

The Semantic Potential of Bases ('roots') and Themes ('patterns'): A cognitive approach

Joaquín Sanmartín - University of Barcelona, IPOA

The languages of the Semito-Hamitic group (on this term cf. Voigt 2002) display a mixed morphology, partly concatenated and partly discontinuous (for example, cf. Schramm 1991; and recently: Schwarzwald 2003; Fox 2003). In consequence, Semitic lexemes can be either monomorphemic or polymorphemic.¹

One of the key concepts forming the basis of Semitic studies today is 'interdigitation'. Semitic polymorphemic lexemes are thought to originate from the interdigitation of two series of phonemic sequences which are discontinuous with respect to one another. Here the terms for these complementary sequences are

→ 'base' (or 'morfema radical', 'root', 'root-tier', 'racine', or ' [Monem-]Wurzel'; cf. Del Olmo Lete 2003) and

→ 'theme' (or 'pattern', 'form', 'template', 'stem', 'schème', ' [Monem-]Schema', 'Bauform', *mškl[ym]* / *bnyn[ym]*, etc.)

The verbal and nominal (deverbal) forms that originated from the interdigitation or 'staggering' of base and theme make up the 'roots': a bimorphemic 'staggered' entity (Harris 1951: 165: 'staggered morpheme'). In the case of nominal forms, the roots are the "entries" in the lexicographic sense.² In a monomorphemic, spontaneous lexeme, which does not undergo the interdigitation process, the 'base' is also the lexemic 'root'. The root is the 'substance' of the lexeme and, as such, is able to receive a set of grammatical 'accidents' (external inflexion) that anchor it syntactically in the discourse.

1. I wish to thank Dr. W. G. E. Watson (University of Newcastle upon Tyne) for his invaluable help in the preparation of this paper for publication.

2. In Semitic lexicography the situation is more unstable in respect of verbs. Lexicographical entries reproduce merely the base, as is the case in Hebrew or Arabic, or the infinitive as in Akkadian.

1. *The 'reality' of interdigitation in language acquisition and processing: psycho-linguistic data*

Several revisions of the formative base–theme model have been proposed in recent decades. Studies carried out from a strictly (and in my opinion unilaterally) synchronic perspective plead for so-called "stem modification" and "template morphology": in verbal forms, for instance, the "disyllabic structure" of the "inner stem", with its fixed vocalism and prosody, and its affixes, would dominate.

Internal, "non-derived stems" would appear most clearly in the singular stem of simple nouns and the imperfect of simple verbs (the so-called "default forms"). These studies have been based almost exclusively on Israeli Hebrew (for example, Ussishkin 1999³), although one can find some excursions into other languages such as Aramaic (Hoberman 1992, 1997), Maltese (Hoberman / Aronoff 2003), Arabic (Benmamoun 2003, Heath 2003), Ethiopic (Rose 2003), etc. According to the revisionists, there is no reason for decomposing the non-derivative roots into "roots, templates, and vocalic melodies": in this respect, Arabic would behave like "the great majority of other languages" (Heath 2003: 129). The processes take place "very much like in English" (Benmamoun 2003: 102).

This criticism is not totally unmotivated. It is evident that Semitic morphology cannot be fully explained by the interdigitation of bases and themes: none of the lexical loans or primary nouns is produced by interdigitation, nor are most of the determiners and pronominal and deictic indicators. Passive diathesis can be explained by means of ablaut processes, at least in the base configuration (B, G, Qal, I). The formative model based on staggered or interdigitated morphology does not apply to all the linguistic domains.

The interdigitation model seems to work in another and very significant sector of Semitic morphology: the verbal system and deverbal nouns. The base (or 'root') is an abstraction, as also are the theme (or 'configuration', 'pattern') and many other linguistic concepts. Nevertheless, bases and themes are real: the reality –in the morphological sense of the term– of both base and theme resides in their functionality, in their "explanatory power" and in their operability (cf. in general Goldenberg 1994; Buccellati 1996: 59f., for Akkadian).

On the other hand, morphological interdigitation finds experimental support from psycholinguistics. Several studies in the fields of language acquisition in infancy and in the linguistic behaviour of adults speaking a Semitic language corroborate the operation of such a mechanism.

→ Prunet / Béland / Idrissi (2000) carried out a series of experiments to establish the nature of the units stored in the mental lexicon. For this purpose they studied mistakes of metathesis produced by aphasias, as well as the phonetic reversals occurring in word-plays and linguistic slips. These tests unequivocally prove the full validity of the model based on the morphemic staggering of bases and themes.

The studies mentioned show that, while Arabic and Israeli aphasics reverse the sequence of the base consonants, and only their sequence, the dysfunction does not affect the thematic melodies (/ħubu:l/ "ropes" > read (!)/ħulu:b/). Aphasia does not affect either the vocalic affixes or the (mistakenly so-called) 'thematic extensions' (/ma-rkaz/ "centre" > written *m-krz* [read presumably: /ma-kraz/]; /mu-s-t-aʕmir/ "colonist" > read /mu-s-t-amʕir/, etc.⁴). Such mistakes prove that the subject selects groups of consonants

3. Following MCCARTHY (1982, 1983); MCCARTHY / PRINCE 1990.

4. Whenever the phoneme /m/ belongs to the consonantal base, it undergoes metathesis, but not if it is part of the theme: /mesa:ʔ/ "evening" > produces written (!)*sma:ʔ*, to be read probably as (!)/sama:ʔ/, whereas /ma-žall-a/ "magazine" > produces (!)*mlžt*, probably read (!)/ma-lažž-a/ (PRUNET / BÉLAND / IDRISSE 2000: 617).

that form a sector of his mental lexicon: we call the consonants that match this morphological string a "base" (or 'root'). This does not happen, or else it happens in a different way, when the patient with aphasia confuses the consonants in French, because the consonants alone do not constitute morphologically significant constellations in the morphologically concatenated Indo-European languages. The analysis of several cases of linguistic slips produced by speakers of a Semitic language corroborates the conclusions which the authors derive from the behaviour of aphasics. Likewise, the authors state that in Arabic word-plays ("ludlings"), based on the commutation or permutation of syllabic segments within a word, the base consonants are rearranged but the vocalic sequence is preserved (Al-Mozainy 1981: 86; Heath 1987: 184).

All these studies point to the possibility of experimental access to the morphological independent unit we call "base". The same applies to the morphemic status of "themes" (or 'patterns', 'templates'). In the same study (2000), Prunet / Béland / Idrissi examine the behaviour of aphasics when selecting such themes. The data obtained for Arab and Israeli aphasics (cf. Barkai 1980) indicate that the disorder produces lexically non-existent or syntactically inappropriate but morphologically correct verbal forms, and that, in all the cases, the aphasia does not affect the correct sequence of the base consonants. So, for example, we detect the confusion in Arabic between the III stem /kaʔtab/ "to correspond (by letter)" and the I stem /katab/, or in Israeli Hebrew between the Hifʔil-stem /higdilu/ "they enlarged" and the Qal-stem /gadlu/. Likewise, also in Israeli Hebrew, we find forms that do not occur actually in the paradigm of a particular verb, such as (!)/yasug/ (a non-existent Qal) instead of the Nifʔal /yisog/ "he will retreat" (/N-S-ʔ/), or (!) /hagati/ (also a non-existent Qal) instead of the Hifʔil /higati/ "I arrived".

Interestingly enough, aphasia does not seem to affect only the strictly vocalic morphemes: in fact, it affects the selection of the whole theme, involving also its consonantal components (the so-called 'extensions' /ma-/, /mu-/, /ta-/, etc.). This fact puts into question the very existence of the ternary scheme (consisting of "root", "template", "vocalic melody") as postulated by the auto-segmental model (cf. McCarthy 1981, 1982; Bohas 1997). Everything seems to indicate that, in effect, the relationship between the vocalic morphemes and the thematic tiers (or 'patterns') is much closer than between either of these and the consonantal base (on the assumption that both tiers –that is vocalic melody tier and template tier– exist independently). Aphasics manipulate the vowels as an integral part of the theme. The vowels seem to be "already linked to templates", so that both supposed levels –"templates" and "vocalic morphemes"– actually make up only one morphological element.

The reversal of syllabic sequences produced by aphasics does not imply incorrect handling of the prosodic constituents of the "template" (in the autosegmental sense of the term), but rather an erroneous selection of the "template" itself (Prunet / Béland / Idrissi 2000: 629⁵).

In any case, fundamentally there are two mistakes due to aphasia:

- First, the creation of 'new' bases, due to the reversal of base consonants.
- Second, the wrong selection of themes or configurations that actually exist in the language but are semantically and syntactically inappropriate.

5. The authors point out only one case of the production of a non-existent theme: (!)/mhas/, instead of /hams/ "whispering" (PRUNET / BÉLAND / IDRISSE 2000: 629).

The cases of mistakes resulting in the production of new themes or configurations among aphasics are extremely rare. Basically, there is a strong tendency to produce possible words, that is, words that may or may not occur, but that actually fit the morphological constraints of the language.

The same can be said concerning the linguistic slips of healthy, non-aphasic adult speakers of a Semitic language (Berman 1981: 619; Abd-El-Jawad / Abu Salim 1987: 149; Prunet / Béland / Idrissi 2000: 631): the vocalic or thematic structures remain in situ, whereas the consonants of the base structures are permuted: /Nabi:l u Kama:l/ > (!)/Kami:l u Naba:l/. The case of word-plays is slightly different. Unlike the cases of mistakes due to aphasia or involuntary slips, word-plays are always conscious and voluntary linguistic productions, endowed with a greater or lesser degree of creativity. In these cases, new bases and new themes are frequent, and the new forms can sometimes become standardized diachronically.

The fact that aphasics produce a high number of consonantal reversals –whilst regularly respecting the thematic structures– is a powerful indication that Semitic base consonants 'float', that is to say, they belong to a peculiar level ("tier") of their own and, therefore, bases and themes are two different morphemes.⁶ Experimental data on the acquisition and production of language in infancy corroborate these conclusions. From the age of three, Israeli and Palestinian children are capable of identifying and extracting the base consonants and of producing new verbs related to nouns and adjectives that are already known to them, applying an appropriate theme to the consonantal base skeleton. The attribution of the selected theme to the extracted base is carried out always "in the classical Semitic way", but obviously without paying attention to the actual or historical occurrence of the generated form (Berman 1999, 2003; Ravid 2003).

The only correct conclusion is that, functionally, bases and themes form two different sets within the mental lexicon of Semitic language speakers. The ability to extract bases and themes from common lexemes –and to create new lexemes (whether lexicographically canonical or not) through the interdigitation of both morphemic sets– is already fully documented at a relatively early age (approximately two to two and half), at least relating to the mechanical control of both morphemic series, obviously leaving out of consideration the syntactic regulators that make possible the contextually correct selection of the concrete form (Berman 1982; Borer 2003a).

The fact that the study of Semitic morphology concerns the organization of morphologically complex lexemes poses some subtle problems. For example, how are morphologically complex words stored in the mental lexicon? Are they stored as complete words (according to the "Full Listing Hypothesis") or classified by their components, that is, separately as simple words, roots and affixes (according to the "Decomposition Hypothesis")? Several psycholinguistic studies carried out among English-speaking aphasics point out that the lexemes are filed as resolved into their components: in fact, the entries of both the phonological and orthographic lexicon consist of independent representations of roots and affixes (Miceli / Camarazza 1988; Badecker / Camarazza 1991; Camarazza / Hillis / Leek / Miozzo 1994). In addition, other experiments demonstrate that adult speakers of Israeli-Hebrew relate faster and more easily to each other words that present a common base structure than the words with a similar lexical meaning but with a different 'base' structure. This becomes especially clear when, in the course of the experiment,

6. "There are more metatheses in Arabic than in Indo-European because consonants float before they are mapped to the template. At this early stage, Arabic consonants are more prone to metathesis than are Indo-European consonants. Once Arabic consonants have been mapped to the template, the usual preference to metathesize adjacent consonants holds in Arabic, as [it] does in other languages. Consequently, Arabic consonants have two configurations under which they can metathesize: before and after they have been anchored. Indo-European consonants have only the latter configuration" (PRUNET / BÉLAND / IDRISSE 2000: 626 n. 17).

the stimuli are separated by a longer interval, which indicates that during the process of recognition the speaker is particularly sensitive to the morphological information about the base in question: recognition of morphology outweighs semantic similarity (cf. Deutsch / Frost 2003: 171).

It was also established that the subjects, while extracting more or less valid triconsonantal sequences from forms provided with authentic themes, identify these base consonants more easily and faster if the consonants to be extracted are productive in the real language, although interwoven with other themes and not with the proposed ones. When the task consisted of rejecting non-existent words (formed, for example, by the interdigitation of consonantal sequences that never actually occur with existing themes), the process of rejection was faster than when the subject tried to identify false forms formed by interdigitation of really existing base sequences with themes which are legitimate but actually do not occur with the bases proposed in the stimulus. All this confirms that the process of identification of a given form includes a stage during which the subject resolves both the morphemic strings (base and theme) that comprise the lexemic core. It confirms also that word recognition does not happen "word for word": otherwise one could not explain the different reaction times of subjects who are set different tasks (Bentin / Feldman 1990; Feldman / Bentin 1994; Feldman / Frost / Penini 1995; Goral / Obler 2003).

2. *Interdigitation phases in language acquisition*

The data set out above can throw some light on the process of language acquisition by Semitic language speakers. Even if many aspects remain unresolved, we can assert that, in infancy, language learning proceeds through several phases. Following Borer (2003a, 2003b), we may call them (1) The naming phase, (2) The morphophonological phase, and (3) The morphosyntactic phase.

→ The naming phase: The first concepts to be lexicalized originate from the conflation of sensation bundles into coherent and highly cohesive conceptual sets: individuals and objects of the environment that are conceived as discrete units. Cross-linguistically, it is a well-attested fact that the first outputs are proper names and common concrete nouns that refer to individualized entities.

Children apply to the concept a name received from their environment, because only socially sanctioned names and nouns assure the efficiency of naming: comprehension precedes production. For children, names and nouns are kinds of phonic calls, able both to evoke a certain amount of reality and to be understood easily as such by the environment. In the learning process, nouns, names and nominalized action verbs come first, as conceptual conflations of sensory information from the outside world. For this reason, the first lexical items form an open set and, at the same time, the set which, cross-linguistically and semantically, appears as the most constant (Gentner / Boroditsky 2001). Studies in language acquisition carried out among Israeli children show that 90% of their first 'nouns' have concrete objects as their referent and that 87% of their first verbs are verbs of action (Dromi 1987). The primacy of nominal forms over the verbal forms is self-evident in this first naming phase.

→ The morphophonological phase (Borer 2003a: 329ff.), which comes next, consists of a series of successive processes. I would stress the following sequence:

(a) First, certain melodic relationships are detected between similar phonemic clusters consisting of intense or exclusively vocalic sounds, such as [A-A (ː)], [Aː-I], [MU-A-I], [MA-A], [TA-A], etc. The persistence of the vocalic melodies acquired in infancy, patently clear in mistakes due to aphasia as well as

in word-plays and 'ludlings',⁷ is a powerful hint of an early and deep implantation of vowels and vocalic melodies in the gears of language production.

(b) Second, the child detects the coincidence of some melodies with particular sememic bundles featuring several specific semantic roles. In simplified form: children begin to associate the melody /A:-I/ with the sememic bundle +AGENT, the melody /A-A:/ with the sememic bundle +ACTION, /MA-A/ with the sememic bundle +INSTRUMENT and / or +PLACE, and so on.⁸

In this way the child collects a set of 'themes' and begins to master them. The set comprises discontinuous morphemes, each provided with relatively definite semantic contents.

(c) Third, the child detects the co-occurrence of certain consonantal sequences in several thematically different constellations. It registers the consonantal coincidence between, for example, /Ra:KiB-/ and /maRKaB/, and notices that /Ša:KiN-/ and /maŠKaN/ behave like the first pair /Ra:KiB-/ and /maRKaB/, thus discovering the coherence between certain phonic sets and certain significant bundles, making the following inference:

"{1a:2i3} relates to {ma12a3} as {4a:5i6} relates to {ma45a6}"

At this stage, the child begins to feel increasingly confident about coining new words, drawing on information bits from its familiar milieu (Berman 2003: 274ff.), and makes a further inference:

"If there is a {7a:8i9}, then there is also a {ma78a9}"

Admittedly, it often generates hypercorrect or wrong phonemic combinations respectively: that is, forms that do not always find the necessary lexemic and pragmatic support in the adult language.

(d) Fourth: Abstraction of the consonantal bases. The stage now reached by the child consists of two crucial elements: first, the knowledge and management of all morphophonologically possible themes and, second, the ability to extract the consonantal base of the input word (/R-K-B; Š-K-N) and to store it independently of the rest of the phonological environment (Borer 2003a: 336). The child now relies on a "consonantal skeleton" that allows him to interpret and to produce new words, or words with which he is not yet fully acquainted.

(e) The last stage of the morphophonological phase consists of the production of new forms by interdigitation of the extracted bases with the themes which are already familiar to the child. All the forms produced at this stage are morphophonologically correct, in so far as they represent morphophonologically possible words. As shown by several studies carried out with Israeli and Palestinian children (Ravid 2003), they seem to be fully acquainted with certain extremely important aspects of Semitic morphology. The children are able:

7. "Ahora que los ladros perran, ahora que los cantos gallan, / ahora que albando la toca las altas suenas campanan; / y que los rebuznos burran y que los gorjeos pájaran, / y que los silbos serenan y que los gruños marranan, / y que la aurorada rosa los extensos doros campa, / perlando líquidas viertas cual yo lágrimo derramas / y friando de tiritio si bien el abrasa almada, / vengo a suspirar mis lanzos ventano de tus dejabas", José Manuel MARROQUÍN (cited in: G. GARCÍA MÁRQUEZ, *Vivir para contarla*, Barcelona: Mondadori 2000: 199).

8. COMRIE 1989: 58f.

- to extract the consonantal bases from existing lexemes,
- to store the extracted bases with their essential meanings, and
- to interdigitate the extracted bases with morphophonologically correct themes.

In fact, Hebrew speaking children know that word-pairs such as [yaš'a] // [hoš'i] "to go out" // "to extract", or [hid'biq] // [nid'baq] "to glue" // "to be glued", do share bases and a conceptual basic meaning. Demonstrably, the ability to extract the consonantal base from a given form and to interdigitate it with a different theme can be detected at the age of 2 and is fully developed by the age of 4 (Borer 2003a; Berman 2003).

→ The morphosyntactic phase (Borer 2003a: 347ff.), the third and final phase, develops between the three and a half and six years of age. It is characterized by the selection of certain themes in function of the syntax of the arguments. The form selected, actually a theme, is often incorrect because it is based on exclusively syntactic criteria, excluding lexical or pragmatic constraints that operate fully in the adult language. At this syntactic-morphological stage, hypercorrections are very common. For example, the wrong selection of a theme with the semantic component +CAUSE (Hebrew Hipšil: hypercorrect (!)/madhi'fa) to express a causativity that is already lexically implicit in the basic meaning of the neutral form (Qal participle feminine /D-Ḥ-P/: /do'ḥefet/), as found in the adult language: the Hebrew base /D-Ḥ-P/ "to push" does not generate Hipšil causative themes. The incorrect selection is obviously due to the component +CAUSE, implicit in the meaning "to push" (< TO MAKE [something] MOVE forwards), which motivates the child to select an explicitly (but lexically non-existent) causative theme, here the Hipšil.

3. *The semantic potential of bases and themes*

Isolated and unrelated phonemic sequences coexist in the Semitic lexicon with sets of phonemic sequences that share some morphemic elements: consonantal bases and themes. Since the interdigitated morphemes are discontinuous and semantically complementary, some syntax is required to establish a relationship between both sequences. Exceptions to the mechanism of interdigitation are bases of all kinds that obey a concatenated morphology of the agglutinative type,⁹ with possible apophonic variants: that is, linguistic loans, Kulturwörter, personal names, pronominal indicators, temporal and spatial localizers, syntactic connectors, determiners, and so on.¹⁰ In Semitic, the interdigitation of discontinuous morphemes operates in the verb, in the deverbal noun, and in a limited number of denominative verbs (Buccellati 1996: 70ff.; Fox 2003: 64).

All bases, whether monomorphemic (spontaneous) or consonantal, are cognitive devices that segment the ontological reality into coherent portions. Bases are the linguistic outputs of a mental mapping of the world. The Semitic base, whether continuous or discontinuous, is therefore a morpheme of lexical classification which demarcates a cognitive area that is more or less concrete, more or less complex. So, for example, /ʔan-t=/ YOU (a pronominal indicator), /bayt=/ DWELLING/ (an isolated morpheme:

9. For example: /Han/ + /Ha/ + /k/ > /ʔan-a/ - /k/, /Han/ + /ta/ > /ʔan-ta/, /Han/ + /di/ > /hanna-di/, etc.

10. SCHRAMM 1991: 1403: "The 'little' words of the language".

concrete countable object), /M-ʔ-D/ TO BE MUCH (a discontinuous morpheme: one-valency reference of condition / state), /H-L-K/ TO WALK / GO (to MOVE ON ONE'S OWN EFFORTS; a discontinuous morpheme: one-valency reference of action), /R-K-B/ TO MOUNT / TO MOVE WITHOUT ONE'S OWN EFFORTS (a discontinuous morpheme: two-valencies reference of action), /Š-P-R/ TO WRITE (a discontinuous morpheme: three-valencies reference of action).

The monomorphemic roots are semantically and morphologically complete and, therefore, already lexemic. Their segmentation into basic (that is consonantal) and thematic strings produces only pseudo-morphemes: seemingly independent phonemic segments (Harris 1951: 177ff.) with no semantic value.¹¹ The concept of pseudo-morpheme is useful to explain the structure of primary / spontaneous nouns, that is, nouns which, while apparently presenting the same thematic morphology as the deverbal nouns, are not the product of interdigitation. In fact, they do not share their consonantal sequences with any other nouns. In other words, it is as pointless to look for a base (!)/ʔ-T-N/ with the theme {qata:l} in cases like Akkadian *ata.mum*, Tiberian-Hebrew *ʔawn* /ʔat(on)/ or Arabic *ʔata.m* "female donkey, she-ass", as to separate a morpheme '-g' from English words like bag, rug or bug or, for that matter, a morpheme 'bo-' from Spanish bobo, bolo or bono. It is not that themes and bases are semantically irrelevant in primary nouns: they simply do not exist as morphemic entities. Unlike monomorphemic lexemes, discontinuous verbal bases are only potentially lexemic. Their semantics is restricted to a coherent bundle of semantic components endowed with some argument structure (monovalency, bivalency, trivalency, admission of certain adjuncts, etc.). In Aristotelian terminology we could define the consonantal bases as the 'materia ex-qua' (*hylē*) of the lexeme. It seems that human learning proceeds along two different channels, or rather, is governed by two kinds of dominance: cognitive dominance and linguistic dominance (Gentner / Boroditsky 2001). In the case of cognitive dominance, "concepts arise from the cognitive-perceptual sphere and are simply named by language". The consonantal bases are linguistic expressions of concepts proceeding mainly from the cognitive-perceptual sphere.

On the other hand, themes ('patterns') are expressions governed by strictly linguistic mechanisms: the world presents perceptual bits, and the themes have a say in how these bits, clustered together as consonantal bases, are conflated into fully significant lexemes. Whilst the bases belong to cognitive-perceptual dominance, in my opinion, the themes belong to linguistic dominance. It is commonly accepted that some semantic universals are the main conceptual engines producing the different verbal and deverbal forms: TO DO (under the control of an AGENT), TO HAPPEN (said of something that takes place as a concrete EVENT or as a PROCESS that is more or less extended in time), TO BE (said of something that is or exists¹²), etc. It is not our aim here to carry out a componential analysis of the verbal and nominal themes of Semitic, or to make a survey of all the morphemic schemes (cf. Kienast 2001: 70ff., Fox 2003). In any case, it seems evident to me that the concepts of +AGENCY (both in its exotropic and reflexive modalities), +CAUSE / +FACTITIVITY, +CAUSATIVITY, +MANNER, +LOCATION (temporal or spatial), +INSTRUMENT, +PATH, and so on, are all semantic universals or cognitive primary categories that –whether innate or not– govern most of the thematic morphemes. These metalinguistic categories belong to the most transparent meanings of the verbal and deverbal themes.

11. See the treatment of the seemingly independent segments *-y* and *-is* in English *wh-y*, or *th-is* in comparison with the partially independent morphemes *th-ere*, *th-en*, *th-at*, *wh-ere*, *wh-en*, *wh-at* in Harris 1951: 177ff., cf. 158f. 'Partially independent' are, for example, '-mundi' in Spanish *mapamundi*, '-familias' in Latin *paterfamilias* or 'him-' in German *Himbeere*, as against the free independent morphemes 'mapa-', 'pater-' or '-beere'.

12. "Conceived as existing, rather than happening", Lyons 1977: 483. Similarly, cf. the 'semantic roles' in COMRIE 1989: 58f.: agent, force, instrument, experiencer, patient (: entity that is said to be in a certain state or to undergo a change in state).

According to all this, the 'theme' or 'morpholexical configuration' is a discontinuous morpheme of lexical realization. In my opinion, the function of the theme is not primarily grammatical but semantic: the theme is not a grammaticizer of the base, but its semantic formalizer. The themes realize semantically the consonantal strings by means of a set of universal semantic features and their resulting syntactic restrictions.

The absence of thematic allomorphs shows the pre-grammaticality or supra-grammaticality of the thematic morpheme. The theme operates primarily at the semantic level. Notice that the different themes are in a paradigmatic relationship to one another: they are commutable morphemes within a series of invariants,¹³ while the grammatical morphemes can concatenate to form sequential syntagms (cf. Goldenberg 2002: 32ff.). Corriente (1996: 26) calls the theme a 'derivational morpheme', and defines it, quite rightly, as "endowed with semantic functionality or aptitude to express [...] certain logical categories". Evidently, the term 'derivational' denotes the semantic potential of the theme. Grammatical accidents, on the other hand, are not concerned *per se* with lexical meanings, but with syntactic relationships. In hylemorphic terms, it may be said that the theme, while interdigitating with the base, works on it like *forma* on *materia*: it bestows conceptual (and therefore linguistic) reality on the base. It changes a mere (nominalized) reference [like /P-R-S/ "to separate"] into a (verbal) predication [*yaPRuS "it separated"; Pa:RiS "it is a separator"].

Like the Aristotelian *éidos* or *morphé* and like the Scholastic 'forma', a theme is the *principium individuationis* (the determinant) of the base it determines, and the output of this determination is the 'word' (lexeme). Therefore the theme is the 'Gestalt' (*skhêma*) and, as such, it is 'more' than the *materia* ('the base') and (cognitively) 'prior' to it (cf. Met. Z. 1). If the 'base' can be conceived as the *materia ex qua*, the theme is the form by which the lexeme is. The theme shapes the base morpholexically and enables syntactical interrelationships (Huehnergard 2002: 125; Fox 2003: 57f.). Studies on semantic universals and on the distribution and planning of the lexicon clearly suggest that any language can be divided into two large lexical subsets: One subset comprising a small, enumerable and restricted quantum of words or morphemes which can be regarded as indefinable (the so-called primary semantemes, or lexical universals), and the other subset consisting of a "wide digest of words capable of being defined in terms of the indefinable" (Goddard / Wierzbicka 1994; Wierzbicka 1992, 1994, 1997). Obviously, 'themes' belong to the first, small subset of enumerable and restricted morphemes, whereas 'bases' belong to the second, open subset of morphemes that can be analysed semantically and defined in terms of the indefinable.

The two basic semantic functions are reference and predication. Semitic bases, whether spontaneous nominal or discontinuous verbal bases, refer to a perceived or imagined world. Bases are references, understanding the term 'reference' in Wunderlich's sense (see Wunderlich 2004). Predication is a statement about something to which reference has been made. Whereas reference is the typical function of nouns, pronouns, determiners and so on, cross-linguistically, predication is the essential function of verbs and deverbals. Discontinuous, referential verbal bases become fully operative predicators in Semitic if, and only if, they are interdigitated with themes. The semantic universals involved in Semitic themes are to be considered as the genuine carriers of the predicative force. Themes are bundles of primitive elements, and predication serves as the link between them and the arguments implied in the base (see Chierchia 2003).

13. The match between a determined argument structure (semantic and syntactic data) and the morphophonological datum is unambiguous: a given thematic configuration ('stem') is either transitive or intransitive, but not both things at the same time. One could compare, on the contrary, Spanish "come" (he / she eats) in the expressions "come patatas fritas" (trans.) and "este niño no come bien" (intrans.). Classical examples from English include *move*, *drop*, *shake*, etc.

It helps to bear in mind that the lexemic result of interdigitation very frequently exceeds the sum of the original semantic potentials of base and theme: the process is not always so 'transparent' (Voigt 1988: 41) as to render unnecessary a lexicographically specific record. Therefore, it is evident that both the meanings "charioteer" and "sexually mature, 'mounting' animal" of the Akkadian term *ra:kibu(m)* go beyond the semantic possibilities of merely adding together the base /R-K-B/ and the agentive theme {qar:til}. The influence over the lexicalization processes of factors that are peculiar to and immanent in the organization of every individual language is an aspect we should be aware of; likewise, we ought to consider certain lexicalizations as due to particular pragmatic constraints.

4. Future study and future tasks

4.1. Functional parallelisms between affixed derivation and internal flexion

For morphologically concatenated languages, three components are usually distinguished: a lexemic root (LEX), a derivational affix (AF) and a grammatical morpheme (GR).¹⁴

So for example, in the case of:

"the instructions"

we would obviously obtain following morphemic distribution:

the	in	struc	tion	s
GR	AF ₁	LEX	AF ₂	GR
DET:PL				PL

Semitic languages make use of seemingly analogous procedures. Actually, these languages exhibit the same three concatenated elements –LEX, AF and GR– in the nomina primaria / spontanea as well as in several deverbal nouns. In the nomina primaria –the bases of which are also their lexical 'roots'– derivation is necessarily produced by concatenation; so, for example, in Akkadian *šarr-ut-um* "royalty", *qaqqad-am-um* "the one with a large head", *a:l-ay-û* "citizen" (Buccellati 1996: 137ff.). Similarly, in verbal and deverbal forms we can detect affixation mechanisms (for example in the Akkadian subjunctive suffix /-u/ and in some deverbal nouns such as *na:din-am-um* "the seller"). Obviously, concatenated derivations always operate on previously interdigitated forms.

Although the existence of semantic and lexical limitations to the general applicability of affirmative derivations is well established in Semitic studies, it has also been pointed out that the role of the derivational Semitic affixes is parallel, functionally and semantically –but not formally– to the role of the themes ('patterns') that operate in internal inflexion (Buccellati 1996: 139). Extrapolating: The themes of internal Semitic inflexion provide interesting functional and semantic parallels with the derivational Indo-European affixes (called *declinatio voluntaria* by Marcus Terentius Varro, in his *De lingua Latina*, in opposition to the *declinatio naturalis* or grammatical endings).¹⁵ Semitic themes, like concatenated Indo-European affixes, comprise a 'closed' morphemic system, a numerically constant set. On the contrary, the

14. *Lexèmes, affixes, modalités* (: *déterminants grammaticaux*) in the terminology of MARTINET (1967: 119).

15. So for example the themes {maPRaS} and {taPRaS} (INSTRUMENT, MEANS, PLACE) are semantically and functionally parallel to several Indo-European suffixes, such as Latin {-bulum} *patibulum* (< *patior*-), {-culum} *vehiculum* (< *veho*:), {-crum} *sepulcrum* (< *sepelio*:), {-?le} *cubile* (< *cubo*:), {-mentum} *documentum* (< *doceo*:), etc.

Semitic discontinuous verbal bases as well as the continuous, spontaneous bases, belong, like the Indo-European 'roots', to an open, extendable lexical set. In a general way, the distinction between the open lexical class and the closed lexical class fits both the main processes of cognition and of language production. The first, the open lexical class, is governed by experience –in the last instance by sensory experience, and the second, the closed lexical class, is governed by internal linguistic structures. We can speak –on the one hand– of an open conceptual and predominantly cognitive array, which responds to incoming perceptual inputs and is endowed by the language with the corresponding linguistic outputs, and –on the other hand– we discern a closed conceptual array, directly governed by inner linguistic structures. That is, the segments of information entering through sensory experience are not already linguistically pre-arranged, so that some internal linguistic structures are required in order to organize conceptually the bits of information from the outside world. Evidently, the Semitic (verbal and nominal) themes are a closed morphemic set which corresponds to operations that originated strictly in the linguistic domain.

We can propose the following chart (cf. Gentner / Boroditsky 2001: 216):

← cognitive dominance		linguistic dominance →			
← open morphemic class		closed morphemic class →			
1	2	3	<div style="border: 1px solid black; display: inline-block; padding: 2px;"> INTERDIGITATION ← BASES THEMES → </div>	5	6
proper names	concrete nouns	kinship terms & other relational systems	verb forms and deverbals	spatial indicators	signallers, determiners, connectors, metapredicates
<i>Šamsi:adad</i> ,	<i>/kalb-/</i> "dog", <i>/qinn-/</i> "nest" <i>/ʔabn-/</i> "stone"	<i>/ʔab-/</i> "father", <i>/šarr-/</i> "king", <i>/qaqqad-/</i> "person"	<i>/ʔarkab/</i> "I rode", <i>/raʔkib-/</i> "rider", <i>/markabt-/</i> "charriot"	<i>/ina/</i> "in", <i>/ki:d-/</i> "outside"	<i>/ʔana:ku/</i> "I" <i>/u/</i> "and" <i>/la:ʔ/</i> "no"

In linguistic systems governed by concatenated morphology, such as Indo-European, the mechanism consists of the sequential ranking of radical lexemes, derivational affixes and grammatical morphemes. In linguistic systems with partial or complete non-concatenated morphology, the mechanism consists of the interdigitation of bases that map the world with themes that animate them with sets of logical categories and semantic universals, in order to produce lexemic 'roots'.

4.2. *The semantic fuzziness of nominal themes*

It has been noticed that often the semantic characteristics of thematic morphemes are not particularly consistent (especially in nominal themes), so that the meaning of a lexeme (especially of a nominal lexeme) cannot be readily predicted from the informative sum of its base and its theme (Deutsch / Frost 2003: 166, 175f.). From a synchronic perspective, interdigitation explains the meaning but does not predetermine it. Fox (2003: 43f.) has touched on a sore point when he says that, while bases ("root") generally have a "fixed meaning", themes ("patterns") do not, or "more rarely":

"The present study [...] takes the position that one pattern can have semantic groups of different fields of meaning within it, since no distinction of form is reconstructible from the languages" (Fox 2003: 22).

The base–theme relationship seems to behave somewhat differently in verbs from the way it behaves in deverbal nouns. In deverbal nouns, the semantic potential of the theme is much weaker and much more unstable than in the finite or nominal verbal forms: in general, the thematic nominal system is less transparent than the verbal one. What is more, the efficiency of the nominal themes is restricted to the semantic level, whereas verbal themes, apart from being semantically much more consistent, also affect the syntactic level. Possibly the wide variety of nominal themes, if compared with the relatively more reduced set of verbal themes, produces some perplexity in recognizing and processing the former. Also, as a matter of fact, it seems that many nouns are filed and handled in the mental lexicon as complete and simple units, apart from their structurally interdigitated origins. In other words, whereas the verbal themes are entitled to a peculiar status as organizational factors of the mental lexicon, there is discussion as to whether the nominal themes enjoy the same clear statute (Deutsch / Frost 2003). That would explain the difficulties in defining the nominal theme (and sometimes also the verbal one) as an autonomous lexical unit (Hoberman / Aronoff 2003: 70ff., 75f.; Fox 2003: 43). It has been stated that the production of illegitimate forms (which are the product of interdigitating morphologically possible themes with bases that actually do not admit them) finds a functional parallel in the concatenated languages: namely in those lexical mistakes due to the use of certain affixes in inadmissible contexts. This raises the problem of the lexical independence of the derivational affixes.¹⁶

But if theme and base are two different morphemes, we cannot help but ask about their respective semantic qualities and about the morphophonological process of their interdigitation. Evidently, not every morpheme is also a lexeme. But any and every morpheme is a linguistic significant and, as a significant, has to mean something. This meaning need not be lexemic. Obviously, this is the case with grammatical morphemes, but the question might not be so evident in respect of derivational morphemes. In concatenated morphologies, the search for a lexical definition of affixes turns out to be neither possible nor desirable, at least from a strictly synchronic perspective. To avoid here the usual and well-worn examples with the English suffixes in {-ly}: It is evident that the Latin suffixes {-aria}, {-arium} in *argenta.ria* and *aerarium* are not, by themselves, lexicographically tangible units. Nevertheless, we might try a naïve definition and say that, synchronically, both the suffixes{-aria} and {-arium} add the 'idea' of "PLACE where to guard ('silver' *argentum* / 'copper' *aes*)". Certain semantic nuances are often detectable but not always definable. In any case, we know that a morpheme is one thing and a lexeme quite another, although every morpheme is necessarily a linguistically significant entity.

16. "Doit-on parler d'un suffixe *-ceté* si un enfant, partant de *méchant*, *méchanceté*, forme *cochonceté* à partir de *cochon*?" (MARTINET 1967: 135).

Future studies on derivative, concatenated affixes and on interdigitable themes should also include the question of the semantic relationship of both with some of the primary semantemes and lexical universals. Another aspect to bear in mind is the possible and very legitimate methodological dilemma between the more or less radically diachronic perspectives (cf. for example Kienast 2001: 108f. on the nominal themes in {ya-, ma-, ta-, *šu-}) and the more or less radically synchronic-diatopic approaches (Fox 2003).

5. Bibliography

- 1981 ABD-EL-JAWAD, H. / ABU SALIM, I., "Slips of the tongue in Arabic and their theoretical implications", *Language Sciences* 9: 145-171.
- 2003 ALEXIADOU, A. / ANAGNOSTOPOULOU, E. / EVERAERT, M. (eds.), *The unaccusativity puzzle: Explorations of the syntax-lexicon interface (Oxford Studies in Theoretical Linguistics)*, Oxford: Oxford University Press.
- 1981 AL-MOZAINY, H., "Vowel alternations in a Bedouin Hijazi dialect: Abstractness and stress", University of Texas at Austin, Ph.D. Dissertation.
- 1991 BADECKER, W. / CAMARAZZA, A. "Morphological composition in the lexical output system", *Cognitive Neuropsychology* 8: 335-367.
- 1980 BARKAI, M., "Aphasic evidence for lexical and phonological representations", *Afroasiatic Linguistics* 7/6: 163-187.
- 2003 BENMAMOUN, E., "The role of the imperfective template in Arabic morphology", in: Shimron 2003: 99-114.
- 1990 BENTIN, S. / FELDMAN, L.B., "The contribution of morphological and semantic relatedness to the repetition effect at long and short lags: Evidence from Hebrew", *Quarterly Journal of Experimental Psychology* 42: 693-711.
- 1981 BERMAN, R., "Language development and language knowledge: Evidence for the acquisition of Hebrew morphology", *Journal of Child Language* 8: 609-626.
- 1982 BERMAN, R., "Verb-pattern alternation: The interface of morphology, syntax and semantics in Hebrew", *Journal of Child Language* 9: 169-191.
- 1999 BERMAN, R., "Children's innovative verbs versus nouns: Structured elicitations and spontaneous coinages", in: L. Menn / N. Bernstein-Ratner (eds.), *Methods in Studying Language Production*, Mahwah, NJ: Erlbraun: 69-93.
- 2003 BERMAN, R., "Children's lexical innovation", in: Shimron 2003: 243-291.

- 1997 BOHAS, G., *Matrices, étymons, racines: Eléments d'une théorie lexicologique du vocabulaire arabe (Orbis Supplementa 8)*, Leuven: Peeters.
- 2003a BORER, H., "Computing argument structure", in: Shimron 2003: 321-362.
- 2003b BORER, H., "The grammar machine", in: Alexiadou et al. 2003: 288-331.
- 1996 BUCCELLATI, G., *A Structural Grammar of Babylonian*, Wiesbaden: Harrassowitz.
- 1994 CAMARAZZA, A. / HILLIS, A. / LEEK, E.C. / MIOZZO, M., "The organization of lexical knowledge in the brain: Evidence from category- and modality-specific deficits", in: L. A. Hisrchfeld / S.A. Gelman 1994: 68-84.
- 2003 CHIERCHIA, G., "A semantics for unaccusatives and its syntactic consequences", in: Alexiadou et al. 2003: 22-59.
- 1981 CLARK, E.V., "Lexical innovations: how children learn to create new words", in: W. Deutsch (ed.), *The child's construction of language*, London: Academic Press: 299-238.
- 1989 COMRIE, B., *Language universals and linguistic typology*, ²Oxford: Blackwell.
- 1996 CORRIENTE, F., *Introducción a la gramática comparada del semítico meridional*, Madrid: Consejo Superior de Investigaciones Científicas.
- 2003 DEL OLMO LETE, G., *Questions de linguistique sémitique. Racine et lexème. Histoire de la recherche (1940-2000) (Antiquités Sémitiques 5)*, Paris: Jean Maisonneuve.
- 2003 DEUTSCH, A. / FROST, R., "Lexical organization and lexical access in a non-concatenated morphology", in: Shimron 2003: 165-186.
- 1987 DROMI, E., *Early lexical development*, Cambridge: Cambridge University Press.
- 1994 FELDMAN, L.B. / BENTIN, S., "Morphological analysis of disrupted morphemes: Evidence from Hebrew", *Quarterly Journal of Experimental Psychology* 47: 407-435.
- 1995 FELDMAN, L.B. / FROST, R. / PENINI, T., "Decomposing words into their constituent morphemes: Evidence from English and Hebrew", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21: 947-960.
- 2003 FOX, J., *Semitic noun patterns* (Harvard Semitic Studies 59), Winona Lake: Eisenbrauns.
- 2001 GENTNER, D. / BORODITSKY, L., "Individuation, relativity, and early word learning", in: Bowerman / Levinson 2001: 215-256.

- 1994 GODDARD, C. / WIERZBICKA, A., *Semantic and lexical universals: Theory and empirical findings (Studies in Language Companion Series 25)*, Amsterdam / Philadelphia: John Benjamins Publishing Company.
- 1994 GOLDENBERG, G., "Principles of Semitic word-structure", in: Goldenberg / Raz 1994: 29-64.
- 1994 GOLDENBERG, G. / RAZ, S. (eds.). *Semitic and Cushitic studies*, Wiesbaden: Harrassowitz.
- 2002 GOLDENBERG, G., "Semitic linguistics and general linguistics", in: Izre'el 2002: 21-41.
- 2003 GORAL, M. / OBLER, L.K., "Root-morpheme processing during word recognition in Hebrew speakers across the adult life span", in: Shimron 2003: 223- 242.
- 1951 HARRIS, Z.S., *Structural Linguistics [Formerly entitled Methods in Structural Linguistics, 1951]*, Reprint edition 1986; Chicago / London: The University of Chicago Press.
- 1987 HEATH, J., *Ablaut and ambiguity: Phonology of a Moroccan Dialect*, Albany NY: SUNY Press.
- 2003 HEATH, J., "Arabic derivational ablaut, processing strategies and consonantal 'roots'", in: Shimron 2003: 115-129.
- 1994 HIRSCHFELD, L. A. / GELMAN, S. A. (eds.), *Mapping the mind: Domain-specificity in cognition and culture*, Cambridge: University Press.
- 1992 HOBERMAN, R. D., "Formal Properties of the Conjugations in Modern Aramaic", in: G. Booij / J. van Marle (eds.), *Yearbook of Morphology 1991*, Dordrecht: Kluwer: 49-64.
- 1997 HOBERMAN, R. D., "Modern Aramaic Phonology", in: A. S. Kaye (ed.), *Phonologies of Asia and Africa*, Winona Lake, IN: Eisenbrauns: 313-335.
- 2003 HOBERMAN, R. D. / ARONOFF, M., "The verbal morphology of Maltese", in: Shimron 2003: 61-78.
- 2002 HUEHNERGARD, J., "Comparative Semitic linguistics", in: Izre'el 2002: 119-150.
- 2002 IZRE'EL, S. (ed.) *Semitic linguistics: The state of the art at the turn of the twenty-first century (Israel Oriental Studies 20)*, Winona Lake: Eisenbrauns.
- 1991 KAYE, A. S. (ed.), *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday, vols I / II*, Wiesbaden: Harrassowitz.
- 2001 KIENAST, B., *Historische semitische Sprachwissenschaft. Mit Beiträgen von E. Graefe (Altaegyptisch) und G.B. Gragg (Kuschitisch)*, Wiesbaden: Harrassowitz.
- 1977 LYONS, J., *Semantics* (1, 2), Cambridge etc.: Cambridge University Press.

- 1967 MARTINET, A., *Éléments de linguistique générale*, ³Paris: Armand Colin.
- 1981 MCCARTHY, J. J., "A prosodic theory of nonconcatative morphology", *Linguistic Inquiry* 12: 373-418.
- 1982 MCCARTHY, J. J., *Formal problems in Semitic phonology and morphology*, Bloomington: Indiana University Linguistics Club [= Ph.D. Dissertation, MIT, New York: Garland Press 1979].
- 1983 MCCARTHY, J. J., "Prosodic templates, morphemic templates, and morphemic tiers", in: H. van der Hulst / N. Smith (eds.), *The Structure of Phonological Representations, Part I (Linguistic Models 2)*, Dordrecht: Foris Publications: 191-223.
- 1990 MCCARTHY, J. J. / PRINCE, A., "Foot and word in prosodic morphology: The Arabic broken plural", *Natural Language and Linguistic Theory* 8: 209-283.
- 1988 MICELI, G. / CAMARAZZA, A., "Dissociation of inflectional and derivational morphology", *Brain & Language* 35: 24-65.
- 2000 PRUNET, J.-F. / BÉLAND, R. / IDRISSE, A., "The mental representation of Semitic words", *Linguistic Inquiry* 31: 609-645.
- 2003 RAVID, D., "A developmental perspective on root perception in Hebrew and Palestinian Arabic", in: Shimron 2003: 293-319.
- 2003 ROSE, S., "The formation of Ethiopian Semitic internal reduplication", in: Shimron 2003: 79-97.
- 1991 SCHRAMM, G.M., "Semitic morpheme structure typology", in: Kaye 1991 (II): 1403-1408.
- 2003 SCHWARZWALD, O. (R.), "Opacity in Hebrew word morphology", in: Shimron 2003: 147-163.
- 2003 SHIMRON, J. (ed.), *Language processing and acquisition in languages of Semitic, root-based, morphology (Language Acquisition and Language Disorders 28)*, Amsterdam / Philadelphia: John Benjamins Publishing Company.
- 1999 USSISHKIN, A., "The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence", *Phonology* 16: 401-442.
- 1988 VOIGT, R., *Die infirmen Verbaltypen des arabischen und das Biradikalismus-Problem (Akademie der Wissenschaften und der Literatur: Mainz. Veröffentlichungen der Orientalischen Kommission 39)*, Wiesbaden / Stuttgart: Steiner.
- 2002 VOIGT, R., "The Hamitic connection: Semitic and Hamitosemitic", in: Izre'el 2002: 265-290.
- 2004 WUNDERLICH, D., "Why assume UG?", *Studies in Language* 28: 615-641.

1992 WIERZBICKA, A., *Semantics, culture and cognition: Universal human concepts in culture-specific configurations*, New York: Oxford University Press.

1994 WIERZBICKA, A., "Cognitive domains and the structure of the lexicon", in: Hirschfeld / Gelman 1994: 431-452.

1997 WIERZBICKA, A., *Understanding cultures through their key words*, Oxford: Oxford University Press.