

2. DECOMPOSING LEXICAL CATEGORIES IN SAKHA

In this chapter we will argue the following:

- (1) There are primitive nouns, verbs and adjectives (sections 2.1 and 2.2)
- (2) Word formation (WF)¹ is possible both in the lexicon and in the syntax (section 2.3)
- (3) Properties of lexical WF:
 1. LWF results in meaning change
 2. LWF can manipulate argument structure by deleting, adding or modifying a θ -role
 3. Thus, there are two types of lexical WF: argument-structure-manipulating and argument-structure-preserving
 4. AS-manipulating LWF can only apply if a concept is associated with a θ -grid, hence AS-manipulating LWF can take as its input verbs and adjectives but not nouns
 5. AS-preserving lexicon-internal manipulation can take as its input any concept irrespective of the presence or absence of a θ -grid
- (4) Properties of syntactic WF:
 1. SWF does not induce meaning change
 2. SWF can only add an argument, it cannot modify or eliminate the original θ -feature clusters (=arguments)
 3. SWF can also apply while preserving the original argument structure which results e.g. in syntactic nominalizations

In this chapter we will concentrate on the morphological aspect of the problem of category resolution which, in turn, has two facets. One is more general: if there are three categories, the presence of three-way asymmetries in morphological marking is expected and the above assumptions can be defended on morphosyntactic grounds. The second is more specific: if there are three categories, there must be three types of markers, for each category, i.e. nominalizers, verbalizers and adjectivizers. It will be shown that the first facet indeed holds true but not so for the second. One of the outcomes of the chapter will be that due to suffixal ambiguity there are no clear-cut morphological means of accounting for the existence of three categories and therefore a more optimal solution is to derive categorial distinctions in terms of thematic properties. A consideration of thematic properties of the three categories is the subject of subsequent chapters.

¹ The terms ‘word formation’ and ‘derivation (both lexical and syntactic)’ will be used interchangeably. The term ‘concept formation’, on the contrary, will be restricted to apply to lexical derivation.

As is clear from above, only one criterion is invoked in distinguishing derivation from inflection – that of meaning change, conditioned by viewing the lexicon as the locus of new concept formation, as is done here. It will be shown in section 2.3 that such approach avoids the many inconsistencies of other criteria² whose relevance is moreover not clear. Thus, the approach entertained here can be summed up by the following quotation from Beard (1998:64): “Derivational morphology differs from inflectional morphology in that it provides new lexical names for objects, relations, and properties in the world”.

The chapter has the following structure. In section 2.1 root and suffix inventories of Sakha will be investigated. The results support the existence of a three-way distinction among the basic/primitive lexical categories. Section 2.2 will argue against the view that roots can become nouns, verbs or adjectives by virtue of being embedded under/attaching appropriate category-determining heads/affixes. Section 2.3 will offer various kinds of evidence against the DM-tenet “all derivation in the syntax”. Sections 2.4 and 2.5 will explore, respectively, lexical and syntactic derivation of nouns and adjectives. Section 2.6 concludes.

2.1. The inventory of roots and suffixes

Appendices 1-2 list a sample of primary/basic/underived words from Sakha, i.e. words which do not contain a base capable of functioning as an independent word. The morphology of these primary words sanctions dividing them into two classes: one consisting of bound root plus suffix combinations and another consisting of roots which are used as independent words. A bound root is a root which cannot be used by itself and needs a suffix in order to become a noun, verb or adjective: e.g. $\sqrt{\text{il}}$ - which becomes the noun *ilii* ‘hand’ when suffixed with the nominalizing suffix –YY; on the contrary, $\sqrt{\text{tar}}$ - and $\sqrt{\text{bos}}$ - become, respectively, a verb and an adjective → *taraa* ‘to comb’, *bosxo* ‘free’. A root word is a root which can be used by itself as a noun, verb or adjective; e.g. *ap* ‘noun: magic’, *as* ‘verb: open’, *teŋ* ‘adjective: equal’. The fact that root words already have either nominal, adjectival or verbal category but not that of prepositions or adverbs and that bound roots also acquire one of these three categories after suffixing is consistent with Baker’s framework but argues against the categorial feature approach which yields a four-member (N/V/A/P) system.

However, within the Theta system we would expect that a given concept can only be categorized in one way because its categorization depends on its thematic properties which, in turn, depend on the causal relations into which the concept enters. If one and the same concept in one case were associated with no feature cluster (thus becoming a noun), in another case with one feature cluster (becoming an adjective) while in yet another case with two feature clusters resulting in a verb, then the theory would be rendered void. If we take a look at the distribution of roots across the three categories, we can see that the theory is supported by the existence of 1) root words which are primarily nouns; 2) root words which are primarily verbs; 3) root words which are primarily adjectives; 4)

² Listed and discussed e.g. in Stump 1998; see section 2.3 below.

bound roots which primarily become nouns after suffixing; 5) bound roots which primarily become verbs after suffixing and 6) bound roots which primarily become adjectives after suffixing. However, this is not the end of the story. Given a double classification of roots and a triple classification of categories, we get 14 logically possible combinations as outlined in (5) and (6). Out of these, the only combination which is not encountered is that of root words which are capable of functioning simultaneously as nouns, verbs and adjectives. Cases which are problematic for TS are boldfaced.

- (5) Bound roots:
1. Bound roots which become nouns
 2. Bound roots which become verbs
 3. Bound roots which become adjectives
 4. **Bound roots which can become nouns, verbs and adjectives**
 5. **Bound roots which can become nouns and verbs**
 6. **Bound roots which can become nouns and adjectives**
 7. **Bound roots which can become verbs and adjectives**
- (6) Root words:
1. Root words which become nouns: nominal roots
 2. Root words which become verbs: verbal roots
 3. Root words which become adjectives: adjectival roots
 4. Root words which are ambiguous between nouns, verbs and adjectives
 5. **Root words which are ambiguous between nouns and verbs**
 6. **Root words which are ambiguous between nouns and adjectives**
 7. **Root words which are ambiguous between verbs and adjectives**

Before these classes are addressed in more detail in section 2.1.3, we would like to consider the nature of roots in general and the nature of the suffixes which we encounter in the appendices.

2.1.1. The evolutionary routes of roots

It is possible to think of bound roots as having a rather general meaning and as capable of going in different categorial directions which would result in the modulation of the basic conceptual meaning. Depending on how flexible/rigid a given conceptual meaning is, one would account for the existence of seven possibilities in (5). Thus, a root like $\sqrt{\text{t}ü\text{ö}l\text{-}}$ can only become a noun resulting in *tüölbe* ‘remote place; dialect’ whereas the meaning of the root $\sqrt{\text{xapta-}}$ can be adjusted so as to fit all three categorial meanings: *xaptahyn* ‘flat wooden board’; *xaptaj/xaptat* ‘become/make flat’; *xaptaqaj* ‘flat’ and so on for the other classes in (5). This would be the track taken by approaches like DM which take the lexicon to be a list of bare roots. However, as argued in chapter 1 (section 1.5.2), coercing a single concept to fit three types of meaning violates conceptual atomism (to which Bare Root theories are most likely to subscribe). On the contrary, the route we are going to follow which is made possible within the Theta system does not require

conceptual coercion (as discussed in 1.5.2). We will argue that a given concept can only be categorized one way and that other categorial meanings are derived from that basic one.

A derivational view does not deny the fact that concepts expressed by bound roots have a rather general meaning. For instance, compare the root $\sqrt{\text{byr-}}$ which connotes something small and useless and becomes the noun *byrdax* ‘mosquito’ after attaching the suffix –TA-x. A variant of this root $\sqrt{\text{byt-}}$ is encountered in words like *bytaryj* ‘crumble up’, *bytarxaj* ‘small, minute’, *bytyryys* ‘scrap’, *bytahyt* ‘mouse’ (all marked by unpredictable morphology). Sometimes a single root is associated with only one independent word, sometimes it can give rise to a number of different phonologically well-formed words which belong to the same or distinct categories and the meanings of which are distantly related. However, it is never the case that a bound root by itself can express an independent meaning. Usually, the meaning of a bound root is too general to draw with any certainty a conclusion as to what it refers to in the world outside our minds: whatever a bound root expresses boils down to some implication or connotation. Consider, for instance the triplet *köñül* ‘freedom’, *köñöö* ‘restrict, limit’ and *köñös* ‘greedy’ which all involve $\sqrt{\text{köñ-}}$. As can be gleaned from the meanings of the independent words and not from $\sqrt{\text{köñ-}}$ itself, the latter implies something which has to do with freedom or restriction thereof. In such cases it is hardly plausible to claim that $\sqrt{\text{köñ-}}$ is a root listed in the lexicon which expresses a vague meaning having to do with freedom and when the root is submitted to the syntactic computation and becomes embedded under a category-determining functional head, this meaning gets modulated. Note that all three members of the triplet involve completely productive and regular morphology: /-AA/ as a verbalizer, /-l/ as a nominalizer and /-s/ as an adjectivizer.

Therefore we will not claim that bound roots with extremely general meanings like “something which has to do with this or that” are listed in the lexicon which seems to be a common assumption in DM which is forced to admit general meanings in order to accommodate concept stretching/coercion³. If the bound root $\sqrt{\text{köñ-}}$ were listed as a lexical item, it would have to correspond to a concept, hence have some denotation. However, this denotation would have to be calculated as a common semantic denominator behind the three members of the above triplet, namely, ‘freedom’, ‘restrict, limit’ and ‘greedy’, which is highly implausible. Within the extended Theta system there is no need to extract a common meaning from the meanings of the members of noun-verb, noun-adjective or verb-adjective doublets or noun-verb-adjective triplets. Therefore we will assume that it is not bound roots (see appendix 1) which are listed in the lexicon but the words to which they give rise, i.e. root-suffix combinations. One possible exception to this may be the root $\sqrt{\text{xan-}}$ or its variant $\sqrt{\text{xaj-}}$ which occurs consistently in a number of wh-words, cf. (7). However, since here we are dealing with a member of the functional lexicon, i.e. a lexicon

³ Cf. Marantz 2001b for whom $\sqrt{\text{STROY}}$ is a manner root that underlies such verbs as *destroy*, *construct*, *instruct*, *obstruct*, *restructure* all of which are derived in the syntax as a result of incorporating to the root $\sqrt{\text{STROY}}$ an appropriate particle the role of which is to introduce an external argument. The meanings of the verbs arise as a result of an interaction between the functional particle and the semantics of $\sqrt{\text{STROY}}$.

which contains functional elements, no problems are presented for the lexicon which lists concepts.

(7) The bound wh-root $\sqrt{\text{xaj-}}/\sqrt{\text{xan-}}$ ⁴

Suffixes	-AA	-A	-Dax	-LYk	-TA	-Tan
xaj-/xan-	Xajaa	Xaja	Xajdax	Xannyk	Xanna	Xantan
Gloss	Do what	Well; what now	How	Which	Where	From where

Note that these considerations do not apply to root words which already have an independent usage and can be easily equaled with concepts/roots listed in the lexicon. Thus, the following conclusion emerges: the (conceptual) lexicon lists roots which must correspond to concepts which must have meanings/denotations (note that the term ‘meaning’ is not used in the technical sense, as *intension*). This conclusion has implications for extended TS because it offers a way to account for the problematic cases in (5) and (6). Before this topic is taken up in section 2.1.3, we will consider evidence for the existence of bound roots and the nature of suffixation in Sakha.

2.1.1.1. Evidence for bound roots

The seven classes of bound roots listed in (5) are real. In order to establish that we are dealing with a bound root plus an affix combination a number of dictionaries of Sakha and related Altaic languages were consulted (see references). Whereas in classes 4 through 7 it makes perfect sense to postulate bound roots, classes 1, 2 and 3 might raise some questions as to the necessity of extracting a bound root. One might object that, since in these cases a root has only way to go, classes 1, 2 and 3 of (5) may be collapsed with classes 1, 2 and 3 of (6). Such collapsing may also be justified by the assumption we have made in the previous section that in classes 1, 2 and 3 of bound roots it is not the bound root itself which is listed in the lexicon but the bound root plus suffix combination. For instance, *ahyy* ‘molar’ ($\sqrt{\text{ah-}}$ plus the suffix -YY) will be listed as it is and not as $\sqrt{\text{ah-}}$. However, apart from historical and typological considerations which lie outside present concerns, extracting bound roots in classes 1-3 is important because, as we hope, it represents the real picture better. The arguments which follow are valid not only for classes 1-3 in (5) but they offer evidence for the psychological salience of bound roots in general.

⁴ The following triplet can also be considered as containing a bound root $\sqrt{\text{ki-}}$ which refers to human referents: *ki-m* ‘who’, *ki-hi* ‘human, person’, *ki-ni* ‘he/she (personal pronoun: human reference only)’. The suffixes -m and -hY do function as nominalizers; but the suffix -ni does not seem to be encountered elsewhere. One more triplet is presented by *tuo-x* ‘what’, *tuo-j* ‘say what; blabber; sing’ and *tuo-hu* ‘witness, evidence’. Again, the suffixes are possible nominalizers and verbalizers. The root $\sqrt{\text{tuo-}}$ may be related to the question particle *duo* (given that t/d-alternations in stem-initial consonants appear elsewhere). In these two triplets involving $\sqrt{\text{ki-}}$ and $\sqrt{\text{tuo-}}$ we are also dealing with functional elements, e.g. the feature [human] for $\sqrt{\text{ki-}}$ and another [wh] for $\sqrt{\text{tuo-}}$. *Tuohu* can be excluded from the second triplet: it probably got reanalyzed as a concept with an independent meaning of its own and is now a full-fledged member of the conceptual lexicon.

First of all, the separate status of classes 1-3 reflects dialectal differences among the speakers of Sakha: in different dialects bound roots from the first three classes sometimes appear with the same or related meaning as in the literary language but with a different suffix. The same applies to bound roots from other classes. The table below gives some examples.

(8) Dialectal differences in Sakha concerning suffix attachment to bound roots

Literary	Gloss	Dialectal	Gloss
Bal-ys	Younger sibling	Bal-ty	Younger sibling
Ad-aqa	1) Foot restrainers for horses 2) Burden, load	Ad-al-qa (also: ad-ar-qa-na, ad-yl-qa)	1) Foot restrainers for horses 2) Wooden tool which resembles <i>adaqa</i> 'foot restrainers for horses'
Xaax-ynaj	Species of fish	Xaax-a	1) The same species of fish 2) Dried fish
Suor-at	Yoghurt	Suor-qa (also: suor-aka)	1) Yoghurt 2) Brain, bone marrow
Ardj-aa-x	Large fish-trap with a wide opening	Ardj-aŋ	Large fish-trap with a wide opening
Köl-ö	Working cattle	Köl-gö (also: köl-gö-m)	Working cattle
Mannj-a	Favour	Mannj-y	Benevolence
		Mannj-yl-ba	Reward
Il-ij	Become wet	Il-ime-x	Wet sponge
		Il-ip-te	Last year's wet grass
Meen-e	Futile, purposeless, loose	Men-ee-k	Wild reindeer
Men-ik	Naughty	Men-ex	Musk deer

Second, bound root extraction (in all classes) shows that there is nothing special about the suffixes which attach to bound roots: they are drawn from the basic suffix inventory (see section 2.1.2).

Third, extracting bound roots is also important from a comparative point of view (with respect to Altaic/Turkic studies). For instance, a perusal of Sakha and Turkish dictionaries yields a large number of correspondences between the two languages when one and the same bound root with the same/related meaning (modulo morpho-phonological modifications) appears in both languages with related or different suffixes. (9) presents just two of such examples.

- (9)
- | | |
|------------|--|
| Böl-öx (S) | Group; division, compartment; section ... |
| Böl-ük (T) | Part, division; compartment; group ... |
| Böl-üm (T) | Portion, slice, division, chapter, part, section ... |
| O-ho-x (S) | Stove, oven, furnace |

O-ca-k (T) Fireplace, hearth, oven, furnace ...

A fourth piece of evidence for bound root vitality concerns degree intensification with adjectives and therefore has restricted applicability (for class 3 only). Adjectives in Sakha can be subjected to two types of reduplication – either of the initial syllable or of the whole word – both of which have the same degree intensifying effect. It is the latter type – full reduplication which is relevant for the present argument. When full reduplication takes place, it can detect internal structure if the adjective is a bound root plus a suffix and modify the vowel of the suffix (third row in (10)). The vowel of the root cannot be modified which explains the impossibility of a third row in (11): adjectival root words are pure roots which are capable of becoming adjectives directly, without the mediation of any suffixes.

- (10) Degree modification with bound root adjectives (class 3 in (5)):
- | | | |
|---------------|----------------------|--|
| A√-suffix | √-suffix-√-suffix | √-suffix _{MODIFIED} -√-suffix |
| <i>Ücügej</i> | <i>Ücügej-ücügej</i> | <i>Ücügij-ücügej</i> |
| ‘good’ | ‘very good’ | ‘very-very good’ |
| <i>Cugas</i> | <i>Cugas-cugas</i> | <i>Cuguus-cugas</i> |
| | | <i>Cugyys-cugas</i> |
| ‘close’ | ‘very close’ | ‘very-very close’ |
- (11) Degree modification with independent root adjectives (class 3 in (6)):
- | | | |
|-------------|------------------|---------------------------|
| A | A-A | *A _{MODIFIED} -A |
| <i>Ciŋ</i> | <i>Ciŋ-ciŋ</i> | * <i>Ciiŋ-ciŋ</i> |
| ‘dense’ | ‘very dense’ | |
| <i>Kieŋ</i> | <i>Kieŋ-kieŋ</i> | * <i>Kiiŋ-kieŋ</i> |
| ‘wide’ | ‘very wide’ | |

A fifth piece of evidence for the vitality of bound roots comes from the restricted use of some bound roots in modern language. This is typically possible inside fixed expressions. For example, √*kys-* is a bound root occurring inside the following independent words – *kyhyn* ‘winter’, *kystyk* ‘winter house’, *kystaa* ‘spend winter’. It can also be used by itself in three contexts: *kys xaar* ‘snow which will stay until May (lit. √winter snow)’, *kys mas* ‘wood prepared in the autumn to last for the whole winter (lit. √winter wood)’ and *kys buolla* ‘winter began (lit. √winter become-past.3)’.

A final piece of evidence is considered in the next section.

2.1.1.2. Evidence from word coining for the psychological salience of bound roots

That bound roots are psychologically salient is supported by the fact that word coining can make use of them. Inventing neologisms not only from independent words but also from roots which are otherwise bound and appear embedded under (often highly idiosyncratic) suffixes emphasizes the ability of native speakers to detect and extract these bound roots. This ability has proven to be quite useful in the

recent years when a number of new words have been coined to replace loanwords: a bound root is extracted and combined with a different suffix or combination of suffixes than previously accepted resulting in a new word. Such new words appear regularly in newspaper language⁵. Whether they will be eventually incorporated into the conservative literary language and accepted by all native speakers, time will show. One such word which has been invented quite recently and already accepted is *oŋ-ku-l* ‘plan, scheme, model’ derived with the suffixes /-KI/ and /-l/ from the bound root $\sqrt{oŋ}$ - which forms the basis of *oŋor* ‘to make’ and *oŋohuk* ‘thing made’. The table below lists five more fully integrated neologisms which are no longer considered controversial.

(12) Neologisms derived from bound roots

Existing words	Gloss	Neologism	Gloss
Ol-ox	Life; seat	Ol-omto	Basis, foundation
Ol-or	Live; sit		
Möl-bö-j	Become large	Möl-üü-k	Profit
Ül-ler	Divide, partition, distribute	Ül-ük	Share, stock
Ül-ex	Reservation, contract, order		
Ura-n	Refined, exquisite	Ura-mnjy	Art
Buo-j	Forbid, prohibit	Buo-qu	Prohibition

It is important to mention here that we disagree with Beard 1998 who makes a distinction between grammatically determined derivation and conscious, extragrammatical irregular phenomena which may look like derivation but are not subject to principles of grammar. The processes which Beard treats as extragrammatical are listed in (13); they all expand the lexical stock of a language, just like derivation.

- (13) **Back formation:** difficulty → difficult, sculptor → sculpt
 Clipping: telephone → phone
 Blends: smoke + fog → smog
 Acronymization: laser, scuba
 Analogical formation: workaholic, shopaholic, fishburger
 Borrowing: sputnik, thug
 Loan translation
 Commonization: aspirin, quisling
 Semantic narrowing: percolator, escalator
 Folk etymology: crawfish

⁵ An informal inquiry on word coining was conducted at the Viluysk Pedagogical College by I. Vinokurov in March 2003: 20 students were asked to invent native substitutes for 15 loanwords from Russian. The students came up with 75 new words, of which 17 were minted on the basis of bound roots, the rest were derived from independently existing, phonologically well-formed words.

Out of these mechanisms, the one most relevant to our present concerns is back formation when a word is (falsely) analyzed as consisting of a ‘root’ plus a ‘suffix’ and the ‘root’ is extracted and made into an independent word. For Beard this is an extragrammatical process: cf. “The point is that these processes tend to be conscious, extragrammatical, and hence grammatically irregular. Rather than filling a position in some lexical paradigm, they create new lexical bases which then generate their own paradigms” (ibid. p. 57). However, the point of excluding this type of derivation from the grammar is not clear. A sentence like *I saw an interesting movie yesterday* is also created intentionally: one pronounces it if one intends to say what it means. Yet this sentence is not created extra-grammatically. This type of argument in fact demonstrates that the boundary between the lexicon and syntax should not be drawn in terms of transparency, regularity or consciousness (see 2.3 below).

2.1.2. Simplex and complex suffixes

A study of suffixation in Sakha shows that the basic inventory consists of simplex suffixes and that all complex suffixes are made up of combinations of simplex ones. Examples of simplex suffixes are reflexive/passive /-n/; reciprocal/comitative /-s/; verbalizing /-AA/; nominalizing /-k/, /-l/, /-m/, /-x/; etc. Examples of complex suffixes are verbalizer /-LAA/ combining /-l/ and /-AA/; adjectivizer /-LAAx/ combining the previous suffix /-LAA/ plus /-x/; nominalizer /-AAx/ containing /-AA/ and /-x/ and so on.

The preceding examination of bound roots showed that in order to become an independent, full-fledged word a bound root needs to be saturated by an affix. There are roots which require just one, simplex suffix and there are roots which require a complex suffix which combines a number of simplex suffixes (usually two but can be more). Whenever we have categorial doublets or triplets built on the same bound root, e.g. noun-verb or verb-adjective pairs, it is impossible to draw any inferences about the direction of derivation from the composition of suffixes since this composition is highly idiosyncratic. For instance, a given bound root can evolve into two distinct categories and both evolutionary routes may extend over the same distance (measured in terms of suffixes). To consider a specific case, the pairs *küüs* ‘N: strength’ – *küür* ‘V: strain/exert oneself’ and *ajdaan* ‘N: noise, uproar’ – *ajdaar* ‘V: be noisy, make a fuss’ are derived from the bound roots $\sqrt{\text{küü-}}$ and $\sqrt{\text{aj-}}$. In the former case there is one simplex suffix: -s for the noun, -r for the verb. In the latter case the suffix is complex involving three simplex ones, the first two of which are identical for both the noun and the verb and only the last one is different: the pair *ajdaar/ajdaan* should be parsed as *aj-L-AA-n* versus *aj-L-AA-r*. The suffixes may also be homonymous. The bound root $\sqrt{\text{san-}}$ can become both a noun and a verb and both categorizations involve the same simplex suffix –AA: *sanaa* ‘N: thought’ versus *sanaa* ‘V: think’. Homonymous suffixes can also be complex as in the case of *kujaar* ‘N: outer space’ versus *kujaar* ‘V: fly up, soar’ derived from $\sqrt{\text{kuj-}}$ with the help of –AA-r.

However, sometimes one of the evolutionary routes may be longer. The root $\sqrt{\text{xal-}}$ related to the concept of the sky has the primary potential of becoming

either a verb or a noun. The former road requires only one suffix resulting in *xalyn* ‘clear up (about the sky)’. On the contrary, becoming a noun, *xal-l-aa-n* ‘sky’, requires three steps – the suffixes /-l/, /-aa/ and /-n/ and there is no cut-off point in between: **xal-yl*, **xal-laa* are not possible words. The reverse holds in the pair *elie* ‘N: kite (bird)’ – *eleer* ‘V: kite (go in a rapid, flighty manner)’ derived from $\sqrt{\text{el-}}$ where becoming a verb requires one step (i.e. one suffix) more than becoming a noun.

These and other examples show that one cannot attach much value to the nature of $\sqrt{\text{-}}$ -suffixation when it comes to determining which member of a $\sqrt{\text{-}}$ -based doublet/triplet is the basic one and which ones are derived.

2.1.2.1. Suffixal promiscuity

Two types of suffixal promiscuity can be distinguished. The first has been analyzed in DM as reflecting different levels of attachment: English *-er*, for instance, can attach either right above the root categorizing the latter as a noun or above a verbalizing functional head. Low level of attachment is shown by *donor*, *rotor*, *debtor*, *malefactor*, *benefactor* whereas in *donator*, *rotator*, *driver*, *writer* the affix is attached high. Under DM, the possibility of different attachment levels is conditioned by the existence of two different places of word formation: the domain of the root (when a particular affix attaches either below a categorial head or as a categorial head itself) and the domain above categorial heads. Within the Theta system, this first type of suffixal promiscuity would also be analyzed as reflecting two different domains of word formation but the relevant opposition will be that of lexicon versus syntax. The occurrence of suffixes across domains will be considered later in sections 2.3-2.5.

More important for the issue at hand, namely, deriving the existence of three basic/primary categories, is the second type of suffixal promiscuity when one and the same suffix is used across categories. For instance, if we take a look at the bound roots in appendix 1, we encounter some suffixes saturating bound roots in all three categories. The suffixes are listed in (14). Assuming, as we have done in section 2.1.1, that bound roots are psychologically salient and that native speakers can detect the $\sqrt{\text{-}}$ -suffix internal structure in words built from bound roots, one must answer the question of what drives successful categorization in these cases if morphological information cannot be relied on.

(14) Ambiguous categorial suffixes appearing on bound roots

Nouns	Verbs	Adjectives
$\sqrt{\text{-AA/-YA}}$	$\sqrt{\text{-AA}}$	$\sqrt{\text{-AA/-YA}}$
$\sqrt{\text{-j}}$	$\sqrt{\text{-j}}$	$\sqrt{\text{-j}}$
$\sqrt{\text{-l}}$	$\sqrt{\text{-l}}$	$\sqrt{\text{-l}}$
$\sqrt{\text{-n}}$	$\sqrt{\text{-n}}$	$\sqrt{\text{-n}}$
$\sqrt{\text{-r}}$	$\sqrt{\text{-r}}$	$\sqrt{\text{-r}}$
$\sqrt{\text{-s}}$	$\sqrt{\text{-s}}$	$\sqrt{\text{-s}}$
$\sqrt{\text{-t}}$	$\sqrt{\text{-t}}$	$\sqrt{\text{-t}}$

√-x	√-x	√-x
-----	-----	-----

The same question arises for some of the suffixes listed in the table above when they are used in category-changing derivation from independent bases. The most ambiguous and therefore problematic is the suffix /-s/: it displays nominalizing, adjectivizing and verbalizing functions. As a nominalizer, it can attach to nouns, verbs and at least one adjective (15). As an adjectivizer, it attaches to verbs and a couple of nouns (16). As a verbal suffix, it has two uses: one is to mark reciprocal and comitative voices (see chapter 6) and another is to derive some verbs from other verbs as shown in (17)⁶.

(15) /-s/ as a nominalizer

Base	Gloss	Base-s → N	Gloss
Kün	N: Day	Künüs	Afternoon, midday
Üge	N: Fable	Üges	Tradition, custom; habit
Saņa	A: New	Saņas	Daughter/sister-in-law
Xamnaa	V: Move	Xamnas	Salary
Kyrbaa	V: To cut into pieces	Kyrbas	A piece
Xat	V: Twist, weave (a rope)	Xatys	Leather rope

(16) /-s/ as an adjectivizer

Base	Gloss	Base-s → A	Gloss
Aņardaa	V: Divide in two	Aņardas	Single, alone
Kuttaa	V: Frighten	Kuttas	Fearful
Kylgaa	V: Shorten	Kylgas	Short
Kömüskee	V: Defend	Kömüskes	Defensive, protective
Baaqnaa	V: Speak in a low voice	Baaqnas	Low (of voice)
Ös	N: Spite, hostility	Öhös	Stubborn
Alyp	N: Magic	Albas	Cunning (also noun: trick)

(17) /-s/ as a verbalizer

V	Gloss	V-s → V	Gloss
Baaj	Tie	Baajys	Carp/cavil at
Iir	Go mad	Iiris	Quarrel with
Yl	Take	Ylys	Undertake
Bier	Give	Beris	Share
Ketee	Watch, keep an eye on	Ketes	Wait

⁶ On verbs which allow idiosyncratic s-marking (i.e. neither reciprocal nor comitative or in addition to these two voice meanings) see chapter 6.

Sit	Catch up with	Sitis	Succeed, achieve
Tap	Hit (one's aim)	Tabys	Get along
Üŋ	Worship, revere	Üŋüs	Complain
Xas	Dig	Xahys	Rake/search through, ransack

It can be objected that s-nominalizer is not as productive as s-adjectivizer/verbalizer and therefore does not have to be taken into account. However, this would still leave unsolved the problem of ambiguity between /-s/ as an adjectivizer and /-s/ as a verbalizer. Second, it is not clear if productivity or lack thereof should be considered a serious threat. On the one hand, productivity is not an absolute notion (see discussion in section 2.3). On the other hand, suffixes which are treated as unproductive in Sakha (labeled in appendix 3 as sporadic) are often invoked in word coining. For instance, the suffixes which form *ura-mnjy* 'art' and *buo-qu* 'prohibition' in (12) are otherwise not productive at all. Also recall footnote 5 which mentions an informal inquiry on word coining conducted in Viluysk: the majority of 75 neologisms were derived with highly idiosyncratic and unproductive suffix combinations.

The suffixes /-x/ and /-k/ from (14) are also problematic because both of them function as extremely productive adjectivizers and nominalizers. Two more suffixes of interest are /-YY/ which derives nouns from verbs and /-A/ which derives adjectives from verbs (see appendix 3 for examples: (23) in section 2 and (21) in section 8). Since these two suffixes are less ambiguous than others, they would be likeliest contenders to claim the titles of a nominalizer and an adjectivizer. However, this would not solve the morphological aspect of the problem of category resolution (mentioned in the opening lines of this chapter) since, on the one hand, not all nouns and adjectives are marked with these two suffixes and, on the other, /-YY/ and /-A/ display promiscuity of the *donor/donator* type (i.e. across domain boundaries) by deriving simultaneous gerunds from verbs.

We can continue this morphological exercise but the upshot of the overall discussion will nevertheless be that, if we cannot rely on clear-cut morphological means of accounting for the existence of three categories, a thematic solution offered by the Theta system, as outlined in the introduction, is a better alternative.

2.1.2.2. Suffixal instability

One property of suffixes in Sakha is their instability. Take, for instance, the bound root \sqrt{y} - from which two independent words can be derived: *ynax* 'N: cow' and *ya* 'V: to milk'. The latter is derived with the verbalizer -AA and the former with the nominalizing suffix -na-x. *Ynax* 'cow' can be verbalized with the universal verbalizer -LAA \rightarrow *ynaxtaa* 'provide with cows; go looking for cows; etc.'. It can also be subjected to a different kind of verbalization which does not respect the integrity of *y-na-x* and derives a new verb by getting rid of the nominalizing suffix -x and replacing it with either -j or -t (both verbalizers) \rightarrow *y-na-j* '(intrans.) resemble a cow', *y-na-t* '(trans.) make resemble a cow'. The same thing happens with *u-ta-x* 'beverage' derived from *uu* 'water' with the suffix -ta-x: *u-ta-x* \rightarrow *u-ta-t* 'become thirsty'. In these cases we cannot say that *ynaj/ynat* and *utat* are derived from bound

roots with –j and –t: there are no bound roots *√yna- or *√uta-; the relevant bases are the bound root √y- and the nominal root word *uu* ‘water’. It is also impossible to claim that the verbs in question are derived from √y- and *uu* with the suffixes –na-j, –na-t or –ta-t because these are not possible verbalizers. /-na-j/, /-na-t/ and /-ta-t/ are only encountered in verbal contexts if there is a corresponding nominal context with the nominalizing suffixes /-na-x/ and /-ta-x/. Rather, in such contexts we are dealing with the instability of the nominal suffixes –na-x and –ta-x: their second (final) constituent suffix –x can be dropped and replaced with another one.

The unstable property of suffixes in Sakha can be explained by the agglutinative nature of the language. All complex suffixes are composed of simplex ones and the boundaries between suffixes are easily detected because of agglutination. Unlike fusion, agglutination does not create a strong link between the stem and the suffix by melting these two together: things glued fall apart more easily than things fused. Therefore it is not surprising that simplex suffixes can be dropped and replaced by others giving rise to rather unusual suffix combinations such as the ones mentioned above: /-na-j/, /-na-t/ and /-ta-t/.

To finalize this section on suffix instability, let’s consider two specific cases of verb to adjective conversion. The verbs *djölörüj* ‘get holes’ and *ildjirij* ‘fall to pieces’ both contain the complex suffix –ryj consisting of –r and –j. In the former case the suffix –ryj derives a new verb from *djöl* ‘make holes’; in the latter case it is attached to the bound root √ildji-. *Djölörüj* ‘get holes’ and *ildjirij* ‘fall to pieces’ can give rise to adjectives: *djölörkøj* ‘having holes’ and *ildjirkej* ‘falling to pieces’. Both adjectives are derived via suffixing –XA-j, a very productive adjectivizer. However, as we can see, the suffix –j of –ryj was forced to drop and replaced by –XA-j. Again, this shows that the verbalizer –ryj is not a stable suffix and word formation processes are sensitive to its constituent structure.

2.1.2.3. Not all derivation is suffixation

Since Sakha is an agglutinative language, almost all derivation involves suffixation. However, modification of a root vowel is also encountered which involves either diphthongization or vowel lengthening. For example, the word *aqabyyt* ‘priest’ is historically derived from *aqabyt* ‘our father (father-1pl)’. Eventually, *aqabyt* meaning ‘priest’ became separated from *aqabyt* ‘our father’, reanalyzed as an independent root and entered in the lexicon as such. To emphasize the fact that a root-plus-suffix combination has indeed been reanalyzed as a single new root, the vowel /-y/ of what was originally the first plural suffix /-byt/ became lengthened which is never possible in a genuine inflectional suffix. As a result, in modern language the original link between the two forms escapes many speakers.

There are also a number of other idiosyncratic cases where the nature of derivation involved is not straightforward suffixation. For instance, the verb *baaj* ‘to tie’ is apparently derived from the noun *bya* ‘rope’: here, in addition to j-suffixation, the vowel of the base is modified. Another unclear case is presented by the pairs of related words: *baj* ‘become rich’ – *baaj* ‘rich’ and *kur* ‘dry, old, stale’ – *kuur* ‘become dry’. Here it is not clear if we are dealing with derivation which involves

vowel lengthening or if one and the same bound root is realized in one case with a short vowel and in another with a long one.

2.1.3. Explaining categorial ambiguities of roots in Theta system

Two conclusions at which we have arrived in the course of the preceding discussion will help explain away the problematic cases boldfaced in (5) and (6). The first one is given in (18). The second concerns suffixal instability in Sakha.

- (18) The lexicon lists roots which correspond to concepts with independent meanings/denotations

First, let's consider bound roots restricting our attention to those which can become nouns, verbs and adjectives (appendix 1, section 4). Other classes of bound roots can be explained along the same lines. In class 4 there are ten cases to account for. (18) renders the following five triplets harmless for TS.

- (19) N/V/A bound roots denoting different concepts

#	Root	Noun	Verb	Adjective
1.	√kōŋ-	Kōŋül 'freedom'	Kōŋöö 'restrict, limit'	Kōŋös 'greedy'
2.	√bar-	Baryl 'sketch, outline'	Baryj 'dimly appear in the distance'	Baraan 'dark'
3.	√xar-	Xarys 'care, protection'	Xaraj 'hide away, take good care'	Xaram 'thrifty, economical'
4.	√taa-	Taabyryn 'puzzle'	Taaj 'guess'	Taamax 'enigmatic'
5.	√sys-	Syhyan 'relation, attitude'	Syhyar 'attach (trans.)'	Syhyamax 'sociable'
			Syhyn 'attach oneself to'	

In the five remaining cases we have four nouns (20) and one adjective (21) referring to concepts different from those to which, respectively, the corresponding V/A-pairs and N/V-pair refer.

- (20) N-concept versus V/A-concept

#	Root	Noun	Verb	Adjective
1.	√xapta-	Xaptahyn 'flat wooden board'	Xaptaj 'become flat'	Xaptaqaj 'flat'
			Xaptat 'make flat'	
2.	√njurgu-	Njurguhun 'snowdrop'	Njurguj 'be first, best, choice'	Njurgun 'first, best, choice'
			Njurgut 'glorify'	
3.	√subur-	Suburqa 'sheaf of wheat'	Suburuj 'stretch, extend, shoot up (intrans.)'	Suburxaj 'stretched, extended'

			Suburut ‘stretch, extend (trans.)’	
4.	√kyr-	Kyryy ‘furthest border, edge’	Kyyrat ‘throw very far; make soar up high’	Kyyraj ‘far away, distant’
			Kyyraj ‘soar up very far, high’	

(21) A-concept versus N/V-concept

Root	Adjective	Noun	Verb
√syl-	Sylaj ‘having bad appetite’	Sylaa ‘tiredness’	Sylaj ‘become tired’
			Sylat ‘make tired’

The question now is how to account for the cases which are not excluded by the principle in (18). For instance, in (20) we have a bound root giving rise to both verbs and adjectives and in (21) the same bound root seems to result in a noun and an adjective and, apparently, in all five cases the same concept underlies the different categories. For bare root theories these would be instances of concept stretching: they would assume that what is listed in the lexicon is e.g. √syl- or √kyr-. This option is precluded for the present framework, given (18): neither √syl-, nor any other bound roots denote anything by themselves. Therefore we have assumed above in 2.1.1 that it is not bound roots which are listed in the lexicon but the words to which they give rise, i.e. root-suffix combinations. With respect to (20) and (21) we will propose the following (anticipating the discussion to come in 2.3-2.5). The four V/A-pairs in (20) and one N/V-pair in (21) actually contain three members because the verb comes in transitive (=causative) and intransitive (=unaccusative) variants. Assuming the TS-analysis of causative/unaccusative alternations as discussed in chapter 1 (section 1.4.3) and as defended for Sakha in chapter 5, it is the transitive verbs in (20) and (21) which are listed in the lexicon and the five intransitives are derived by unaccusative reduction. The adjectives in (20) are derived from the intransitive verbs in the syntax by replacing the unstable suffix -j with the regular adjectivizer -XA-j or the idiosyncratic -n⁷ (on deriving adjectives from intransitive verbs see section 2.5). The noun in (21) is also derived from the intransitive verb but in the lexicon because, although V→N derivation can take place either in the lexicon or syntax (section 2.4), the latter must preserve the original argument structure of the verb which is not the case in (21): *sylaa* ‘tiredness’ has no arguments.

Now let’s consider root words (appendix 2). Three classes are problematic: root words which are ambiguous between nouns and verbs (class 5); root words which are ambiguous between nouns and adjectives (class 6) and root words which are ambiguous between verbs and adjectives (class 7). After the principle in (18) combs through these classes, we are left with only six cases.

⁷ The derivation *kyyraj* ‘V: soar up very far, high’ → *kyyraj* ‘A: far away, distant’ displays a promiscuous suffix.

(22) Root words which are ambiguous between nouns and verbs

#	Root Word	Noun	Verb
1.	Kös	Migration; a walking mile	Migrate, move from one place to another
2.	Saat	Shame	Be ashamed
3.	Tyyn	Breath	Breathe

(23) Root words which are ambiguous between verbs and adjectives

#	Root Word	Verb	Adjective
1.	Toŋ	Freeze	Frozen
2.	Tot	Eat one's fill, glut oneself	Well-fed, with a full stomach
3.	Köp	Fluff; rise, heave; become stronger (disease)	Fluffy

The ambiguous items in (23) are amenable to the same analysis as (20), namely, unaccusative reduction. The three intransitive verbs *toŋ* 'freeze', *tot* 'eat one's fill' and *köp* 'fluff, rise' all have irregular causatives which, as will be shown in chapter 5, is one of the diagnostics of establishing the causative verb as the basic entry in the lexicon from which unaccusative alternants are derived by reduction. Therefore in (23) the listed lexical items are transitive verbs *toŋor* 'make freeze', *totor* 'feed to the fullest, satiate' and *köbüt* 'make fluff, make rise' which, when their external argument is reduced, give rise to intransitive *toŋ*, *tot* and *köp*. The relevant adjectives *toŋ*, *tot* and *köp* are derived syntactically from these intransitive verbs. As for the three cases in (22), we will assume a V→N argument-structure reducing, hence lexical derivation. In these six cases the morphological markers involved turn out to be phonologically null. However, there is no proliferation of zero affixes and therefore the use of zero morphology in these contexts is not problematic from the point of view of learnability.

2.2. Against bare roots

In this section we will argue against an assumption maintained in Distributed Morphology and Borer's exo-skeletal framework that the lexicon lists only roots – meaning-sound pairings devoid of any other information including argument structure and categoriality. On the contrary, in TS the roots listed in the lexicon are associated with a particular number of arguments depending on their thematic properties and it is this argument structure information which determines whether a particular root will be passed on to the syntactic component as a verb, noun or adjective. In section 2.2.1 we will argue against the exo-skeletal (XS) model of bare roots, in 2.2.2 against the DM-model. Since both models of categorization have already been considered in the introduction (DM in section 1.3.2.1; XS in section 1.3.2.3) and some objections have been raised against both of them (against DM in

section 1.3.2.2; against XS in section 1.3.2.4), the following discussion will be quick.

2.2.1. Against XS-categorization

Categorization in the exo-skeletal framework can proceed along two different dimensions: either by morphological structure or by functional heads.

Morphological categorization would only be possible in a language which has an inventory of nominalizing, verbalizing and adjectivizing affixes (bound f-morphs) at its disposition. Therefore this type of categorization is language-specific and cannot hold in a language like Sakha where affixes (both inflectional and derivational) are syncategorematic (i.e. promiscuous), as argued in section 2.1.2.1.

Categorization by functional structure is also problematic because it creates circularity, as shown in 1.3.2.4.1. One cannot derive verbhood and nounhood from morphosyntactic properties of verbs and nouns such as occurrence with tense or determiners because it is precisely these morphosyntactic properties which need to be derived. Assuming that a root is turned into a verb in the environment of tense simply restates the basic descriptive generalization that verbs are those categories which can inflect for tense but does not explain it.

The exo-skeletal framework results in massive overgeneration because of its prediction that in principle all roots should behave like the English *form* which can be inserted in virtually any syntactic context – nominal, verbal as well as adjectival. As shown in section 1.3.2.4.4, the proposed means to curb overgeneration is highly unsatisfactory. It is assumed that those lexical items which are not flexible like English *form* or *siren* are listed in the lexicon as internally complex: for instance, if they have a rigid verbal behaviour, they contain verbalizing structure. Taking rigid lexical items to contain nominalizers, verbalizers or adjectivizers already in the lexicon contradicts the very foundation of the exo-skeletal model because the syntactic skeleton can no longer drive the interpretation of such complex entries.

2.2.2. Against DM-categorization

We have argued in section 1.3.2.2 that categorizing roots with the help of category-determining functional heads such as *n*, *v* and *a* has the same amount of explanatory power as the approach based on categorial features and turns out to be merely a translation of Chomsky's original (1970) account into syntactic terms: if for Chomsky 1970 what makes a lexical item a noun is a [+noun] feature, for DM what makes a noun is small *n*. Such an approach is clearly circular and in the end we have no independent account of what it is to be a noun, a verb or an adjective. Apart from this, DM raises another theoretical objection discussed in chapter 1, section 1.5.2, namely, the incompatibility of conceptual coercion adopted within DM with Fodor's conceptual atomism.

If we turn away from these conceptual problems to empirical matters, the resulting picture is also not very satisfactory because it results in a proliferation of zero affixes. Let's consider root words first (repeated below in (24) from (6)). In the first three classes, the roots will be marked with overt morphology in two of their

categorial uses whereas in the third incarnation there will be a zero affix: in class 1, for instance, verbalizing and adjectivizing suffixes will be overt whereas the nominalizers will be null. The same considerations carry over to the other classes in (24) as outlined. Proliferation of zero's will clearly raise problems for learnability. On the contrary, if we compare the DM-account with the TS-account discussed above in section 2.1.3, the latter does not result in the explosion of null morphemes: with respect to phonologically independent root words, only in six cases was an appeal made to zero morphology – i.e. derivation which was not appropriately marked with overt morphological markers⁸.

(24) Root words: the DM-account

1. Root words which become nouns: n – zero affix, v and a – overt affixes
2. Root words which become verbs: v – zero affix, n and a – overt affixes
3. Root words which become adjectives: a – zero affix, n and v – zero affixes
4. Root words which are ambiguous between nouns, verbs and adjectives: none
5. Root words which are ambiguous between nouns and verbs: n and v – zero affixes, a – overt affix
6. Root words which are ambiguous between nouns and adjectives: n and a – zero affixes, v – overt affix
7. Root words which are ambiguous between verbs and adjectives: v and a – zero affixes, n – overt affix

With respect to classes of bound roots which become either nouns, verbs or adjectives (see (5) above), the problem faced by DM will be even more severe for the following reason. Consider, for instance, the first class of bound roots which become nouns after attaching some affixes: the roots in this class are not phonologically independent words and therefore need an affix which would saturate them and turn them into a noun. For DM this would be a case par excellence when a root is inserted in the nominalizing context under a category-determining little *n*. The issue can be exemplified with the specific root *uruu* 'relative' (or any other root from appendix 1, section 1): the root in question is $\sqrt{\text{ur-}}$ which is saturated by the nominalizing suffix *-YY* (in word formation from independent bases this suffix functions as a very regular and productive nominalizer). However, consider what happens with the same bound root in other contexts: namely, verbalizing (*uruurqaa* 'treat/consider as a relative', *uruulas* 'become relatives', etc.) and adjectivizing (*uruuluu* 'related'). Whereas in the case of the corresponding class of root words (i.e. root words which become nouns without making use of overt nominalizers) we could say that a root word would become a noun with the help of a zero nominalizing affix while becoming a verb or adjective via overt verbalizers/adjectivizers, what happens in the case of bound roots from class 1 is quite problematic for DM because when a bound root from class 1 becomes a verb (or adjective), the original nominalizing suffix is not lost despite the fact that

⁸ DM will face an identical problem in the case of English flexible words like *form* or *siren* which can be inserted in virtually any context: in all of these cases there will be three types of zero affixes realizing three types of category-determining little functional heads.

suffixes in Sakha are unstable. Rather, this suffix is further accretioned with the overt verbalizer or adjectivizer resulting in a structural representation like $\sqrt{n-v}$ ($\sqrt{ur-uu-rqaa}$; $\sqrt{ur-uu-las}$) or $\sqrt{n-a}$ ($\sqrt{ur-uu-luu}$). This result clearly supports the assumption defended in this dissertation that such bound roots (\sqrt{n}) are primary nouns which clearly explains why they become $\sqrt{n-a}$ when adjectivized or $\sqrt{n-v}$ when verbalized. This argument is based on the presence of morphological asymmetries in categorial marking that follows from the impossibility of conceptual coercion. The latter ensures that there will be primary nouns, verbs and adjectives and that secondary nouns, verbs and adjectives will be derived from the primary ones – a process which will be overtly marked by derivational morphology. A similar argument but the other way around is raised in Baker (2003:295): if conceptual coercion can apply freely, what is the need for derivational morphemes which do not add new meanings but only shift categories?

Moreover, if we turn to the literature we find numerous arguments against not only zero morphology⁹ but also against viewing {noun-verb-adjective} triplets or doublets (where the same concept underlies different categorial incarnations) as lacking a direction of derivation. Don (1993) argues convincingly that conversion must be viewed as a directional process involving a base, a derivative and a zero morphophonological affix. The use of zero morphology in Don's model raises no objections from the perspective of learnability because it offers a number of independent ways to determine the direction of derivation, such as whether the conjugation class involved is regular or irregular, the type of nominal gender, phonological constraints on syllabification, consonant clusters, etc. Also note that the number of zero morphemes is drastically reduced: whereas DM would treat a noun-verb conversion pair as involving two zero affixes – one realizing little *n* and another realizing little *v*, within Don's model there will be one zero less: if the verb is primary/underived, there will be a zero affix realizing the nominalizer and vice versa^{10, 11}.

⁹ We refer the reader to an excellent argument against the proliferation of zero morphemes in DM offered in Stump (1998:40-41) (see Stump also for references to other arguments against the DM approach).

¹⁰ Among authors who adhere to the same derivational view of conversion as Don 1993, we can mention Beard 1998, cf. the following passage from Beard (1998:62): "Transposing a lexeme from one category to another without affixation is sometimes called *conversion*. The evidence weighs against a separate operation of conversion, however, for we find precisely the same semantic relations between conversional pairs as between derivational pairs. Thus for every conversion *to dry*, *to wet*, *to empty* we find at least an equal number of affixed derivatives with the same relation: *to shorten*, *to normalize*, *to domesticate*. Moreover, precisely those stems which affix are precluded from conversion (*to *short*, **normal*, **domestic*), and precisely those which convert are precluded from affixation: *to *endry*, **wetten*, **emptify*. The simpler account of such forms is that those without affixation are null marked variants of the same derivation which is otherwise marked by a variety of affixes". For review of literature on conversion and various other authors for whom conversion is directional, see Don 1993.

¹¹ However, Don 2003 gives an account of conversion in Dutch and German in accordance with DM which involves two zero affixes. For instance, it is not the case that a root like \sqrt{feest} is a primary noun which can be converted to a verb with the help of a zero affix. Rather \sqrt{feest} is first nominalized through insertion under little *n* and then the *n- \sqrt{feest}* structure is verbalized by being embedded under little *v*: both little *v* and little *n* are realized by zero affixes (ibid. p. 42). Still, Don's (2003) account in DM-terms is different from our interpretation of DM-treatment of conversion which maintains that for DM any root is insertable in any syntactic context (under any category-determining head) whereas for Don 2003 a root like \sqrt{feest} apparently cannot be inserted directly under *v*.

In the next two subsections we would like to consider two apparent pieces of evidence against bare roots in general.

2.2.3. Against bare roots in general: Onomatopoeic words in Sakha

Onomatopoeic words in Sakha may be argued to provide an argument for the categorality of roots. Sakha has a large number of onomatopoeic roots which can be mono- and disyllabic (trisyllabic roots are also possible but are not frequent). The vocabulary of onomatopoeic roots is freely expandable: the coinage of new roots is one of the distinctive features of epic poems and other kinds of folklore as well as everyday colloquial speech. The table below lists some of the onomatopoeic roots and their suggestive meanings which are to some extent subject to idiolectal variation.

(25) A sample of onomatopoeic roots in Sakha

Root	What it imitates
Baa	Low voice, bass
Njaa	Squeaking/meowing
Daa	Croaking
Xaa	Hoarseness
Buu	Wheeze
Küü	Buzzing
Ta	Tapping
Ca/sa/ha	Laughing
Ar	Growling
Bar	Roaring

What is important for the issue at hand is that these roots, whether already existent or newly coined, always come with a category – as nouns. Two points illustrate their categorality. First, they can appear in light verb constructions involving the verb *gyn* ‘do’ which can only take an NP as its direct object: *njaa gyn* ‘to squeak’, *bar gyn* ‘to roar’, etc. Second, the roots in question cannot function as verbs unless they are converted into verbs with the verbalizer –(L)AA as in *lihirdee* (lihir-LAA) ‘to tramp, walk heavily’¹². Thus, an onomatopoeic verb has the following template:

¹² –(L)AA can attach directly to di- or trisyllabic onomatopoeic roots. With monosyllabic roots, it is (almost) always the case that an onomatopoeic formant from a limited inventory (given below) is inserted between the root and the verbalizer as in *baaqynaa* (√baa-qY-LAA) ‘to speak in a low voice’, *kykkyraa* (√kyk-KYr-AA) ‘to creak’, *kylyrqaq* (√kyl-YrG-AA) ‘to rattle’.

(i) The inventory of onomatopoeic formants in Sakha

Onomatopoeic formant	Suggestive meaning
-qY-; -gY-; -kY-	Extended duration of sound
-djY-; -cY-	Energetic sounding
-KYr-	Energetic vibrating sound
-YrG-	Rhythmic sound

$\sqrt{\text{onomatopoeic}}$ – (onomatopoeic formant(s)) – (L)AA_{verbalizer}. As for onomatopoeic adjectives, they must be derived from the corresponding verbs.

The following route emerges for sound-imitating roots: root (=noun) → verb → adjective. That these roots which expand the language's lexical vocabulary always get categorized as nouns is in line with the fact that loanwords which are also an instance of lexicon enrichment are always borrowed as nouns independent of their category in the source language (for details and explanation see 3.5).

2.2.4. Against bare roots in general: Some notes on parsing

Aronoff and Anshen (1998:240) make the following remark with respect to the psychological phenomenon of blocking: “if a person has temporarily forgotten the word *fame*, then that person may in fact use the word **famousness*, which *fame* would otherwise block. This seeming failure of blocking is especially common in children, who coin new words quite freely, because their vocabulary is not as entrenched as that of adults” (p. 240). This seems to go against bare roots: if *fame* and *famousness* involve the same root, then forgetting *fame* would translate DM-wise as forgetting the root $\sqrt{\text{fame}}$ and forgetting to embed this root under n. However, if the same person would use *famousness* instead, this would translate into DM as embedding $\sqrt{\text{fame}}$ under a and then under n. Clearly, the result is contradictory.

Second, if *fire* as a noun and *fire* as a verb both involve the same bare root embedded in one case under n and in another under v, then the question of ambiguity resolution in parsing arises. For DM, the structural context of *fire* or any other ambiguous item should be an important factor in guiding the parser because if, for instance, *fire* is embedded under n, it would also be embedded under D.

Frazier and Rayner 1987 report the results of experimental studies on the resolution of categorial ambiguities in English such as the ones in (26). Three initial hypotheses formulated are 1) the first analysis strategy when the processor immediately adopts the first analysis of *warehouse* categorizing it as a noun; 2) the multiple analysis strategy when the processor constructs and maintains in active memory more than one analysis of the ambiguous string; 3) the delay strategy when the processor delays syntactic integration of the ambiguous item into the structure assigned to the preceding items¹³.

- (26) The warehouse fires ...
 a. numerous employees each year.
 b. harm some employees each year.

¹³ Cf.: “categorial ambiguities tend to be resolved locally and superficially, often by the category of the immediately following word. Thus, determining the basic category of an item (e.g., whether it is a noun or verb, though not necessarily whether it is a main verb or auxiliary or whether it is transitive or intransitive) can often be accomplished by examining the category of the following one or two words, without engaging in extensive syntactic computations or inferences. For example, in a sequence of three adjacent words, if the first word is the determiner *the* and the third is unambiguously an active verb (e.g., *hires*), the second word must be a noun (i.e., the head noun of the noun phrase introduced by the determiner)” (ibid. p. 507).

The results of three experiments confirm the predictions of the last hypothesis providing clear evidence for the delay strategy. F&R suggest that the processor, when using the delay strategy, considers all the available analyses of the ambiguous item in question but delays selecting and computing “all of the (global) structural consequences of just a single analysis” “in case helpful information is more likely to arrive immediately than further downstream” (ibid. p. 522)¹⁴.

It is crucial for the TS-based model of categorization that syntactic category assignment patterns together with thematic frame selection in invoking the same delay strategy. The findings in F&R contradict the expectations of bare root theories like DM which categorize lexical items by embedding them in the appropriate syntactic structures. For DM, there is no motivation to wait until later if the preceding structural context has already made clear that an ambiguous item like *warehouse* in (26) occurs under the determiner *the*: this occurrence should signal to the parser that an ambiguous item is a noun inserted in the D-n-√ structure. Thus, DM would be compatible either with the first analysis or multiple analysis strategies the predictions of which, however, are not confirmed. On the contrary, in a theory like TS, information about argument structure of an ambiguous lexical item is crucial for determining its syntactic category. Therefore the parser is expected to delay its decision about category assignment until it inspects the complement structure of the ambiguous element¹⁵.

2.3. Derivation: Lexicon versus syntax

In this section we will argue that word formation is possible in the lexicon because the lexicon is a computational module. It will be shown that lexicon-internal word-formation is principle-governed contrary to the commonly held view that lexicon is for the unruly. On the contrary, it will emerge that the nature of syntactic word-formation is more unruly than that of lexical one because only syntactic word formation can give rise to unfixed conceptual values such as when *djie-lee* (house-verbalizer) in Sakha is allowed to acquire a wide range of different meanings – a state of affairs impossible for the outputs of lexical derivation.

Before we proceed it is necessary to give some consideration to the traditional distinction between inflection and derivation, in particular, to the criteria which are often appealed to in accounting for such a distinction. The criteria discussed below are drawn from Stump 1998. Some of the criteria considered by Stump are also considered by Beard 1998 and Marantz 2001b but the latter two works make use of far less criteria than Stump (where Beard and Marantz also invoke the same or similar criterion as Stump will be mentioned in the course of discussion). Below when we talk about derivation, we will mean lexicon-internal word formation whereas meant by inflection is not only inflection as understood

¹⁴ The distance over which delay is tolerated (calculated in the number of words) was not investigated.

¹⁵ Note that the delay strategy may be well-motivated for a language like English where complements appear to the right of the complement-taking head. In ‘head-final’ languages like Sakha or Turkish there may be no need to invoke the delay strategy because the arguments of an ambiguous item would precede rather than follow the latter.

traditionally (e.g. case, tense, agreement) but also syntactic word formation (e.g. participles, syntactic nominalizations, etc.).

One of the most frequently cited criteria concerns that of meaning change (also considered by Beard): derivation results in meaning change, inflection cannot. Cf. Beard (1998:45): “derivation does change the meanings of words so as to allow the derivate to become a lexical entry in the lexicon”. For Marantz this criterion is connected to the notion of transparency: inflection (=syntactic derivation outside the root domain) is transparent, derivation (syntactic derivation inside the root domain) is not transparent. Transparency has to do with the fact that affixes are allowed to attach to a particular root before the attachment of a category-determining head *x*: only when the latter is added, can the *x*-root combination be submitted to the interpretational component and the affixes enter into semantic interaction with the meaning of the root negotiating all kinds of idiosyncratic meanings for the *x*-affix(es)-root combination. In a configuration like affix2-*x*-affix1-root, only affix 2 can trigger meaning change on the root: affix2 does not have this ability since it occurs above *x* and is subject to transparency.

A second criterion discussed and criticized by both Stump and Beard is category-change: derivation unlike inflection can change category. Beard points out that this criterion has no applicability in the absence of a clearly defined theory of categorization. Thus, in order to decide if derivations like *violin* : *violinist*, *cream* : *creamery*, *zip* : *unzip* change category or not, one first has to make decisions about the categorial nature of the bases involved (Beard 1998:46).

The criterion of syntactic determination as formulated in Stump (1998:15) comes third: “A lexeme’s syntactic context may require that it be realized by a particular word in its paradigm, but never requires that the lexeme itself belong to a particular class of derivatives”. In other words, only inflection determines syntax, not derivation. This criterion is a non-issue but its being a non-issue provides indirect support for the fact that it is only syntactic computation which is sensitive to particular categories whereas for lexicon-internal computation there is no need to make use of categorial information.

A fourth criterion which is rejected is that of productivity. Stump (1998:16) notes, on the one hand, the existence of “highly productive morphological phenomena which (by the other criteria) are derivational” and, on the other hand, the possibility of inflectional paradigms “which are *defective* in that some of their cells are left empty”. Aronoff and Anshen 1998 also argue against viewing productivity as an absolute notion. As we saw above in section 2.1.2.1, the data from word coining in Sakha confirm Aronoff and Anshen’s view: suffixes normally treated as unproductive turn out as productive or semi-productive when inventing new words¹⁶. Marantz is also against invoking productivity: one and the same suffix may be productive when it occurs outside the root domain as in *donator*, *driver*, etc. and unproductive or semi-productive when it occurs inside the root domain as in *donor*, *debtor*.

According to the fifth criterion discussed by both Stump and Beard, inflection is semantically more regular than derivation. Stump attributes this

¹⁶ Cf. Sproat (1985:499): “Productivity, in particular, does not serve as a criterion for dividing word-formation from the rest of the grammar”.

difference to lexical listing. However, given “the existence of highly productive classes of derived forms and irregular or defective paradigms of inflected forms”, he concludes that “listedness is neither a necessary nor a sufficient correlate of the inflection/derivation distinction” (pp. 17-18). For Marantz semantic regularity, just like meaning change, is connected to the transparency of syntactic derivation outside the root domain discussed above.

Finally, one more criterion which cannot hold water discussed by Stump: inflection closes words to further derivation, while derivation does not. This criterion is rejected by Stump because it is not supported by the data. A corollary of this criterion is that inflectional affixes must occur outside derivational ones (also endorsed by Beard who, just like Stump, comes to a negative conclusion because the “inflectional markers occur widely inside derivational markers” (Beard 1998:45).

Thus, Stump, after rejecting many criteria, argues neither in favor nor against the inflection/derivation distinction. However, he makes an important note that “the theoretical appropriateness of the inflection/derivation distinction” is only important in as far as it helps answer the fundamental question “Does a theory that incorporates this distinction furnish simpler (more learnable) grammars than one that doesn’t?” (Stump 1998:19).

We will assume that parameters are more easily formulable in a theory which makes use of such a distinction and therefore proceed to show that all derivation cannot take place in the syntax, viz. there is an essential difference between derivation and inflection which cannot follow under the assumption “all derivation in the syntax”. We will maintain that lexicon-internal derivation (word formation; new concept formation) and syntax-internal derivation (word formation; *new concept formation) are independent phenomena belonging to different modules (although their mutual independence does not preclude interaction). Instead of letting too many cooks spoil the broth, the criterion of meaning change will be taken as the defining one in drawing the distinction between derivation in the lexicon and the derivation in the syntax. The necessity of this criterion is determined not only by treating lexicon as the locus of computation distinct from syntactic one but also follows from the fundamental nature of the lexicon as the module of grammar where new concepts are formed.

2.3.1. Untenability of the tenet “All derivation in the syntax”

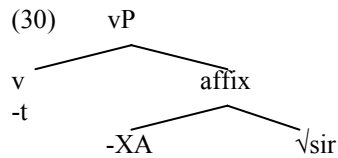
Since not all derivation is fully regular and productive, allowances must be made within any framework to explain morphological idiosyncrasies. As discussed above and in the introduction, DM assigns all such irregularities to the domain of the root which is the domain below the category-determining x 's. However, this kind of approach only sweeps the problem under the rug by creating an anomalous domain which is part of syntax and yet not subject to principles of syntactic composition. Let's consider this in more detail. The minimum structural requirement any phonologically well-formed word must satisfy is that it must embed an x -root configuration. Derivation can add more structure by attaching affixes either below x to the root itself or above x to the whole x -root complex. Since a number of affixes are promiscuous, the same affix can attach in the two different domains making possible a structure like [affix₂-affix₁- n -affix₂-affix₁- $\sqrt{\quad}$] where affixation above x

replicates in the exact order affixation below *x*. However, whereas the combination of affix₁ with [n-affix₂-affix₁-√] and of affix₂ with [affix₁-n-affix₂-affix₁-√] is fully compositional, this is not the case for the combination of affix₁ with the root and of affix₂ with [affix₁-√]. In fact, the meanings of these combinations are not determined each time an affix merges: meaning determination is delayed until the phase level because only a phase can be submitted to the interpretive module. Since it is the little *n* which defines the phasal edge in the [n-affix₂-affix₁-√] configuration, no intermediate structure like [affix₁-√] or [affix₂-affix₁-√] can be shipped off to LF before the merger of *n*. At the point when the phase is submitted, the Encyclopedia is searched for a suitable meaning of [n-affix₂-affix₁-√]. If no such meaning were available, the structure would be uninterpretable resulting in a crash.

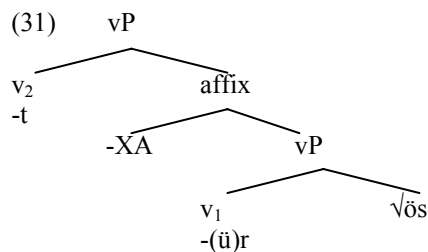
Thus, what happens after a particular root is selected and merged in the syntax and before this root is merged with a functional *n/v/a*-head is left completely ungoverned (apart from being partially conditioned by the (un)availability of idiosyncratic pieces of semantic knowledge stored in the Encyclopedia): neither the number nor the order nor the nature of affixes can be regulated. The kind of asymmetry created by DM with respect to the two places of affixation is reminiscent of the lexicalist assumption that lexicon is the storage place of all idiosyncrasies, except that in DM the storage of idiosyncrasies is divided between the syntactic domain below an *x*-head and the encyclopedic semantic module.

This argument is based on the conceptual unattractiveness of the asymmetry assigned to the syntactic component. To illustrate the practical side of the problem we will consider one particular group of verbs from Sakha which bear identical morphology. First, we will consider a TS-based account and then compare it with one based on DM. The verbs in question are, on the one hand, *sir* ‘dislike, disdain, reject’, *sirget* ‘disgust’, *sirgen* ‘become disgusted’ and, on the other, *öhür* ‘bear a grudge against so’, *öhürget* ‘offend’, *öhürgen* ‘become offended’. These verbs are interesting from the point of view of their morphology as well as their thematic structure. With respect to the latter, *sir* ‘dislike, disdain, reject’ and *öhür* ‘bear a grudge against so’ have the following theta feature specification: +c+m, -c-m. The transitive/causative counterparts *sirget* ‘disgust’ and *öhürget* ‘offend’ also have identical argument structures: they are associated with a +c causer, a -c+m experiencer and a -c-m theme. In the corresponding intransitive experiencers the causer has been reduced giving the following argument structure: -c+m, -c-m. Whereas the causative/experiencer alternation is derivationally motivated, it is much less clear what to do with the first verbs (*sir* ‘dislike, disdain, reject’ and *öhür* ‘bear a grudge against so’) which, from a purely morphological point of view, can be considered as the bases in deriving the causative/experiencer pair. *Sirget* ‘disgust’ decomposes into *sir-ge-t* where -ge is the allomorph of the nominalizing suffix -XA and -t is the verbalizer. Deriving *sirget* ‘disgust’ from *sir* ‘disdain, reject’ would involve splitting the agentive role into [+c] and [+m]: the former will be interpreted as causer, the latter as [-c+m]-experiencer.

- (27) *Sir* ‘disdain, reject’ → *sir-ge-t* ‘disgust’ → *sir-ge-n* ‘get disgusted’
 [+c+m; -c-m] [+c; -c+m; -c-m] [-c+m; -c-m]



With respect to the other verb *öhürget* ‘offend’, given DM-guidelines, there is no way to derive it from the root $\sqrt{\text{ös}}$. In order to negotiate for the root $\sqrt{\text{ös}}$ the special meaning ‘offend’ (different from ‘bear a spite’ – the meaning induced by the verbalizer $-r$ in (29)), $\sqrt{\text{ös}}$ must be affixed with $-XA$ and then verbalized with $-t$. However, more deeply embedded inside *öh-ür-ge-t* is another verbalizer v_1 which already closed off the root domain – the domain of special meanings. Therefore whatever is suffixed above $/-r/$ can never induce a special meaning.



Diachronic evidence also weighs against assigning all derivation to the syntactic component. In the framework which is argued for in this dissertation not all derivation is syntactic and lexical word formation is possible. New words derived lexicon-internally become new lexical entries listed alongside their bases. For instance, the word *künüs* ‘afternoon’ is derived from *kün* ‘day; sun’. Since derivation involved here is lexical, both the base *kün* ‘day; sun’ and the derivative *künüs* ‘afternoon’ are listed in the lexicon. Derived lexical entries can be ‘frozen’ where ‘freezing’ entails severing the original derivational link: for example, if *künüs* ‘afternoon’ is frozen, one would be unable to restore its derivational history and the word in question would be analyzed as underived, having no relation whatsoever to *kün* ‘day; sun’.

Evidence for freezing comes from diachrony. Imagine Word.Z which has the internal composition Word.Y+suffix (in other words, Word.Z is derived in the lexicon from Word.Y and the process is marked overtly by the suffix). Both Word.Z and Word.Y will be listed in the lexicon as full-fledged members. Now suppose that the base Word.Y over the course of centuries has undergone some phonological changes and, from a synchronic point of view, now looks like Word.X. Its derivative Word.Z, however, did not change its internal composition from Word.Y+suffix to Word.X+suffix. Facts like these are certainly of great help to linguists concerned with phonological laws or etymology but they also prove the possibility of freezing, whence, indirectly, the possibility of lexicon-internal word formation which offers a readily available account of freezing in terms of listing both the base and the

derivative and severing the derivational link between the two. On the contrary, if all derivation were syntactic as in DM, we would expect the process of derivation to use as bases the actual roots of the language: Word.Z would be derived from Word.Y in the syntax with an appropriate suffix. Only Word.Y will be listed in the lexicon: if it underwent some historical change becoming Word.X, then now Word.X will be chosen from the lexicon in the derivation of Word.Z. Hence, we would expect parallel development for both Word.Y and Word.Z: if Word.Y becomes Word.X, then Word.Z (originally Word.Y+suffix) must also change its internal structure to Word.X+suffix.

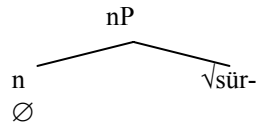
Fortunately for the present approach and unfortunately for DM, facts like these are abundant in Sakha. Let's consider some particular cases of old (original) roots conserved. *Durda* 'stronghold, refuge, shelter' is a common common noun in modern language and consists of $\sqrt{\text{dur-}}$ and $-\text{TA}$. Historically, $-\text{TA}$ is related to the locative case marker $-\text{TA}$. Since locative case is now extinct, $-\text{TA}$ has been preserved as a sporadic nominalizer attaching to nouns (see appendix 3: section 1) and inside a complex productive nominalizer $-\text{l-TA}$ or its variant $-\text{m-TA}$ ¹⁸. The bound root $\sqrt{\text{dur-}}$ is the archaic form of the modern verb *tur* 'to stand'. Under current assumptions, if a verb is nominalized and its argument structure reduced, this kind of derivation is lexical and both the base verb and the derived noun are listed as lexical entries. Long time ago when modern *tur* 'to stand' was still **dur*, a noun was derived with the nominalizer $-\text{TA}$ using **dur* as the base. This resulted in the creation of a new lexicon entry *durda* 'stronghold, refuge, shelter' in addition to the already existent **dur* 'to stand'. The new entry was frozen and therefore did not replicate the change which was undergone by **dur* \rightarrow *tur*. This type of freezing derivation exemplified by *tur* versus *durda* is best accounted for within a framework which allows lexicon-internal derivation: syntax-only derivation cannot accommodate freezing.

Another interesting case is presented by the word *sür* 'soul, internal spirit' and its three derivatives: two nouns *sürex* 'heart', *sürge* 'mood' and one adjective *sürün* 'basic, fundamental, essential'. Before the revolution all four were in regular use. After the revolution, the base was forced to go out of use because it reflected religious beliefs. The derivatives, however, stayed. In the last decade the word *sür* made a comeback with its original meaning. Assuming a syntactic derivation, two possibilities are conceivable none of which is plausible enough. Since all four involve special meanings, there must be the root $\sqrt{\text{sür-}}$ which is embedded in four different contexts either as in (32) or as in (33). In (32) there are three types of n's and one type of little a (all realized by different suffixes) which induce different meanings on the root. According to this analysis, the disappearance of (32a) would mean either the disappearance of the particular type of nominalizer used in (32a) realized by \emptyset (less likely) or the disappearance from the encyclopedia of the first meaning (more likely). However, adopting the analysis in (32) would eventually result in the proliferation of functional heads, be them n's, v's or a's, each one of

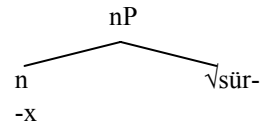
¹⁸ $-\text{TA}$ also occurs inside a complex productive adjectivizer $-\text{TAAqY}$ which derives locative/temporal adjectives (see appendix 3: sections 7 and 9).

which would induce a different meaning on the root depending on which affix realizes it¹⁹.

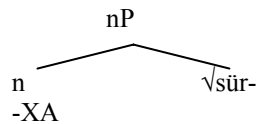
(32) a. *sür* ‘soul, internal spirit’



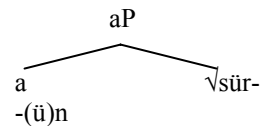
b. *sürex* ‘heart’



c. *sürge* ‘mood’



d. *sürün* ‘basic, essential’



(33) avoids this pitfall by postulating only one type of nominalizer uniformly realized by the zero affix but allows idiosyncratic derivation inside the root domain negotiate with the encyclopaedic knowledge special meanings for $\sqrt{\text{sür-}}$. For instance, if (33a) were possible, the $[\text{n-}\sqrt{\text{sür}}]$ substructure would be sent LF, the meaning of $\sqrt{\text{sür-}}$ would be looked up and (33a) would acquire the meaning ‘soul, internal spirit’. The same would happen in the case of (33b) and (33c) but the substructures shipped off for interpretation will be, respectively, $[\text{n-affix}(x)\text{-}\sqrt{\text{sür}}]$ and $[\text{n-affix}(XA)\text{-}\sqrt{\text{sür}}]$ with the same nominalizer as in (33a). This time one will have to look if the encyclopedia makes available special meanings for $\sqrt{\text{sür-}}$ in the context of the pertinent affixes. Since there is a special meaning of ‘heart’ for $\sqrt{\text{sür-}}$ in the context of $-x$ and a special meaning of ‘mood’ for $\sqrt{\text{sür-}}$ in the context of $-XA$, the derivations in (33b) and (33c) will be interpretable. However, it is quite unclear how one would negotiate these special meanings if the meaning in (33a) disappeared from the encyclopedia: the meaning in (33a) is the basic one and it would make little sense to try to determine the meaning of $\sqrt{\text{sür-}}$ in the context of either $-x$ or $-XA$ if one cannot even determine the meaning of $\sqrt{\text{sür-}}$ by itself. On the contrary, under lexical derivation all four items will be listed in the lexicon and freezing the derivatives will not eliminate them from the lexicon when the base is eliminated. This analysis offers a much simpler solution than the one based on consulting the encyclopedia.

¹⁹ A similar caution is expressed by Kihm 2000.

- (33) a. *sür* ‘soul, internal spirit’ b. *sürex* ‘heart’
- nP

n
∅

√sür-

nP

n
∅

affix

-x √sür-
- c. *sürge* ‘mood’ d. *sürün* ‘basic, essential’
- nP

n

affix

-XA √sür-

aP

a

-(ü)n √sür-

Diachrony can also support the analysis given above for the verb *öhür* ‘bear a spite’ which is a denominal verb and as such must be derived in the syntax: *öhür*, however, has been reanalyzed as a new concept and acquired its own listing in the lexicon alongside the base noun *ös* ‘spite’. If such reanalysis were indeed possible, we would expect freezing to be possible in denominal verbs. The expectation is justified and can be illustrated with *suuj* ‘wash’ and *tajaa* ‘put together pieces of material’, *tajyn* ‘dress oneself’. The verbs have as their bases bound roots √suu- and √taj-. Historically, these bound roots correspond in modern language to the nouns *uu* ‘water’ and *son* ‘coat’. The history can be relived as follows: the verbs *suuj* ‘wash’, *tajaa* ‘put together pieces of material’ and *tajyn* ‘dress oneself’ were derived from the then nouns **suu* ‘water’ and **taj* ‘coat’ in the syntactic component; then the verbs were reanalyzed as new concepts, given a separate entry in the lexicon and the link between the base noun and the derivative verb has been frozen. As a result, when the nouns were subjected to phonological laws eventually becoming *uu* ‘water’ (←**suu*) and *son* ‘coat’ (←**taj*), the other nouns frozen inside the verbs were exempted from the process.

Discussion in this section concentrated on arguments against the DM-tenet that all derivation is syntactic. We have shown that moving morphological derivational idiosyncrasies to the syntactic component creates within the latter a conceptually unattractive asymmetry. Even if the asymmetry is allowed in, DM still cannot account for meaning negotiations which take place above clear instances of category-determining functional heads (recall the discussion of *öhürget* ‘offend’). Finally, DM cannot accommodate freezing.

2.3.2. Predictions for lexical and syntactic derivation

Given what has been said in the preceding section, it is not feasible to keep all derivation in the syntax. Therefore we will assume that derivation is possible both in the lexicon and in the syntax. In 2.3 we considered a number of criteria which have

been advanced in the literature to formulate the distinction between syntactic and lexical derivation and arrived at the conclusion that the only relevant criterion is that of meaning change. The sole necessity of this criterion is motivated by the fundamental nature of the lexicon as the module of grammar where new concepts are formed. All properties of lexical and syntactic derivation stated in the opening lines to this chapter in (3) and (4) and repeated below follow directly from this criterion. First, with respect to lexical WF, property 1 goes without saying. The second property captures the intuition that one of the ways to arrive at a new meaning is by modifying the original argument structure as in *John broke the window* versus *The window broke*. Since a meaning can change without an accompanying change in the number and θ -feature specification of arguments, property 3 emerges. Fourth property is connected to the basic thesis defended in this work, namely, that only verbs and adjectives can take arguments. Property 5 follows for nominal concepts without saying and it will be shown below that lexicon-internally the meanings of verbs and adjectives can change while their θ -grids remain unaltered. As for properties of syntactic WF in (35), (1) goes without saying. (2) follows from the fact that eliminating or modifying one of the original theta roles will necessarily change the meaning of a concept but not so for adding an extra argument. Third property follows from the fact that categorization is a property of syntax not connected with conceptual content and therefore changing a category from e.g. verb to adjective or noun to adjective does not violate the criterion of new concept formation.

(34) Properties of lexical WF:

1. LWF results in meaning change
2. LWF can manipulate argument structure by deleting, adding or modifying a θ -role
3. Thus, there are two types of lexical WF: argument-structure-manipulating and argument-structure-preserving
4. AS-manipulating LWF can only apply if a concept is associated with a θ -grid, hence AS-manipulating LWF can take as its input verbs and adjectives but not nouns
5. AS-preserving lexicon-internal manipulation can take as its input any concept irrespective of the presence or absence of a θ -grid

(35) Properties of syntactic WF:

1. SWF does not induce meaning change
2. SWF can only add an argument, it cannot modify or eliminate the original θ -feature clusters (=arguments)
3. SWF can also apply while preserving the original argument structure which results e.g. in syntactic nominalizations

In accordance with (34) and (35) we can formulate specific predictions of the present approach for lexical and syntactic derivation.

(36) Predictions for lexical WF and syntactic WF

Input	Output	Domain of application	
Nouns	Nouns	Lexicon	Syntax
Nouns	Verbs		Syntax
Nouns	Adjectives		Syntax
Verbs	Nouns	Lexicon	Syntax
Verbs	Verbs	Lexicon	Syntax
Verbs	Adjectives	Lexicon	Syntax
Adjectives	Nouns	Lexicon	Syntax
Adjectives	Verbs	Lexicon	Syntax
Adjectives	Adjectives	Lexicon	Syntax

These particular types of derivation particularized to domains and categories will be considered in sections 2.4 through 2.5. Before we proceed to these sections a couple of general remarks are needed in order to clarify our stand on the issue of correspondence between derivation and morphological marking.

2.3.3. Some notes on morphological marking of derivation

It is usually assumed that syntactic derivation is compositional and transparent, lexical derivation is not (this is the criterion of semantic regularity discussed above). There are various ways to approach this problem. One can be rejected immediately. This difference cannot be taken to reflect the existence of a bifurcation in the inventory of affixes because of suffixal promiscuity (discussed in section 2.1.2.1), when one and the same affix is allowed to participate in both compositional (regular, transparent) and non-compositional (irregular, non-transparent) derivation. In addition, we showed in 2.1.2 that in Sakha the basic inventory of suffixes counts a limited number of simplex suffixes which can appear in semantically non-transparent as well as transparent contexts and which, furthermore, can combine with each other resulting in combinations that, again, have transparent as well as non-transparent uses. Thus, suffixal promiscuity supports Aronoff's (1994:126) observation that "derivation and inflection are not kinds of morphology but rather uses of morphology: inflection is the morphological realization of syntax, while derivation is the morphological realization of lexeme formation" (cited by Stump 1998:19).

In 2.3.1 we also argued against a second possibility, namely, a DM-postulate that semantic regularity (compositionality) is connected to the transparency of syntactic derivation outside the root domain while its lack is blamed on the non-transparency of derivation inside the root domain.

A third option that cannot be adopted here is Reuland's (1988) suggestion that the difference between inflectional and derivational suffixes is due to the fact that only the former bear a specific lexical meaning (e.g. –ed and –ing) enabling a compositional semantic interpretation whereas the latter (e.g. the derivational affix –al in *refusal*) are not associated with a meaning for compositional semantics to operate on. This suggestion cannot go through given the existence of complex suffixes in Sakha which bear a specific meaning but which are composed of simplex

suffixes for which it is very hard to specify any meaning. For example, /-msYj/ and /-msAx/ are composed of, respectively, /m+s+j/ and /m+s+x/. /-msYj/, illustrated in (37), derives verbs from nouns and adjectives (see appendix 3: sections 4 and 6) with the meaning ‘pretend to be/act like N/A’. /-msAx/, illustrated in (38), derives adjectives from nouns (appendix 3: section 7) with the meaning ‘N-loving’. The sole morphological difference between the two complex affixes is the j/x alternation from which the difference in meanings cannot follow. Nor can the respective meanings themselves be derived from the simplex suffixes –m, -s, -k and –x.

(37) N/A-msYj → V: ‘pretend to be/act like N’

N/A	Gloss	N/A→msYj	Gloss
Njirej	A calf	Njiremsij	Act stupid, naïve
Tojon	Boss	Tojomsuj	Act bossy
Uus	Master, craftsman	Uuhumsuj	Pretend to be a master, show off one’s craftsmanship
Öjdöox	Clever	Öjdööqümsüj	Pretend to be clever
Sytyy	Cunning	Sytyymsyj	Pretend to be cunning

(38) N-msAx → A: ‘N-loving’

N	Gloss	N-msAx → A	Gloss
Et	Meat	Etimsex	Meat-loving
Balyk	Fish	Balygysax	Fish-loving
Kuoska	Cat	Kuoskamsax	Cat-loving
Beje	Self	Bejemsex	Selfish

Our approach to morphology can be summarized as follows. Along with Borer, we will assume the existence of two lexicons – conceptual (which will often be referred to simply as the lexicon) and functional²⁰. Items in the conceptual lexicon are not phonologically abstract concepts but rather sound-meaning pairings and therefore enter syntax phonologically specified. Items of the functional lexicon come in two varieties: on the one hand, the f-lexicon contains grammatical features with no phonological content (e.g. [plural], [aorist], [negative]) and, on the other, phonologically specified functional morphemes (e.g. /-LAAX/ in Sakha). Therefore syntax operates on syntactic trees whose terminal nodes may be either phonologically abstract features or phonologically specified morphemes. The former will be supplied with phonological matrices at the point of Spell-Out when the Vocabulary is accessed. As conceived in DM, the Vocabulary is a list of phonological forms for (bundles of) grammatical features (e.g. /-LAR/ for [plural], /-AR/ for [aorist], /-MA/ for [negative] in Sakha). We are thus adopting, following DM and Borer, a postsyntactic morphological component which is responsible, apart from vocabulary insertion, also for allomorphy.

²⁰ In DM, on the contrary, there is one lexicon which lists both roots and grammatical features.

With respect to the issue of correspondences between derivation and morphological marking (realized as suffixation in Sakha), we will take the following stand (which is particularized to Sakha and does not present a global picture). Since the lexicon lists concepts, we argued above that bound roots must be listed together with their suffixes (a bound root by itself has no conceptual content). Lexicon-internal derivation changes meanings and can manipulate argument structure. These operations are expected to be marked overtly. Thus, affixation can take place in the lexicon. Consider, for instance, nouns derived from nouns as in *törüt* ‘origin’ → *törüccü* ‘genealogy’, *xappax* ‘cover, lid’ → *xappaxy* ‘store-room’.²¹ They will be derived and marked with –CI already in the lexicon so that they will enter syntax as *törüccü* and *xappaxy* and not as, say, $\{törüt/xappax\} + \text{suffix}(to\ be\ chosen\ in\ the\ postsyntactic\ morphological\ component)$.

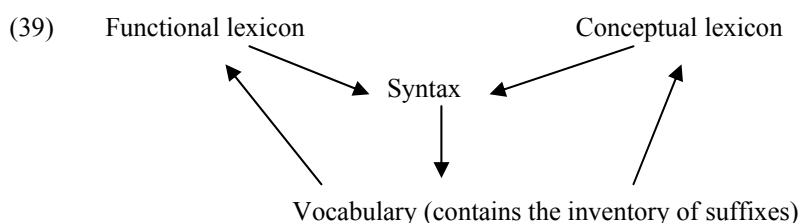
However, when inputs from the lexicon, whether basic or derived, are submitted to syntax, their internal structure is still visible to syntactic computation making possible such phenomena as suffix modification under degree intensification described in section 2.1.1.1 (examples (10) and (11)) or suffix instability. Recall the example with *ynax* ‘cow’ which is built on the bound root \sqrt{y} - plus the complex suffix –na-x. When *ynax* is verbalized with /-j/ (N→V conversion must take place in the syntax according to the present assumptions; see below), the x-part of the nominalizer –na-x is dropped resulting in *ynaj* ‘resemble a cow’. If a suffix attaching to a bound root were completely non-transparent to syntax, then it would be impossible to derive a verb like *ynaj*. Derivations like *ynax* → *ynaj* show that word boundaries are not preserved through derivations²².

Marking conventions also apply to derivation within the syntactic component proper which does not change meaning but which can change category and add arguments: argument addition and category change should be appropriately marked. However, the difference of syntactic derivation from lexical is that it involves syntactic heads whose terminal nodes may be either phonologically abstract features or phonologically specified functional morphemes like –LAAX which derives adjectives from nouns. In the latter case we have an instance of syntactic affixation, i.e. we are dealing with affixes which operate in the syntax by combining with other syntactic constituents. In the former case we have syntactic nodes whose abstract feature content will be replaced by affixes chosen from the vocabulary at the point of spellout. Note that in both cases idiosyncratic affixation is a priori excluded: in the case of functional morphemes the syntactic use of each particular morpheme must be negotiated at the interface between syntax and the f-lexicon whereas in the case of abstract features, which affix will be inserted to replace a particular feature is regulated by the competition among vocabulary items. In the lexicon, on the contrary, bound root suffixation and lexicon-internal derivation are

²¹ Note that the latter pair presents another counterexample for DM: the base *xappax* ‘cover, lid’ is derived from the root \sqrt{xap} which also gives rise (with the help of a zero morpheme) to the verb *xap* ‘catch; encompass’. Hence, in *xappax* ‘cover, lid’ a special meaning has already been negotiated for the root \sqrt{xap} and no more negotiations are possible: *xappax* should be closed for further idiosyncratic derivation and it should be impossible to arrive at the special meaning ‘store-room’ for *xappaxy*.

²² Lexicon-internal derivation also does not preserve word boundaries: when causative *xaptat* (built on a bound root) ‘make flat’ yields, under [+c]-reduction, unaccusative *xaptaj* ‘become flat’, the causative suffix –t drops and is replaced by –j.

marked in a rather unpredictable way. Nevertheless, the actual suffixes used are drawn from the basic inventory of suffixes. Therefore we can assume the following tentative organization of grammatical modules. The arrow from the vocabulary to the f-lexicon indicates the fact that those functional morphs which are listed there are (composed of) the same suffixes as the ones made available by the vocabulary. The arrow from the vocabulary to the conceptual lexicon is also meant to capture the fact that all suffixation inside the lexicon involves the basic inventory.



As a final remark on transparency of morphological structure, we would like to maintain that the scope of the affix is determined by the position it occupies and no affix movement inside words takes place. A theory, in which there is no correspondence between the degree of embedding of a particular suffix and the material over which it takes scope, may have to resort to affix hopping (cf. Chomsky 1981) or affix raising (cf. Pesetsky 1985, Lebeaux 1986). However, affix movement inside words has many conceptual and empirical disadvantages as shown extensively in Reuland 1988.

Before presenting in the following two sections the specifics of the above assumptions about morphological marking of derivations, we would like to compare these assumptions to the ones advanced in Baker 2003. Baker assumes that morphological derivation and syntactic derivation can take independent paths because they belong to two different modules independent from each other – morphology and syntax. For instance, his analysis of verbs has both syntactic and morphological aspects. On the syntactic side, all verbs have a uniform derivation: given the underlying $[_{\text{PredP}} \text{Pred} [_{\text{AP}} \text{A}]]$ structure (where A is an abstract adjective), the A head moves to Pred: the resulting complex, conflated head Pred-A is interpreted as a syntactic V node and whatever is inserted into this X^0 node in the syntax is interpreted as a verb. There are virtually no restrictions on the complexity of a morphological unit that can be inserted into V. It can be a morphologically simple verb root (*fall*, *die*), a verbal stem derived from a category-less bound root plus a verbalizer (*magnify*, *colonize*) or a root compound (*panfry*, *handwash*) (ibid. 276-7). It can also be a deadjectival or a denominal verb, both derived in the morphological component by affixing a verbalizer to an adjectival or a nominal root. In the case of deadjectival verbs their internal morphological structure (adjective plus affix) directly corresponds to their syntactic derivation (adjective conflates with Pred). In the case of denominal verbs such as *fossilize*, *crystallize*, *symbolize*, *classify*, *originate* and *knight*, it goes counter since syntactically these verbs must be analyzed, like any other verbs, as involving an adjective conflated to a verb but morphologically speaking, they are derived from a noun root. Such counter-derivation is allowed because morphological structure is not syntactically

transparent which follows from the independence of the two modules. What is derived within morphology can be as complex internally as one can get but syntactically it counts as an X^0 -level category.

Current framework converges with Baker's in that morphological derivation is possible outside syntax (lexicon-internally in our view) but diverges from his theory in rejecting a drastic split between morphological and syntactic derivation (lack of any correspondences). We demonstrated above²³ that syntactic processes are sensitive to the internal structure of complex words which enter syntax from the lexicon and therefore the effects of lexicon-internal derivation must be visible to syntactic computation.

2.4. Deriving nouns in the lexicon and syntax

In this and the following two sections we investigate in detail the predictions formulated in (36) and consider the specific properties of deriving nouns (2.4) and adjectives (2.5) in the lexicon and syntax. The practical materials to these sections are given in appendix 3. We begin with nouns which can be derived from nouns, adjectives and verbs. All three types of derivation can take place either in the lexicon or syntax.

2.4.1. Deriving nouns from nouns

The defining property of nouns is that they take no arguments (they can take possessors but these are introduced syntactically; see chapter 3). Thus, argument-structure manipulating derivation cannot by definition apply to nouns in the lexicon, and outputs of lexical $N \rightarrow N$ derivation also must be argument-less. However, $N \rightarrow N$ lexical word formation must bring about a meaning change. This is indeed confirmed by the data in appendix 3 (section 1): a sample is given in (40).

(40) $N \rightarrow N$ LWF

N	Gloss	N-suffix \rightarrow N	Gloss
Kölö	Working cattle	Kölöhün	Sweat
Alyp	Magic	Albyn	Deceit
Uu	Water	Utax	Beverage
Djyl	Year	Djylqa	Destiny
Xaa	Box, receptacle	Xax	Cover, shell, skin
Xax	Cover, shell, skin	Xaxxa	Shelter
Xaa	Box, receptacle	Xabax	Bladder
Töbö	Head	Töböt	Urchin, rascal

²³ Recall examples (10) and (11) in section 2.1.1.1 (suffix modification under degree intensification) and the syntactic instability of suffixes attached lexicon-internally. For instance, *utax* 'drink, beverage' is derived in the lexicon from *uu* 'water' with the complex suffix $-ta-x$. In the syntax *u-ta-x* can be subjected to verbalization with the suffix $-t$ resulting in *u-ta-t* and not in **u-ta-q-yt*: syntactic verbalization detects the complexity of the nominalizer $-ta-x$ consisting of two nominalizers $-ta$ and $-x$ and replaces the latter with the verbalizer $-t$.

In some very few cases, the derivation appears not to change the original meaning and the sole difference between the base and the derivative is the presence of an extra suffix in the latter: e.g. *kurun* ‘dried up tree/forest’ versus *kurunax* ‘dried up tree/forest’. In such cases it is plausible to assume, given independently motivated suffixal instability, that the suffix can drop optionally: in other words, the fully suffixed derivative (*kurunax*) is the base and the base (*kurun*) is the reduced derivative.

Since N→N derivation does not affect argument structure, nothing prevents it from occurring in the syntax provided that the original meaning of the base is preserved. Indeed, such case is presented by the examples in (2) of appendix 3, section 1. A couple of examples are repeated below. The pertinent suffixes are complex: -CI is used to derive agentive nominals whereas -hY-t shares the -hY part with -hY-n, a suffix which derives syntactic action/process nominalizations from verbs. The nouns derived with these suffixes can be considered syntactic in the sense that the original meaning of the base noun is preserved and the new noun can be roughly paraphrased as ‘person whose occupation/hobby pertains to N’. Therefore -CI-t (-hY-t) should be viewed as a syntactic affix which comes from the functional lexicon with a meaning and when it combines with a noun during syntactic computation, the meaning of the derived syntactic unit is calculated from the meanings of the suffix and the noun. This treatment is similar to the ones adopted for syntactic affixation in both Baker 2003 and DM²⁴.

- (41) Nouns derived from nouns with the suffix -CI-t (or its variant -hY-t)

N	Gloss	N-CI-t → N	Gloss
Ot	Grass	Otcut	Person who mows grass
As	Food	Ascyt	Cook
Balyk	Fish	Balyksyt	Fisherman
Olonxo	Epic poem	Olonxohut	Epic poet

2.4.2. Deriving nouns from adjectives

The defining characteristic of adjectives is that they are one-place predicates. Therefore A→N derivation should involve argument reduction which is only possible in the lexicon. There are indeed cases of A→N conversion with accompanying meaning change which are marked overtly with derivational morphology, e.g. *cuor* ‘sharp, keen (of hearing)’ → *cuoraan* ‘bell’, *saja* ‘new’ → *sajas* ‘daughter-in-law’, *sürin* ‘basic, fundamental’ → *sürdjüges* ‘a pole serving as a basis for a fence’ (see section 3 of appendix 3). However, such cases are extremely rare for a language which otherwise makes great use of derivational morphology. We will argue in chapter 4 (section 4.7) that the reason for this is that grammar gives an opportunity in the syntax to create nouns from adjectives by binding their open

²⁴ Syntactic affixation deriving nouns from nouns is also encountered in the productive formation of diminutives with the suffixes -CIk, -kA and -CAAn which will not be considered.

argument position with a determiner (in line with Higginbotham 1985). Syntactic θ -binding creating nouns from adjectives is very productive in the sense that any adjective can be nominalized. Compared to lexical derivation, syntactic nominalization offers one advantage: whereas lexically derived nouns have highly idiosyncratic meanings, nominalized adjectives are very flexible meaning-wise and allow a range of contextually determined interpretations. We will assume that this is one of the properties of syntactic word formation, namely, that it creates an opportunity for contextual interpretation of the meaning of the derived expression. On the contrary, the function of lexicon-internal word formation is to bring about a meaning change creating new concepts: it is therefore inconceivable that LWF could give rise to unfixed conceptual values.

2.4.3. Deriving nouns from verbs

Verbs can be nominalized both in the lexicon and syntax. As predicted in the present framework, lexical nominalization reduces verbal argument structure and effects a meaning change whereas syntactic nominalization preserves the original arguments. Some examples of lexical nominalization under which the meaning of the resulting noun cannot be predicted are given in (42).

(42) Lexical V→N derivation

V	Gloss	V-suffix → N	Gloss
Tüs	Fall	Tühük	Grammatical case
Üün	Grow	Üünüges	Shoot, sprout; puppy
Aj	Create	Ajylqa	Nature
Uop	Take a mouthful	Obot	Appetite
Xamnaa	Move	Xamnas	Salary
Tüm	Sum up, conclude; unite	Tümen	Parliament
Tep	Kick	Tebilik	Support; bicycle

The cases presented in (43) differ from the ones presented above in that one can say that, even though arguments are reduced, no apparent change of conceptual content has occurred contradicting the defining criterion of lexical derivation – meaning change. However, we would like to maintain the following. In both (42) and (43) the meaning of the resulting noun has been fixed once and for all. This is quite different from what happens in denominal verb formation where the meaning of the resulting verb must be determined depending on the context. Thus, *xaardaa* ‘snow-LAA’ can mean ‘to snow’, ‘remove snow’, ‘provide with snow’, ‘consume snow’, ‘gather snow’. Similarly, *ülelee* can mean ‘to work’, ‘to provide with work’, ‘to go to work’. Denominal verb formation, as argued in chapter 7, is a syntactic process. It is not the job of the syntax to provide new lexical names and therefore it makes sense to expect those words which are derived syntactically not to have a fixed meaning but rather a meaning which will be adjusted later, e.g. in the semantic component. On the contrary, lexicon is the engine for supplying the lexical stock of the language and therefore new words derived lexicon-internally are expected to have their

meaning ultimately/terminally determined. Therefore meaning change is probably better understood as final meaning determination/fixing – a procedure which can take place inside the lexicon but not inside the syntax.

(43) Lexical V→N derivation

V	Gloss	V-suffix → N	Gloss
Tüm	Sum up, conclude; unite	Tümük	Conclusion
Tuluĵ	Tolerate, endure	Tuluk	Endurance
Solbuĵ	Replace	Solbuk	Replacement
Taptaa	Love	Taptal	Love
Xajqaa	Praise	Xajqal	Praise
Iteqeĵ	Believe	Iteqel	Belief
Boldjoo	Set a timeframe, deadline	Boldjox	Timeframe, deadline
Baaccaj	Complicate	Baaccax	Complication

Moving to V→N derivation which takes place in the syntax, we find two cases of syntactic affixation in section 2 of appendix 3: (23) and (24). In (23) we have two suffixes, -YY and -hYn, deriving action nominalizations from verbs (regularly and productively). The suffixes have the same syntactic effect but differ in their morphological properties: -YY attaches to verbs ending in a consonant, -hYn attaches to verbs ending in a vowel. (44b) shows the nominalization of a simple transitive sentence in (44a). We will assume that in such nominalizations the external argument is saturated and this saturation is marked by the morphemes – YY/-hYn. This assumption is compatible with the interpretation of (44b) as implying that the action of reading the book is performed by someone. In (44c) an optional possessor is introduced which corresponds to the external argument of the corresponding transitive sentence. However, as argued in chapter 3, possessors are not external arguments and they can be introduced syntactically. This is what we are witnessing in (44c). (44d) shows that the case assigning properties of the underlying verb must be preserved as well such that the accusative object of (44a) cannot be expressed as a nominative possessor in (44d).

- (44) a. Sardaana kinige-ni aaq-ar.
Sardaana book-acc read-aor
'Sardaana reads the book.'
- b. kinige-ni aaq-yy
book-acc read-nom
'reading the book'
- c. Sardaana kinige-ni aaq-yy-ta
Sardaana book-acc read-nom-3
'Sardaana's reading the book'
- d. *kinige aaq-yy-ta / book read-nom-3
'the book's reading; the reading of the book'

Nominalization can affect all kinds of verbs: (in)transitive, (un)derived, marked with various voices such as reflexive, reciprocal, causative, passive²⁵. (45) and (46) show nominalizations of a causative and passive verbs. In the latter case it is the derived subject of passive that is saturated testifying to the syntactic nature of nominalization. That the nominative possessor which is optionally possible in (46b) does not correspond to the external argument of (46a) is shown by a difference in interpretation. While in (46a) it is Keskil who is chosen, in (46b) this is not the only reading available for Keskil who can also be the person contemplating on someone being chosen, can be the author of a theory on political elections, etc.

- (45) a. Sardaana Lena-qa Erel-i möx-tör-dö.
Sardaana Lena-dat Erel-acc scold-caus-past.3
‘Sardaana made Lena scold Erel.’
- b. Lena-qa Erel-i möx-tör-üü / Sargy Lena-qa Erel-i möx-tör-üü-te
L.-dat E.-acc scold-caus-nom / S. L.-dat E.-acc scold-caus-nom-3
‘making Lena scold Erel / Sardaana’s making Lena scold Erel’
- (46) a. Keskil tal-ylyn-na.
Keskil choose-pass-past.3
‘Keskil was chosen.’
- b. tal-yll-yy / Keskil tal-yll-yy-ta
choose-pass-nom / Keskil choose-pass-nom-3

The suffixes –YY/-hYn are also encountered on bound roots, e.g. *ar-yy* ‘butter’, *malaahyn* ‘feast, banquet’. This shows that the suffixes can attach across domains: both in the lexicon and in the syntax. The difference in the domain of application is reflected in whether the nouns derived with –YY/-hYn allow denominal verb formation: those derived from bound roots (i.e. in the lexicon) do, those derived from verbs with argument structure preservation (i.e. in the syntax) do not, cf. *aryy-laa* ‘to butter’, *malaahyn-naa* ‘organize a banquet’ versus **yrytyy-laa*, **bosxoloohun-naa* (*yryt* ‘to analyze’ → *yrytyy* ‘analysis’, *bosxoloo* ‘to liberate’ → *bosxoloohun* ‘liberation’).

However, there are also cases when nouns derived from independent verbs with these two suffixes have ‘lexicalized’ (non-transparent) meanings alongside the fully transparent reading of action nominalization: e.g. *bys* ‘to cut’ → *byhyy* ‘cutting’ and ‘piece; shape, character’, *üün* ‘grow’ → *üünüü* ‘growing’ and ‘harvest’, *yjaa* ‘to weigh’ → *yjaahyn* ‘weighing’ and ‘scales’, *daqaa* ‘to touch lightly’ → *daqaahyn* ‘touching lightly’ and ‘adjective’, *battaa* ‘to press’ → *battaahyn* ‘pressing’ and ‘atmospheric pressure’. Whereas in the case of action nominalization reading all of the original verbal arguments are preserved, the argument structure is completely eliminated in the case of lexicalized meanings. Thus, *byhyy* ‘cutting’, *üünüü* ‘growing’, etc. pattern exactly like nominalizations in (44)-(45) whereas their lexicalized versions *byhyy* ‘piece; shape, character’, *üünüü*

²⁵ Also, syntactically nominalized verbs can be complex, including aspectual auxiliaries. In contrast, lexically nominalized verbs do not support aspectual modification.

‘harvest’, etc. are not associated with any arguments and behave like ordinary nouns. These data show that one and the same verb can be subjected to nominalization both in the lexicon and in the syntax. The two types of nominalization display systematic differences. First, meaning-wise syntactic nominalizations have an action reading derived from the underlying verb, lexical nominalizations are assigned a completely new meaning related but independent from the meaning of the corresponding verb. Second, with syntactic nominalizations the original argument structure of the verb is left intact, with lexical nominalizations it is reduced. If the base verb assigns accusative, this ability is preserved in syntactic nominalizations but disappears in lexical ones. Third, syntactic nominalizations cannot give rise to denominal verbs whereas lexical nominalizations pattern together with nouns like *malaahyn* ‘feast, banquet’ derived from bound roots in the lexicon (see above) in allowing denominal verbs. For instance, *daqaahynnaa* (derived from *daqaahyn* which is ambiguous between syntactic ‘touching lightly’ and lexical ‘adjective’) can only have meanings related to ‘adjective’ such as ‘provide with an adjective’, ‘turn into an adjective’, ‘invent an adjective’, ‘create epithets’, etc.

Another case of syntactic affixation is presented in (24) with the suffix –CI which derives agentive nominals (if a verb ends in a consonant, the suffix becomes –AAccY). We will assume that this suffix also saturates the external argument. If the base verb is transitive, the ability to assign accusative is preserved as (47) shows.

- (47) Terilte-ni salaj-aaccy kel-le.
 Company-acc manage-nom_{agent} come-past.3
 ‘The manager of the company came.’

The suffix has another syntactic use: it derives habitual participles.

- (48) Kini terilte-ni ücügejdik salaj-ar / salaj-aaccy.
 He company-acc well run-pres / run-habit
 ‘He runs the company well.’ / ‘He usually runs the company well.’

Whereas in the habitual function the suffix can apply to any predicate with no restrictions, as an agentive nominalizer it cannot apply to passives and unaccusatives. It is therefore similar to the English –er, cf. *aaq-aaccy* ‘read-er’, *üören-eecci* ‘study-er’, *oonnjoo-ccu* ‘play-er’, *tañas tik-ter-eecci* ‘tailor’s client (lit. clothes sew-caus-er)’ but **üün-eecci* ‘grow-er’, **timir-eecci* ‘drown-er’, **tökünüj-eecci* ‘roll-er’ (note that all these words are fine as habitual participles).

Thus, although we assumed that –CI saturates the external argument, its effect is not the same as the one caused by –YY/-hYn. The latter, unlike the former, can saturate the external argument of passives and unaccusatives. We believe the difference is structural in nature, witnessed by the grammaticality of denominal verb formation from –CI agentive nominals illustrated in (49) and its impossibility from –YY/-hYn syntactic nominalizations as discussed above (recall **bosxoloo-hun-naa* ‘liberate-nom-verb’).

- (49) a. Redaktsija-qa suruj-aaccy-laa-ty-byt.
 Editorial.office-dat write-nom_{agent}-verb-past-1pl

- ‘We worked as writers at the editorial office.’
- b. Redaktsija-ny suruj-aaccy-laa-ty-byt.
 Editorial.office-acc write-nom_{agent}-verb-past-1pl
 ‘We provided the editorial office with a writer.’

We would like to argue that syntactic –YY/-hYn nominalization is a high-level process taking an IP as input whereas denominal verb formation must take place before IP is merged. On the contrary, –CI suffixation takes a VP-input and derives an agentive nominal which can be further subjected to denominal verb formation: it will be shown in detail in chapter 7 that denominal verb formation, being a syntactic process, can apply at different levels reflected in the scope of the verbalizing suffix –LAA. Let’s consider why –CI suffixation is impossible with an unaccusative verb like *tökünij* ‘roll’. According to Theta system, this verb is derived from the corresponding transitive *tökünüt* ‘make roll’ by the lexical operation of causer reduction. *Tökünüt* ‘make roll’ is a two-place verb associated with [+c] and [-c-m] arguments. By the lexicon marking conventions of (75) in chapter 1,²⁶ the [+c] cluster will be marked with index 1 and [-c-m] with index 2. After [+c]-reduction has applied, only one argument marked with index 2 remains and this argument will be merged inside VP in the internal (object) position. Thus, when agentive nominalization applies there is no external argument for –CI to saturate, hence –CI suffixation is ruled out. Similar considerations carry over to passives. This result is very attractive because semantic restrictions on the type of external argument capable of being saturated follow not from the idiosyncratic properties of the suffixes but from the difference in the syntactic levels of application: one type of nominalization saturates IP-subjects (hence all kinds of subjects including the derived ones), the other type applies to VP-subjects. What kind of VP-subjects can be saturated receives a natural explanation under the Theta system. We will be coming back to the issue of IP- and VP-nominalizations in later chapters as well²⁷. Thus, the results of this section confirm the long-standing and well-studied distinction between, on the one hand, action/process nominals (also referred to as complex event nominals, argument-structure nominals or V-nominals, depending on the author) and, on the other hand, result nominals (also referred to as referential nominals or N-nominals) investigated in Abney 1987, Grimshaw 1990, Schoorlemmer 1995, Borer 2003b, Lebeaux 1986.

²⁶ Lexicon marking: Given an n-place verb-entry, n>1,

- a. Mark a [-] cluster with index 2.
- b. Mark a [+] cluster with index 1.
- c. If the entry includes both a [+] cluster and a fully specified cluster [/ α , /-c], mark the verb with the ACC feature.

²⁷ Our analysis of syntactic nominalizations bears similarity to the ones advanced in Abney 1987 who distinguishes three types of *ing*-nominals depending on the level of *ing*-attachment (IP, VP or V) and in Borer 2003b who treats argument-structure nominals as syntactic verbal structures dominated by nominalizing suffixes such as *-tion* or *-ing*.

2.5. Deriving adjectives in the lexicon and syntax

New adjectives can be derived from nominal, verbal as well as adjectival bases.

2.5.1. Deriving adjectives from nouns

Since nouns have no arguments and adjectives have one, deriving the latter from the former involves argument addition which is only possible syntactically. In chapter 4, section 4.5.2 we argue that one productive means of turning nouns into adjectives involves type-shifting ($\langle e \rangle \rightarrow \langle e, t \rangle$) along the lines proposed in Chierchia 1998. The adjectival nature of such property-turned nouns is supported by the fact that they take negation specific to adjectives, as will be shown in 4.5.2.

Type shifting is not marked overtly: a property-turned noun will appear in its basic shape when functioning as an adjective. Other syntactic means of converting nouns into adjectives involve overt morphology. These are cases of syntactic suffixation in which meanings of derived adjectives are determined compositionally, from the meanings of the base noun and of the suffix. In (38) one such affix was given. Other examples include –KI which derives various relational (directional, temporal) adjectives, –TAAqY which derives locative adjectives and –LAAx. The latter which decomposes into –LAA and –x will be considered in more detail in chapter 7 where it will be compared with the universal verbalizer –LAA.

Apart from these fully productive and transparent cases, other examples of noun-based adjective formation which are presented in section 7 of appendix 3, namely, (5) through (19) can hardly be viewed as syntactically derived. Note that out of 15 suffixes listed in (5) through (19) ten involve singleton adjectives and the other five are implicated in deriving either two or three adjectives. In these cases the most optimal solution is lexicon-internal derivation with subsequent freezing as discussed above in section 2.3.1: one of the adjectives discussed there was *sürün* ‘basic, fundamental’ derived from the noun *sür* ‘soul, internal spirit’ which has been out of use for several decades, yet its adjectival derivative had high frequency during this period; in this case we are literally dealing with freezing. Another pair with a history similar to *sür* → *sürün* is *bas* ‘head’ → *bastyn* ‘best’: the adjective is very frequent, the base noun *bas*, on the contrary, has a very limited use both because of its archaic flavour and because it has been losing ground to another noun *töbö* which also means ‘head’. Freezing is also supported by the fact that in the majority of derivationally related pairs in (5) through (19) the meaning connection has been broken and native speakers without linguistic training do not even suspect the existence of an original link between the two.

2.5.2. Deriving adjectives from adjectives

In this case both the input and output of derivation take one argument, therefore derivation is in principle possible both in the lexicon and syntax. This is indeed what we encounter. On the one hand, listed in (47) (appendix 3, section 9) are some highly idiosyncratic, sporadically occurring cases of A→A conversion in the lexicon. On the other hand, we find three cases of productive derivation involving

syntactic suffixes: -TYŋY which indicates incompleteness of quality, -CAAn which derives diminutives and -TAAqY which derives relational (locative) adjectives.

2.5.3. Deriving adjectives from verbs

All adjectives listed in section 8 of appendix 3 can be grouped into four types depending on their derivational history. The first two types are adjectives derived in the lexicon either from transitive (50) or intransitive (51) verbs. The former type involves argument reduction, the latter does not.

(50) $V_{\text{TRANS}} \rightarrow A$ lexical derivation: argument reduction; meaning change

V	Gloss	V-suffix → A	Gloss
Yk	Squeeze; urge, demand	Ygym	Tightfisted
Bys	Cut	Byha	Short, quick, direct
Ketee	Watch, keep an eye on	Ketex	Private
Kömüskee	Defend	Kömüskes	Defensive, protective
Bul	Find	Bulugas	Resourceful
Sie	Eat	Siemex	Carnivorous
Bil	Know	Biliges	Curious, inquisitive
Sim	Stuff in	Simik	Shy

(51) $V_{\text{INTRANS}} \rightarrow A$ lexical derivation: no argument reduction; meaning change

V	Gloss	V-suffix → A	Gloss
Xaam	Walk	Xamaqaj	Able, skilful
Aas	Pass	Aahygan	Transient, transitory
Bar	Go, leave	Barbax	Unimportant, insignificant
Xaam	Walk	Xamaqa	Saleable, marketable
Büt	End, finish	Bütün	Whole
Köp	Float to the surface (e.g. junk, dead fish)	Köppö	Lazy

The third type are adjectives derived in the syntax from intransitive verbs. This kind of derivation involves only syntactic category change from V to A: neither argument reduction nor meaning change can take place.

(52) $V_{\text{INTRANS}} \rightarrow A$ syntactic derivation: no argument reduction; no meaning change; category change

V	Gloss	V-suffix → A	Gloss
Kön	Straighten	Könö	Straight
Kilbij	Be timid, shy	Kilbik	Timid, shy
Tögürüj	Become round	Tögürük	Round
Kylgaa	Shorten	Kylgas	Short

Elbee	Multiply	Elbex	Multiple, a lot
Möltöö	Weaken	Möltöx	Weak
Sahar	Become yellow	Saharxaj	Yellow
Incej	Become wet	Inceqej	Wet
Djadaj	Become poor	Djadany	Poor
Kyryj	Become old	Kyrdjaqas	Old
Minnjij	Become sweet	Minnjiges	Sweet, delicious

The fourth type is also represented by adjectives derived syntactically. Whereas the adjectives in (52) are less finicky with respect to the choice of suffixes, the adjectives belonging to the fourth type attach the suffix *-m-t-YA* which results in the meaning ‘capable of V-ing, V-able’.

(53) ‘V-able’ adjectives

V	Gloss	V-mtYA → A	Gloss
Bar	Brew, intrans.	Barymtya	Strong (of tea)
Tuluj	Tolerate, trans.	Tulujumtuo	Enduring, tough
Soj	Cool down, intr.	Sojumtuo	Quickly cooling down
Kör	See	Körümtüö	Able to see, observe
Bul	Find	Bulumtuo	Able to find
Ihit	Listen, obey	Istimtie	Able to listen/to obey

The interesting thing about able-adjectives is that they can be derived from transitive verbs (e.g. *tuluj* ‘endure’, *kör* ‘see’, *bul* ‘find’, *ihit* ‘listen, obey’ in (53)) in which case they preserve the internal argument of the base verb along with accusative case. This proves once again that syntactic processes of category change, unlike lexical ones, are incapable of affecting the original argument structure of the derivational base.

- (54) a. Sargy itii-ni tuluj-ar.
Sargy heat-acc endure-aor
‘Sargy endures heat.’
- b. Sargy itii-ni tuluj-umtuo.
Sargy heat-acc endure-able
‘Sargy is capable of enduring heat.’

Intransitive verbs derived in the lexicon by unaccusative reduction can give rise to an adjective either in the lexicon or syntax. For instance, *köt* ‘fly’ is derived by unaccusative reduction from the corresponding transitive/causative verb *kötüt* ‘make fly’ (which takes a [+c] external argument). In its turn, it can be converted to an adjective either in the lexicon resulting in *kötümex* ‘negligent, slipshod’ (non-transparent, lexically fixed meaning) or in the syntax resulting in *kötümtüö* ‘able to fly’ (compositional able-meaning).

One important question has not yet been addressed. It is claimed here that no categorization takes place in the lexicon and that categories such as nouns, verbs and adjectives constitute a purely syntactic phenomenon. What takes place in the

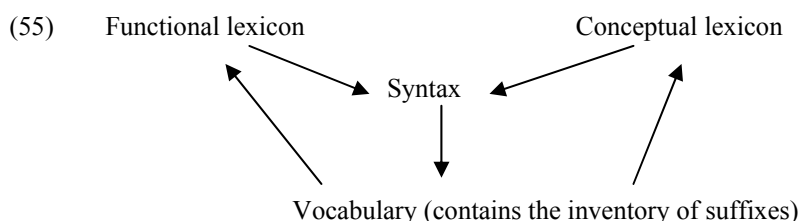
lexicon, however, is a classification of concepts according to their thematic properties. At the interface of the theta system with the syntactic component it is this prior classification which determines whether a particular concept will merge under a syntactic N, V or A node: two-place concepts will be merged as verbs, one-place concepts associated with a [-c-m] feature cluster will merge as adjectives. However, we are also assuming a number of lexicon-internal operations on θ -grids of verbs and adjectives which can alter the original number of arguments, either by increasing or decreasing. Whereas it is clear that a basic adjective with its sole argument reduced will come out as a noun in the syntax, the situation is much less clear for verbs. Consider, for example, unaccusative reduction which applies to two-place concepts with a [+c] argument, reduces the latter and the result is an unaccusative verb. As already mentioned, such unaccusative verbs can in turn give rise to adjectives either in the syntax or lexicon. Some of the important questions which must be answered are why causer reduction which results in a one-place concept yields an unaccusative verb and not an adjective, how exactly an intransitive verb can be converted into an adjective in the lexicon and in the syntax and how do these two processes differ from each other. (50) analyzed as argument structure reduction leading to the formation of new adjectives also raises the question of how to distinguish cases of transitive verbs giving rise to adjectives and of transitive verbs giving rise to unaccusatives. In this section we have only touched upon the tip of the iceberg by uncovering four possible types of derivation leading to the formation of new adjectives. Doing more than that requires better familiarity with the thematic composition of verbal concepts in Sakha – a topic undertaken in chapter 5. Therefore we will postpone considering the questions just mentioned until then.

2.6. Concluding remarks

This chapter considered the morphological aspect of categorization. On the one hand, we argued that given suffixal ambiguity it is hard to rely on the morphological means of categorizing lexical items. On the other hand, the existence of three-way morphological asymmetries supports the assumption that there exist three basic categories: concepts will always be less marked in their primary categorial incarnation than in their secondary one. However, it is not claimed that categorization takes place in the lexicon. On the contrary, categories such as nouns, verbs and adjectives are a purely syntactic phenomenon. What takes place in the lexicon is a classification of concepts according to their thematic properties. At the interface of the theta system with the syntactic component it is this prior classification which determines whether a particular concept will merge under a syntactic N, V or A node (i.e. become a primary noun, verb or adjective).

We also argued that lexicon is not only the storage place for existing concepts but also the locus of new concept formation. A number of arguments were presented against the DM-tenet that no derivation is possible lexicon-internally. Lexical derivation is opposed to syntactic derivation along the criterion of meaning change from which it follows that only in the lexicon can thematic properties of concepts be altered. In other words, only lexical derivation is capable of reducing or

modifying the original θ -grid of a concept. The following organization of the modules having a relation to morphological derivation was proposed.



To conclude, we would like to present one more piece of evidence for the existence of three basic categories – from morphophonology. Don (1993:171-3) gives a summary of Trommelen 1989 where “it is argued that there is a relation between the syllable structure of Dutch underived words and their category. The general observation is that verbs have a very restricted syllable structure, while the syllable structure of adjectives is less restrictive and nouns have the most ample possibilities” (Don 1993:171). The maximal number of consonants allowed in the syllable onset in Dutch words is three, whereas the rhyme can have five segmental positions as in *herfst* ‘autumn’. However, whereas all three categories can have the maximal onset, the maximal rhyme is only encountered in nouns: “there are hardly any verbs or adjectives boasting five positions in their rhyme; nor are there any verbs having a four positional rhyme, although this is possible in adjectives and nouns” (Don 1993:172). Apart from syllable structure, Don 1993 also demonstrates, based on Trommelen 1989, that there is a correlation between the number of syllables and the type of category such that underived verbs in Dutch are mostly phonologically monosyllabic. The same kind of correlation also exists in Sakha. Basic underived verbs and adjectives (root verbs and root adjectives) are mostly monosyllabic, root nouns can be polysyllabic. Verbs can never end in a short vowel, nouns and adjectives can. These and other phonological diagnostics help decide the category of nonce words which do not otherwise exist in the language. For instance, both *sam* and *mellej* can be assigned to all three categories, *soxxoroon* can be viewed as a noun or adjective but never as a verb whereas *soxxoroonju* can only become a noun. Thus, whereas lightness fits all categories and middle weight qualifies for nouns as well as adjectives, the heavier a word the more likely it is to become a noun. An explanation for these curious facts seems not yet to have been offered but the correlations themselves have a practical significance: on the one hand, they can be used to determine the direction of derivation in conversion as is done in Don 1993 and, on the other, they clearly argue for the existence of three basic categories²⁸.

²⁸ Morphological subclasses may also be phonetically determined. For instance, Bybee & Moder 1983 (also Bybee & Slobin 1982) show that the most productive class of strong verbs in English (win/won) is phonologically defined. Zubin & Köpcke 1981 argue for German that there is a strong correlation between the initial consonant clusters in monosyllabic nouns and the assignment of nouns to the masculine gender: 73% for those with an initial two-consonant cluster and 83% if the number of consonants is three. Phonetic determination of major word subclasses is also encountered, e.g. all adverbs

in Yoruba are ideophonic whereas relatively few nouns, verbs and adjectives are (Schachter 1985). Another example is Edo where all nouns begin with a vowel (Baker 2003:121n.).