfaza* “television”, *talfazī* (adj.),

talfaz “to “television, telecast”

transmit by television”, *tīlfāz*

“television set”¹⁴

(1236):

i: *baṭrak, baṭrik* “id.”; *baṭrakī, baṭrikī*
(rel. adj.);

i: *baṭārika*, plur.;

baṭrakīya = baṭriyarkīya ;

(1235/1236):

al-baṭālīma (Wehr), *al-baṭālīsa* (Schr.)
“the Ptolemies”;

(1234/1236):

al-baṭālīma, al-baṭālīsa (see above);

2.3. When comparing the two classes, it becomes obvious that the transition from mono- to bimorphemic constructions is in both cases due to the consonant reduction, typically coinciding with what we call quadrilateralization (q). The latter process, however, assumes quite different forms with any of the two classes. With constructions reducible to compounds (see 2(1)), the q-process is typically bound to junctural phenomena (see 1(1.2) above) in contrast to the operative freedom of the latter with constructions irreducible to compounds (2(2)). The lavish selective freedom, mapped with the 2(2) subclass of morphological constructions, can tentatively be reduced to the following two types: (a) one element deleted: 12345: 1234, 1245, 1235, 1345, and (b) two elements deleted: 123456: 1234, 1236, 1235/1236, 1234/1236.

2.3.1. A more substantial procedural difference may be observed between structurally assimilated bimorphemic constructions, resulting from quadrilateral-

¹⁴ So far only a cautiously accepted neologism.

ized root-words, and those resulting from quadrilateralized syntactic phrases with no intervening root-word stage, as in:

(1) bimorphemics resulting from quadrilateralized root-words:

(1.1) reducible to compounds, as in:

ra's-māl > *ra'smāl* > *rasmāl* (see above);

(1.2) irreducible to compounds, such as:

(123456) *baṭriyark* > (1236) *baṭrak*, *baṭrik* (see above);

(2) bimorphemics resulting directly from quadrilateralized syntactic phrases (see 1.3), as in (selected radicals for the q-process will be written as majuscule letters):

– *lā ḤaWla wa-lā Qūwata 'illā bi-llāh* (< *bi-l-Lāh*) > *ḥawqal*, *ḥawqala* (see above);

– *taḤlīL bi-l-Mā'* (*ḥ-l-m* + ' /glottal stop has no majuscule representation/) > *ḥalma'a* (see above), and the like.

2.3.2. As is evident from what precedes, the q-process with bimorphemic constructions reducible to root-words consists in confronting nonradicals with radicals but, with those irreducible to root-words, in confronting radicals with radicals. In the former case the q-process may be qualified as radicalization through reduction while, in the latter, as a mere reduction.

REFERENCES

- al-BA^cALBAKĪ, Munīr, 1998: *al-Mawrid. Qāmūs 'inklīzī-ʿarabī* (*al-Mawrid. English-Arabic Dictionary*) 32nd ed.. Beirut, Dar el-ilm lil-malayēn.
- BADAWI, El-Said and Martin HINDS, 1986: *A Dictionary of Egyptian Arabic. Arabic-English* (Issued under sponsorship of The American University of Cairo). Beirut, Librairie du Liban.
- BARANOV, Kh. K., 1976: *Arabsko-russkiy slovar* (*Arabic-Russian Dictionary*), 5th revised and enlarged edition. Moscow, Izdatel'stvo Russkiy yazyk.
- BIEŁAWSKI, Józef, 1956: "Deux périodes dans la formation de la terminologie scientifique arabe", in: *Rocznik Orientalistyczny*, xx, 263-320.
- BROCKLEMANN, Carl, 1908, 1913: *Grundriss der vergleichenden Grammatik der semitischen Sprachen*, i: *Laut- und Formenlehre*; ii: *Syntax*. Berlin, Reuther & Reichard.
- COHEN, David, 1963: *Le dialecte arabe ḥassāniya de Mauritanie*. Paris, C. Klincksieck.
- EL-AYOUBI, Hashem, Wolfdietrich FISCHER, Michael LANGER, 2001: *Syntax der Arabischen Schriftsprache der Gegenwart*. Teil I, Band 1. *Das Nomen und sein Umfeld*. Unter Mitarbeit von Dieter Blohm und Zafer Youssef bearbeitet und herausgegeben von Wolfdietrich Fischer. Wiesbaden, Reichert Verlag.
- ERWIN, Wallace, M., 1963: *A Short Reference Grammar of Iraqi Arabic*. Washington, D.C., Georgetown University Press.
- FLEISCH, Henri, 1961: *Traité de philologie arabe, I, Préliminaire, phonétique, morphologie nominale*. Beyrouth, Imprimerie catholique.
- FLEISCH, Henri, 1968: *L'arabe classique. Esquisse d'une structure linguistique*. Beyrouth, Dar El-Machreq Editeurs, Imprimerie catholique.
- Georgetown University dialectological series: ERWIN, HARRELL et al.

- GRANDE, B.M., 1963: *Kurs arabskoy grammatiki v sravnitel'no-istoricheskoy osveshchenii*. Moscow, Izdatel'stvo vostochnoy literatury.
- GREENBERG, Joseph H., 1960: "A quantitative approach to the morphological typology of Language", in: *International Journal of American Linguistics*, xxvi, 3, 178-194.
- HAMZAOU, Rachad, 1965: *L'Académie arabe de Damas et le problème de la modernisation de la langue arabe*. Leiden, E.J. Brill.
- HARRELL, Richard S., 1962: *A Short Reference Grammar of Moroccan Arabic*. Washington, D.C., Georgetown University Press.
- HOLES, Clive, 1995: *Modern Arabic. Structure, Functions and Varieties*. London and New York, Longman.
- al-KHATIB, Ahmed Sh., 1971: *A New Dictionary of Scientific and Technical Terms. English-Arabic*. Beirut, Librairie du Liban.
- KURYŁOWICZ, Jerzy, 1961: *L'apophonie en sémitique*. Wrocław – Warszawa – Kraków.
- LANE, E.W., 1954 (reprint of the 1860 edition) : *Manners and Customs of the Modern Egyptians*. London, J.M. Dent & Sons; New York E.P. Dutton & Co Inc.
- Mağalla* > *Mağallat Mağma^c al-luğa al-^carabīya* (Journal of the Academy of Arabic Language, Cairo).
- al-Mawrid* > al-BA^cALBAKĪ.
- MAYER-LAMBERT 1897 (in: 'Studien zu Ehren Alex. Kohut, pp. 354 ff) in: BROCKELMANN 1908.
- MONTEIL, Vincent, 1960: *L'arabe moderne*. Paris, C. Klincksieck.
- RAAD – Revue de l'Académie arabe de Damas (Mağallat al-Mağma^c al-^cilmī al-^carabī)*;
- SA'ID, Majed F., 1967: *Lexical Innovation Through Borrowing in Modern Standard Arabic* (Princeton Near East Papers No 6). Princeton University.
- as-SUYŪTĪ, ^cAbdarrahmān Ġalāladdīn, 1958: *al-Muzhir fī ^culūm al-luğa wa-^canwā^cihā*, i-ii. Cairo, Dār 'iḥyā al-kutub al-^carabīya.
- Syntax* > EL-AYOUBI et al.
- WEHR, Hans, 1979: *A Dictionary of Modern Written Arabic*. Edited by J. Milton Cowan. Fourth (enlarged) edition. Wiesbaden, Otto Harrassowitz 1979.