

Pauli Brattico

A Category-free Model of Finnish Derivational Morphology

Abstract

The lexicon is traditionally understood as consisting of lexical items, which are categorized in lexical categories such as verbs, nouns or adjectives. Recently, this assumption has been challenged by a theory which posits no lexical categories in the lexicon. Rather, lexical items are taken to be categorially underspecified roots. This article presents a theory of Finnish word formation which supports, and is based on, a category-free model of the lexicon. It is argued that the category neutral layer of word formation in Finnish is recursive, hence likely to be part of the syntax proper. Some implications are discussed.

1. Introduction

Recently, some linguists have argued that the lexicon contains categorially underspecified roots, rather than nouns, verbs and adjectives (Adger 2003, Alexiadou 2001, Barner & Bale 2002, Chomsky 1970, Chomsky 2004, Giegerich 1999, Farrell 2001, Jeanne & Hale 2000, Marantz 1997, 1999, 2000, Hale & Keyser 2002, Harley & Noyer 1999, Pesetsky & Torrego 2004, Salo 2003, Whitman 2004).^{1,2} According to one such view (e.g.,

¹ Preparation of this paper was supported by Finnish Cultural Foundation and the Academy of Finland (project number 106071). The basic ideas presented here were first formulated in my PhD thesis (Salo 2003). The present paper was first presented at the SKY symposium “The lexicon: its status in the theory of language” (2004, Turku). I thank the audience of the symposium for comments. Saara Huhmarniemi, Christina Krause, Markus Mattsson, Jouni Rostila and two anonymous referees gave me valuable comments at various aspects of this work. I can’t imagine what this work would be like without their time and patience. Julie Uusnarkaus helped me with the English language, although all the remaining mistakes are of course mine.

² Jouni Rostila (p.c.) pointed out that this proposal is essentially compatible with various Construction Grammar approaches. Within Construction Grammars, lexemes can be conceived of as having rich frame-semantic meanings, and their categorial properties

Marantz 1997, 1999, 2000), categorially underspecified roots become nouns, verbs and adjectives when inserted into specific syntactic contexts. The question then arises whether this new model can be applied to Finnish or, more interestingly, whether properties of Finnish word formation support or reject such a view. Furthermore, we want to know whether the category free model could help us to solve some of the remaining mysteries of Finnish word formation. I argue that the answer to both questions is positive. Section 3 presents one category free model of Finnish word formation and section 4 applies the theory to a number of open problems in Finnish. What emerges is not so much a completely new model of Finnish word formation, but a fine-tuning of the standard theory (e.g. Karlsson 1983) under a somewhat different theoretical orientation.

Before going to the main business of this article, I want to clarify some terminological matters and explicate certain features of my theoretical orientation. I take “lexicon-as-listedness” to refer to elements which are the output of no grammatical rules. It is a storage of all primitive elements, often referred to as “linguistic features.” These can include derivational morphemes, inflectional morphemes, lexical category features, agreement features, semantic features, concepts, wider cognitive categories and phonological features – in short, all the grammatical elements that are primitive. This list has to be composed by means of empirical, not conceptual investigation. Some of these feature combinations produce possible words and phrases in a language. Lexicon as ‘derivational morphology’ refers to the output of applying word formation rules to the elements in the lexicon-as-listedness, so that the resulting objects constitute the domain of syntax proper. To follow standard terminology, I call them lexemes. Intuitively, these constitute ‘possible words’ in a language, assuming that inflection takes place in syntax. ‘Psycholinguistic lexicon,’ or lexicon-in-use, is the domain of cognitive processes involving actual language use, parsing, the effects of word frequency, automatization, and so forth. This list contains a catalogue of words memorized by an individual speaker or a list of words shared by a community of speakers. These are potentially very large feature bundles chunked together.

Take the word *juoksu-tta*- ‘run-CAUSE.’ This is a complex lexeme from the perspective of derivational morphology, but it might constitute

arise when they are unified with schematic constructions. Since schematic constructions are Construction Grammars' means of capturing syntactic phenomena, this means in practice that categorial properties arise in the syntax.

one atomistic unit in the psycholinguistic sense if, for example, it is used with considerable frequency. It might, furthermore, be accessed by using a whole-word strategy instead of a decompositional strategy in an actual context of use (see Baayen 1993). We expect and will later find a lot of variation between speakers concerning the constitution of their lexicon-in-use. From the perspective of lexicon-as-listedness, *juoksuttaa* contains at least three elements: *juokse-* (the base), *-tta-* (causative morpheme) and *-a* (the marker for the first infinitive). It could be found to contain much more, if we are allowed to penetrate deeper into its structure. The word *hammastaa* ‘teeth-CAUSE’ is a potential word in Finnish as well, yet it is seldom (if ever) used in modern Finnish, so it belongs to the lexicon as derivational morphology, but possibly not to the typical lexicon-in-use. It might belong to the lexicon-in-use of a dentist specialized in gerontology, for example. Unless otherwise stated, ‘lexicon’, as it is used here, does not mean lexicon in the psycholinguistic sense, but in the linguistic sense.

I recognize that this orientation differs significantly from those of many others. For what it is worth, my motivation for separating the lexicon-in-use from the two other linguistic notions is the fact that no theory of language *use* can be restricted only to the linguistic domain. For instance, almost any type of cognitive material can be stored and manipulated as a single item or by means of a “whole-item strategy,” given enough practice and repetition (Logan 1988). Thus, to understand the notion of lexicon-in-use, we would need a general psychological theory of automatization, rather than only a linguistic theory of word formation. To assume that such a general theory could succeed without the more fundamental linguistic theory is as mistaken as the assumption that chess psychology could start without acknowledging the rules of chess, but equally it makes little sense to try to explain such general psychological matters only inside of the linguistic domain. Hence, some current trends notwithstanding, I think that we have to keep the two domains distinct. Ultimately, the issue is empirical. We will have the chance to return to this matter on several occasions.

Since the lexicon-as-listedness contains only primitive features, it cannot contain linguistic elements which have been assigned to some lexical category; such elements are automatically ‘complex’ by virtue of being composed out of something plus a lexical category feature such as V, N or A. The interesting question here is whether morphemes, derivational or inflectional, are complex elements in the sense of being provided with one

of these categories. I will argue in this paper that derivational morphemes and the resulting complex lexemes are categorially underdetermined.

To introduce some terminology, by “categorially underspecified Root” I mean a morpheme, stem, or lexeme which does not belong to any lexical category. This terminology comes from Giegerich (1999). Intuitively speaking, these are constituents of words that do not contain features such as +N, V, A. To follow Pesetsky (1995), I use the notation $\sqrt{\text{ROOT}}$ to refer to Roots. Roots are lexical elements which can become verbs, nouns and adjectives when they become Words. Words are lexical items which are, in themselves and without further ado, pronounceable as grammatically well-formed units.³ Whether a given Root becomes a verb, noun or an adjective depends on its syntactic context. For instance, the Finnish Root $\sqrt{\text{BUY}}$ can become a verb (*osta-a*), noun (*osta-minen*, *ost-o*) or adjective (*osta-va*). *Osta-* is the phonological exponent of a categorially underspecified Root that cannot be pronounced alone without certain minimal inflectional markers and/or a marked stress pattern.⁴ Each Root projects thematic roles, which are associated with argument DPs (determiner phrases). The syntactic realization of these DPs depends on the categorial status of the Root, as shown in these Finnish examples (1a–d) (all examples, unless otherwise stated, are from Finnish). When the Root is combined with relevant inflectional material, they become phonological Words, which are pronounceable as such.

- (1) a. *isä* *osta-a* *auto-n* (V)
 father-NOM buy-3SG car-ACC
 ‘the father buys a/the car’
- b. *isä-n* *auto-n* *osta-minen* (N)
 father-GEN car-GEN buy-N.NOM
 ‘the buying of a/the car by the father’

³ While also Roots, like cranberry morphemes, have pronounceable phonological exponents, these elements are not grammatical and understandable without further affixation.

⁴ The root *osta-* by itself can be used as an imperative form in Finnish. However, in the case of derivationally complex Roots, the imperative form and the Root form differ. Thus we have *osta-tta-* ‘to cause to buy’ as a causative Root, whereas the imperative form of the same Root is *osta-ta*. The causative Root *osta-tta-* cannot be used as a Word in any context.

- c. *isä-n* *osta-ma* *auto* (A)
 father-GEN buy-A car
 ‘a/the car bought by the father’
- d. *auto-n* *osta-va* *isä* (A)
 car-GEN buy-A father
 ‘a/the father who buys a car’

Note that according to my definition, Roots can be either complex or simple. Indeed, I will argue that there are complex Roots along with simple roots. This terminology comes from Giegerich (1999). The notion of Root is almost identical to the notion of ‘categorially underdetermined lexeme,’ except that it includes the individual morpheme constituents of lexemes in addition. For instance, I will argue that causatives such as *osta-tta-* ‘buy-CAUSE’ are category-free Roots, but so is the causative morpheme itself. This lexical element has been composed by merging two Roots together, which together constitute another, complex Root ($\sqrt{\text{BUY}} + \sqrt{\text{CAUSE}} = \sqrt{\text{BUY} + \text{CAUSE}}$).

Furthermore, I do not reject the reality of lexical categories altogether; rather, I assume that they are part of the conversion process of Roots into Words. Inside of Roots, they do not have any status. For instance, it is the Root–Word conversion process which explains why $\sqrt{\text{FISH}}$ appears as a zero-derived form inside of an NP and with a copula inside of a VP, while for $\sqrt{\text{RUN}}$ the situation is the converse in that the VP context gives the zero-derived form, and the nominal context requires the presence of overt morphemes. In the case of complex stems such as *osta-tta-* ‘buy-CAUSE’, categorization is always overt, since *ostatta-* does not constitute a Word in *any* context, so there is no asymmetry in its phonological form with respect of the lexical categories.

I assume that any complex linguistic element, be it a word or a phrase, can obtain an idiomatic semantic interpretation without losing its syntactic/morphological complexity. Thus, *kick the bucket* means in some contexts ‘to die,’ but it is inflected as if it were a complex phrase (*kicked the bucket*, not **kick the bucket-ed*). Moreover, it still has the literal interpretation. Thus, semantic opaqueness is by no means a good argument for syntactic atomicity (Marantz 1997). The same is true of words: *girlfriend* is a complex word even if it means (in some contexts) something else than ‘a friend, who is also a girl.’ Here the morphologically and syntactically complex element has obtained idiomatic semantic features, which are stored in the lexicon-in-use. This assumption is motivated also

by the fact that in some cases my informants were aware of both the idiomatic and the compositional meanings of complex words, which thus exists side-by-side.

2. Cross-linguistic evidence for category neutral roots

Before developing and testing the model on Finnish word formation, I will explain briefly why many linguists have been interested in the category-free theory. These constitute in my mind persuasive arguments in favor of the model, which do not rely specifically on Finnish word formation.

In some languages there is direct evidence of category neutral Roots. For instance, in Semitic various semantically related words can be produced by altering the vowels between a sequence of consonants. The consonant sequence itself is not associated with any lexical category. Example (2) comes from Hebrew.

- (2) a. g.d.l – ‘big’ as a root that is never used in isolation
 b. gadol – ‘big’ as an adjective
 c. giddel – ‘be magnified’ as a verb

In addition to the Semitic languages, several other languages such as Wintu (Pitkin 1984), Tagalog (Gil 1995), Jimgulu (Pensalfini 1997), Tuscarora (Williams 1976), Salom (Mosel & Hovdhaugen 1992), Wolof (Mc Laughlin 2004) and Cayuga (Sasse 1993) are arguably best described as containing categorially underspecified roots. Mosel and Hovdhaugen (1992) describe Samoan as follows:

Many, perhaps the majority of, roots can be found in the function of verb phrase and NP nuclei and are, accordingly, classified as nouns and as verbs. This does not mean that a noun can be used as a verb or a verb as a noun or that we have two homophonous words, one being a noun and the other being a verb. Rather, it means that in Samoan the categorization of full words is not given a priori in the lexicon. It is only their actual occurrence in a particular environment which gives them the status of a verb or a noun [...] What is given in the lexicon, is not a particular word class assignment, but the potential to be used in certain syntactic environments as a noun or a verb. (Mosel & Hovdhaugen 1992: 77)

In Finnish, in contrast, many complex stems are non-words, which then leads essentially to the same kind of model. Secondly, from the typological perspective it is reasonably clear nowadays that dichotomous features such as N, V and A do not exist (Baker 2003, Dixon & Aikhenvald 2004,

Hengeveld, Rijkhoff & Siewierska 2004, Rijkhoff 2002, Vogel & Comrie 2000). Perhaps nowhere is the need to reject the dichotomous lexical categories as evident as in the case of adjectives (Dixon & Aikhenvald 2004). Rather, there is a loose category of nounhood and verbhood, which are constituted by prototypes listing the typical properties of nouns and verbs. Lexical categories must be dissolved into several independent properties which correlate with each other. In statistical terms, a lexical category is like a factor: a cluster of properties correlating with each other. Thus, labeling lexical elements with the dichotomous categories N/V/A prior to syntactic computations does not seem to lead to a tenable theory.

Studies of agrammatic patients with selective Noun–Verb dissociations show that while it is true that agrammatic patients can have selective problems with verbs and nouns, these deficits seem to extend also to pseudonouns and pseudoverbs (Caramazza & Shapiro 2004). This indicates that the deficit has to do with some form of productive morphosyntax, suggesting that categorization is part of some rule-based component of the grammar.⁵ More importantly, virtually all Words are associated with a lexical category. If lexical categories were not be associated with lexical elements in some rule component, the fact that there are twenty thousand nouns in somebody’s lexicon would be a miracle, much as if all nouns in somebody’s lexicon would happen to represent entities which are all red (blood, the flag of former Soviet Union, fire truck, etc.) and of no other color. Because lexical categories are based on grammatical rules, they must be, in principle, dissociable from the lexicon-as-listedness which is the domain of all rule-like processes. To put it simply: if the standard theory of the lexicon says that lexical elements are constituted by structure [$\sqrt{\text{CAT}+\text{N}}$], then these complex lexical elements must have been composed *somewhere* out of the feature +N and a sublexical feature $\sqrt{\text{CAT}}$, as follows:

⁵ It is also worthwhile to note that one type of mistake children commonly make is that they nominalize and verbalize beyond adult grammars (*I’m going to basket those apples*), use verbs in nominal contexts (*Where’s the shoot*) and vice versa (*Mommy trousers me*) (Barner & Bale 2002, Berman 1999). As noted by an anonymous referee, the relevance of this evidence is weakened due to the fact that English has much categorial homonymy.

$$(3) \quad \begin{array}{c} \sqrt{\text{CAT} + \text{N}} \\ \wedge \\ +\text{N} \quad \sqrt{\text{CAT}} \end{array}$$

Here one may assume that $\sqrt{\text{CAT}}$ and $+\text{N}$ originate from the lexicon-as-listedness, whereas the complex entry $[\sqrt{\text{CAT}+\text{N}}]$ belongs to derivational morphology, being a possible feature combination in many languages. But then $\sqrt{\text{CAT}}$ must initially be a category-free element. It thus looks as if the postulation of category-free roots is simply inevitable, at least in relation to the lexicon-as-listedness. Furthermore, it would make no sense to say that the category-free features are excluded from the domain of linguistics, since (3) is a linguistic rule on the basis of its inputs and outputs. To me, the only controversial question is then the issue of how the category-free elements interact with derivational and inflectional morphology, and especially with syntax, not whether they exist or whether they are “linguistically relevant.” For instance, Karlsson’s (1983) model of Finnish word formation does not acknowledge any status to rule (3), but there is no convincing case to be made, I think, for the hypothesis that rule (3) is not “linguistic” or otherwise irrelevant.

There is a *prima facie* argument which supports the contention that lexical category is attached to the root only in the syntactic component of the grammar, so that process (3) becomes part of the core syntax, as argued by Marantz (1997, 1999, 2000). Nominalization, verbalization and adjectivization are productive and systematic processes which seem to take place in syntax, at least in some cases. As we will see, this is so in the case of the Finnish *-minen* nominalization, for instance (Hakulinen & Karlsson 1979). If it is true that the category of a stem/word can be determined in the syntactic component, the question arises whether we need to assume, on top of this, that it can be assigned *also* in the lexical component. Since there are independent reasons to assume that the lexical category of a word can be determined in the syntax, the null hypothesis is to assume that it is determined *only* in the syntax. As it turns out, Finnish word formation provides further evidence for this hypothesis.

The fact that the X-bar theory, created in the 70s for purposes of describing certain generalizations concerning syntax, seems descriptively correct, represents one of the strongest cases in favor of the category-free lexicon. The leading idea of the X-bar theory is that a substantial amount of regularity in syntax is insensitive to lexical categories, but refers only to

abstract relational notions such as complement-of-X and specifier-of-X where X can be a head of any category – thus, for example, either a verb, noun, preposition, or an adjective. The following list provides examples of a verb phrase, an adjective phrase and a noun phrase, each instantiating the same underlying structure containing a head (*envy*), complement (*Mary*) and a specifier (*John*).

- (4) a. John envies Mary (VP)
 b. John's enviousness of Mary (NP)
 c. John's envy of Mary (NP)
 d. John is envious of Mary (AP)

This raises another question of no less importance: why is this true of human language(s)? The assumption that the syntax does not see lexical categorial information in the first place provides one explanation: if lexical categorial information is invisible at the level of syntax, then substantial evidence in favor of the X-bar theory would emerge, X^0 being the category neutral lexical element. Thus, extracting categorial information out of the lexicon provides a way to explain some properties of the X-bar theory itself (Salo 2003: 106–107).

Finally, categorially underspecified roots are semantically relevant, as each of the *envy* morphemes in (4a-d) are semantically related: basically, they represent enviousness. The root $\sqrt{\text{ENVY}}$ can capture this common conceptual meaning behind the various words. I propose in this article that lexical roots are the linguistic counterparts of concepts: mental symbols which are constituted over and above by their meaning (Chomsky 2005, Fodor 2003: 152–158, Salo 2003: 69–76, 123–126). This is the standard assumption in formal semantics literature, which piles up intransitive verbs, nouns and adjectives all into the semantic category of one-place predicates and thus claims that they all correspond to the same category-neutral semantic entity (e.g., Heim & Kratzer 1998).

To sum up, there are a number of reasons to assume that there is a notion of lexicon which contains only categorially underspecified roots, not verbs, nouns or adjectives. I will argue next that properties of Finnish word formation support the same conclusion.

3. The proposal

In this section I will define the proposal that describes the properties of Finnish word formation in the best possible way and is in agreement with the kind of category-free lexicon that the evidence cited in the previous section supports. I present a few preliminary arguments in favor of the theory and show how it organizes the facts around Finnish derivational morphology.

I begin with Karlsson's (1983) comprehensive model of Finnish derivational morphology. According to Karlsson's theory, suffixes can be located in specific positions after the root. The root is specified for its category, so (3) falls out of the domain of Finnish word formation. There are ten positions, in the following order: three for V affixes (1–3), one for passive (4), two nominal positions (5–6), two adjective positions (7–8), one nominal position (9) and, finally, all inflections (10) (Karlsson 1983: 244). Each position can be left empty, and each position can contain a closing suffix so that the derivation stops at that point. For example, nominal marker *-minen* occurs in the first nominal position (5) but positions (6–9) must be left empty. To illustrate the model, consider a nominalized form of a triple causative:

- (5) *tee-tä-ty-ttä-minen*
do-cau-cau-cau-n
'causing ... to do'

The three causative suffixes fill the three verbal positions, after which comes the nominalizer from position 5.⁶ From this position, one cannot continue derivation. Yet Karlsson admits (*ibid.*: 241) that more than three verbal affixes can be stacked on top of each other; see also section 4.3 of this paper. Even four causative morphemes can be added to a root, resulting in only a sense of marginality:

- (6) *?tee-tä-ty-tä-ttä-minen*
do-cau-cau-cau-cau-n
'causing ... to do'

⁶ To be exact, the correct analysis of *tee-tä-ty-ttä-minen* seems to be *tee-tä-tä-y-ttä-minen* 'do-CAU-CAU-REFL-CAU-N,' where *-y-* is the exponent of reflexivization. I will ignore this detail here.

Karlsson excludes such words from the model because they are “marginal” albeit not “ungrammatical” (241) and because their semantics is no longer transparent (238). Similarly, Hakulinen et al. (2004: § 306) found that only three V affixes can be stacked on top of each other if the actual use of the language is used as a criterion. Longer words become marginal, pragmatically deviant, more difficult to understand, hence such words are unusable (compare *under-undercover agent* and *anti-anti-missile*), but the combinatorial process itself is without doubt recursive. I will allow recursive stacking of V affixes instead of three independent positions, but there is hardly any empirical issue at stake here, given my ignorance of the lexicon-in-use.⁷

In Karlsson’s model, nominal suffixes follow verbal suffixes. Some of these suffixes are closing suffixes, which is a stipulated fact in Karlsson’s model. Karlsson however claims that, excluding some exceptions, complex nominals cannot be verbalized (*ibid.*, 236–237). Because of this, he ends up modeling Finnish word formation by means of rigid positions (1–10) without loopback. Assuming that the three V positions are actually filled with a recursive loop of verbal affixes, such as causatives, this implies that nominal affixes are merged after the verbal loop. Yet as Karlsson himself admits, there are several complex nominals which allow recursion back to the beginning. Furthermore, in some cases complex nominals can be causativized, which brings the derivation back to the inner recursive V layer. Potential examples are provided in (7a-k).⁸

⁷ In chess psychology, to pursue again another well-studied example, we have to make a difference between possible positions which do not make sense to experienced chess players and positions which do (Saariluoma 1995). The distinction is largely irrelevant for novice players. The difference is a function of the frequency which such positions are encountered in actual games and in the chess literature, more exotic variations being infrequent and largely unexplored territory. Here we can choose to study chess either as a dynamical, recursive system of rules, or as something represented in the mind of experienced chess players. The former offers a more abstract but at the same time more fundamental basis of inquiry, whereas the latter is a function of the particular experiences of the given players. Both views are equally important; which one we choose to study does not, in and itself, commit us to any empirical claims about chess or chess psychology.

⁸ When a particular example did not exist in the current Finnish lexicon-in-use, I sought an analogous derivation. An actual record of usage was obtained from the Finnish corpus composed by the Research Institute for the Languages of Finland, the Finnish IT Centre for Science and Department of General Linguistics, University of Helsinki. The

- (7) a. *hampa-*, *hampa-isto*, *hampa-isto-ttaa* (*hampaistottaa*) (>*puistottaa*)
teeth, teeth-COL, teeth-COL-CAUSE (tree-COL-CAU)
'teeth, a collection of teeth, to cause a collection of teeth (to cause to have parks)'
- b. *hiki*, *hiki-tys*, *hiki-tys-tää* (*hiestystää*)
sweat, sweat-N, sweat-N-CAUSE
'a sweat, causing to sweat, to cause to sweat'
- c. *jrysi-*, *jrysi-in*, *?jrysi-in-tää* (*jrysintää*) (>*viestintää*)
bite, bite-IN, bite-IN-CAUSE (message-N-V)
'to bite, the thing that is used for biting, to cause (to have) the things that are used for biting (to broadcast)'
- d. *toimi-sto-tta-* (*toimistottaa*), *kalu-sto-tta-*, *vesi-stö-ttä-*, >*pui-sto-tta-*
act-N-V, thing-N-V, water-N-V, tree-N-V
'to cause to have/be an office, to cause to have/be furniture, to cause to have/be water, to cause to have many trees/cause to have parks'
- e. *hampai-stus-ta-* (*hampaistustaa*), *kalu-stus-ta-*, *kala-stus-ta-*, >*avu-stus-taa*
teeth-N-V, thing-N-V, fish-N-V, help-N-N
'to cause to be/have a collection of –'
- f. *juoksu-(t)t(a)-in-* (*juoksutin*), *paalu-(t)t(a)-in-*, *syö-(t)t(ä)-in-*
run-V-N, pole-V-N, eat-V-N
'an instrument for causing to –'
- g. *tutki-nto-ttaa* (*tutkinnottaa*)
research-N-V
'to cause (academic etc.) degrees'
- h. *lomaile-u-ttaa* (*lomailuttaa*), *arvele-u-ttaa* (*arveluttaa*)
holidaying-N-V, suppose-N-V
'cause to have a vacation, to cause to suppose/wonder'
- i. *laahaa-us-taa* (*laahustaa*), *etu-us-taa* (*edustaa*)
drag-N-V, front-N-V
'crawl, represent/stand for'
- j. *pehmeä-us-taa* (*pehmustaa*), *helma-us-taa* (*helmustaa*)
soft-N-V, hem-N-V
'soften, to cause to have hems'

- k. *ime-u-ri-oida* (*imuroida*), *puske-u-ri-oida* (*puskuroida*)
 suck-N-N-V, buck-N-N-V,
 ‘vacuum-clean, to buffer’

Based on Karlsson’s own examples and data such as this, there appear to be two kinds of nominal affixes: those which allow the derivation to continue (e.g., *-o* in *tule-o-ttaa*) and those which do not (e.g., *-inen*, *-minen*, *-ma*, *-na*; **punainentaa* ‘red-CAUSE’).

An alternative explanation for the data in (7) is to claim that the nominal stems in (7) are not just complex nominals, but lexicalized complex nominals. It is well known that causativization can apply to bare nominals in Finnish, as well as in many other languages (*paalu-tta-* ‘to pole’). This theory is supported by the fact that many causatives of complex nominals, of which there exists a record of actual usage in my sample, are highly lexicalized (Hakulinen et al. 2004: § 306). The word *puistottaa* ‘to cause to have parks’ (7b) is derived from *puisto*, which is a lexicalized noun in Finnish referring to parks. But it is itself composed from *puu* ‘tree’ and *-isto* ‘a collection of –,’ so that the compositional meaning of the word is ‘a collection of trees.’ Because of its lexicalized use, there can be a *puisto* even if there are no trees, which contradicts the compositional reading of the term. The question is then if it is possible to form a novel nominal and causativize it as well. To me, words such as *hampa-isto-ttaa* ‘to cause a collection of teeth’ are novel, understandable and possible words in Finnish, especially when given some extra-linguistic context. Therefore, it is necessary to look also at possible but nonactual words, since these are items for which the effects of lexicalization can be best controlled. Because the rule which causativizes complex nominals works even in the case of nonactual, but possible words, the data cannot be explained solely by relying upon lexicalization.

These considerations bring us to the problem of saying what counts as a “possible word” in a language. The above judgments are based on the present author’s judgment of grammaticality and semanticity, which could be contaminated by theoretical bias. On the other hand, in order to study the combinatorial potential in a language, whether in the domain or syntax or lexicon, we cannot rely solely on instances that the subject has actually heard and used frequently, if only because such instances are also contaminated by properties of the lexicon-in-use which we have to control experimentally. To investigate the matter further, I presented the above data to a range of Finnish-speaking informants to obtain their judgments on

semanticality and grammaticality.⁹ The subjects were asked first to rate the words in the test by their grammaticality (on a scale of 1–5: 1 completely impossible word, 3 possible word, 5 very possible word) and then to provide a semantic interpretation (if it had any). Semantic interpretation was encoded quantitatively so that 0 = no interpretation, 1 = interpretation. Four categories of words were used: (1) monomorphemic Finnish words as fillers and control items (*virsi* ‘gospel’, *talo* ‘house’, *Suomi* ‘Finland’, *matto* ‘carpet’), (2) derivationally complex words which are used idiomatically in current Finnish (*imuroida*, *laahustaa*, *rokottaa*, *tulospalvelu*), (3) derivationally complex words which are possible according to the above model but not idiomatic (*hampaistottaa*, *juoksutin*, *hikistyttää*, *kalustottaa*, *puistottaa*, *juoksuttaa*, *tutkinnottaa*, *puistollinen*, *puistollistaa*, *hampaallistaminen*) and (4) derivationally complex words which involve an ungrammatical lookback from a closing N suffix (**hyppääminen*, **juokseminentaa*, **hyppimineninen*, **punainentaa*, **puistotintaa*). The present model predicts that items on the group 4 should not have a coherent semantic interpretation, whereas items on the three other lists should have. Furthermore, the model predicts that items in category 2 should be classified on average as “fully possible words” (belonging to the lexicon-in-use), items in category 3 as “possible words” (belonging to the class of possible words only) and items in category 4 as “impossible words.” The results agreed with these predictions. Average grammaticality and semanticality judgments obtained from this test were as follows:

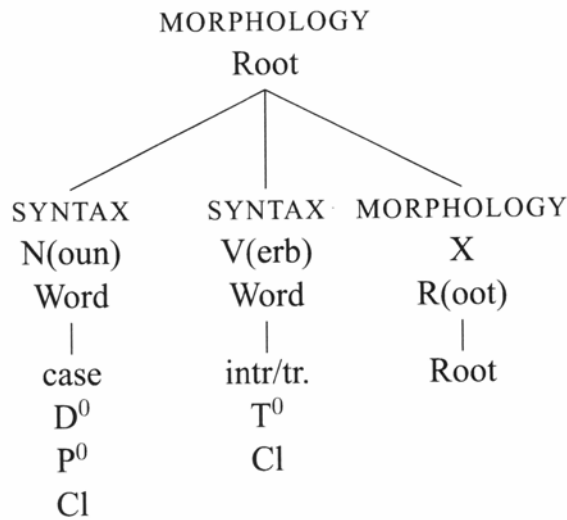
⁹ All informants (7 male, 9 female) were above 20 years of age, most but not all studying or working at the university. Their judgments were obtained by means of a questionnaire.

GROUP	TYPE	semanticity	grammaticality	rating
1	simple word	1.00	5.00 ± 0.00	“fully possible word”
2	complex word, idiom	1.00	5.00 ± 0.00	“fully possible word”
3	complex word, novel	0.89 ± 0.07	3.30 ± 1.04	“possible word”
4	complex word, impossible	0.20 ± 0.11	1.29 ± 0.14	“impossible word”

Table 1. Average grammaticality and semanticity judgments obtained in the present study. The behavioral semanticity and grammaticality data (dependent variables) were analyzed separately. These data were entered into repeated measures analyses of variance (rmANOVA) with one independent factor (experimental category; four levels: 1, 2, 3, 4). The main effect for the factor experimental category was statistically significant for both semanticity ($F(2, 14) = 118.7, p = .000$) and grammaticality ($F(2, 24) = 287.4, p = .000$). These effects arose due to differences between groups 2, 3 and 4 $p < .01$ for both semanticity and grammaticality, specifically. Note that for judgments of semanticity, the novel words were very similar to the established idioms (groups 2, 3).

This seems to confirm my own intuitions about the grammaticality and semanticity of these words, and thus the distinction between the two kinds of nominal suffixes. I therefore hypothesize that there is a first layer of processes which turns stems into new stems, and which may include both V affixes and N affixes, as traditionally understood. The resulting stems have the property that they are still free to turn into adjectives, nouns or verbs when suffixed with appropriate suffixes or inflectional markers. All causatives can be verbalized (*juoksu-tta-a*), nominalized (*juoksu-tta-minen*) or adjectivized (*juoksu-tta-va, juoksu-tta-ma*). These causative stems as such cannot be used as Words, since they require inflectional markers (tense/aspect/case) or further suffixes (nominalization, adjectivization). But when the nominalization, adjectivization or verbalization is attached to the stem, no further derivation is possible. These level 2 suffixes are thus closing suffixes, in that they block further derivation (group 4 violations). Because level 1 stems often require further suffixes in order to be used as Words, I therefore make a distinction between level 1 affixes which turn Roots into Roots, and level 2 affixes which turn Roots into Words. Words are the exit points from the derivation. The model that will emerge as we proceed is illustrated in (8).

(8)



Examples (9a–h) give a general outlook of the model with some concrete examples.

(9) a. [Af_R + Af_R + ... + Af_R] + n/v/a + inflection + clitics

b. [*osta+tta-*] + *misen* + *han*
 [buy-CAU] + N.case + Cl
 ‘causing to buy’.

c. [*osta+tta-*] + *nen* + *ko*
 [buy-CAU] + v.Mood.1SG + Cl
 ‘whether I should cause somebody to buy’.

d. [*puu+isto+tu+tta-*] + *minen* + *han*
 [tree-col-cau-cau] + N.case + Cl.
 ‘the phenomenon of causing a collection of trees’.

e. [*puu+isto+tu+tta-*] + *va* + *mpi* + *han*
 [tree-col-cau-cau] + A + COMP + Cl.
 ‘the property of causing collections of trees’ (comparative).

f. [*pu(u)+isto+t(a)+in+ta*] + *a*
 [tree-col-cau-instr-cau] + v.1inf
 ‘to cause to have an instrument of causing a collection of trees’.

g. [*pu(u)+isto+ll*] + *inen* + *ko*
 [tree-COL-LL] + N.case + Cl
 ‘whether it is something that has a collection of trees’.

- h. [*pu(u)+isto+ll+sta-*] + A + *ko*
 [tree-COL-LL-CAU] + v.1inf + Cl
 ‘whether to cause something to have a collection of trees’.

Af_R refers to categorially underspecified Root affixes, which together form a Root stem marked by the brackets. After that comes inflection: first the category label, then category-specific inflection (case for nominals, transitivity/tense/aspect for verbs), then clitics.

I assume that the causative morpheme is a level 1 Root, based on the fact that it does not produce phonological words and that it requires separate morpheme pieces to be used as a noun, verb, or an adjective. This provides the diagnostic tool to detect other Roots under model (8). If the causative morpheme can be suffixed to a stem S, then this gives a sufficient condition for S being a level 1 Root. On the other hand, if the causative morpheme cannot be merged to S, then either (i) S is a level 2 word (**punainentaa*) or (ii) some independent constraint prevents the combination (**hammas-llinen*). A necessary property of conclusion (i) would be that S as such can be used as a phonological word. Hypothesis (ii) can be tested by using proper controls on the independent constraints, for example, by controlling the number of syllables in the word, its pragmatical plausibility, phonological output, and so forth.

It is well-known that many semantic attributes correlate with syntactic verbhood and nounhood. I will later discuss two such features in this study, eventiveness and referentiality (see Baker 2003: Ch. 3). Eventiveness refers to the property that the phenomenon represented by a word has a ‘temporal contour’ or that it evolves over time. Referentiality captures the property that the word refers to a complete thing, either abstract or concrete, in the world. The first grammarians circa 100 B.C noted that eventiveness is typical of verbs and referentiality of nouns (Robins 1989). This is the default view in functionalist literature (Givón 1984, Langacker 1986). However, it is not possible to correlate these notions one-to-one with the lexical categories. For example, there are eventive nouns, such as *juokseminen* ‘running,’ and non-eventive verbs (or verb phrases), such as *olla talo* ‘to be a house.’ I will return to this problem later. I therefore conclude that lexical categories cannot be *reduced* to features such as eventiveness or referentiality, although they play an important role in grammar. More specifically, there is evidence that they are semantic features of the Roots, unlike lexical categories, which reflect the syntactic contexts of Roots.

An anonymous referee has pointed out that it is not certain that the complex words listed in (7) are interpreted compositionally in the manner that I have decomposed them. The little experiment I conducted reveals that speakers of Finnish are quite good at decomposing the words semantically. Consider the word *juoksutin*. From the perspective of feature combination, this word can be decomposed into *juokse-u-ta-in* ‘run-N-CAU-INST,’ which thus involves elements of running, causation, instrument, and nounhood. When the word did not exist in their lexicon-in-use, the subjects gave precisely this meaning.¹⁰ Much of the same is true of more complex words, such as *hampaallistaminen* ‘the causing of somebody to have a collection of teeth’.¹¹ Finally, even in the case of fully lexicalized complex

¹⁰ The subjects’ interpretations were as follows (without glossing): juoksettamiseen käytettävä ainesosa, vrt. juoksute, miksei myös juoksumatto; jokin mekaaninen laite tai ihminen, joka juoksuuttaa tyyppettä; kyseessä on jokin juoksumaton tyyppinen laite, joka kuittaa ilkeillä sähköiskuilla tms. huonon harjoittelun; eikö ruuanvalmistuksessa käytetä juoksutinta, että saadaan aine juoksevaksi; juoksupyörä; Jonkinlainen houkutuslintu, esim. juoksukoiraradoilla. Tai sitten käänös tietotekniikkatermille ‘iterator’; eläimen juoksutin, esim. kilpakoiran juoksutin on jänis; Juoksuliina koiralle; Juustoa tehdessä tarvitaan juoksutin (naudoista saatava tai synteettinen), joka juoksettaa juustomassan eli hera erottuu; juuston valmistuksessa käytettävä ainesosa; Koiran/hevosen liikunta-avustin; laite jonka tehtävä on juoksuuttaa; Juustonjuoksutin ainakin on olemassa; Aine tai väline, jolla esim. maidon tai jonkun muun valkuaisperäisen aineen voi juoksuuttaa niin että siitä tulee rakenteeltaan epätasaista; Jonkinlainen kone jonka läpi neste virtaa eli juoksee. Tulee mieleen jonkinlainen tuotantovaihe meijerissä, varsinkin juustontuotannossa...; aine jolla voi juoksettaa toisen aineen; käsittääkseni juuston valmistuksessa käytetty aine, jolla maito saadaan juoksettumaan.

¹¹ The subjects’ interpretations were as follows: tehdä jollekin hammas (jolla ei aiemmin ole); lähes sama kuin hampaistottaa eli laittaa jollekin hampaat; hampaiden suuhunlaitto, esim. vanhuksia voidaan hampaallistaa valtion toimesta. Toisaalta substantiivina voidaan käyttää myös kuvaannollisessa merkityksessä, jolloin se tarkoittaa esim. henkilön mielipiteiden terävöittämistä. Vrt. ylioppilaskunnan vaalien ehdokkaan mielipiteitä hampaallistettiin; ei hampaistoiteta koko suuta kuten aiemmin, vaan vain pari kolme hammasta hampaallistetaan uudestaan. se on hampaallistamista; hampaat suuhun; Ehkä uusien hampaiden asentaminen hampaattomalle henkilölle; tehdä hampaalliseksi joku, jolla ei ole hampaita. voisin ehkä kuvitella sanottavan, että "tällä papalla on menossa hampaallistamisprosessi"; Asettaa hampaat johonkin (lähinnä koneeseen); Tehdä hampaita eli "hampaistottaa". Prosessin nimi on siis hampaallistaminen, ehkä; hampaisiin liittyvää toimintaa epäilemättä. Ehkä meidän kulttuurissa tätä sanastoa ei ole tosiaan hiottu tarpeeksi nyansoidulle tasolle?; Virkamies hammastaa; hampaiden lisääminen johonkin; Tapahtuma, jossa jollekin annetaan hampaat suuhun; tehdä hammasrattaan muotoiseksi; tehdä jotain sellaista jonka ansiosta jollekin kohteelle tulee hammas/hampaita; hampailla varustaminen, tuskin kuitenkaan

stems such as *puisto* ‘a park’ (lit. a collection of trees), subjects were sometimes spontaneously aware of the compositional reading, as in the following interpretation of *puistollistaa* ‘tree-COL-POSS-CAU-V’:

- (10) *alueen varustaminen puilla*
 ‘to supply the area with trees.’

This is because the compositional and idiomatic meanings can exist side-by-side. Although a larger and better controlled psychometric study should be considered for the future, I think that these facts speak in favor of the two-layer model: words are interpretable compositionally as long as we do not merge V affixes after the layer 2 closing suffixes (or as long as we do not violate some independent constraint).

4. Finnish word formation without lexical categories

4.1 General remarks

In this section, I will apply the category-free theory of word formation to Finnish by concentrating on certain (to me, at least) interesting features of Finnish word formation. This analysis relies on a few essential differences from the standard description of Finnish word formation. Most importantly, I do not assume that suffixation must create phonological Words. This leads to a more fine-grained and, I think, more simple analysis. Second, as we have much word formation which does not generate phonological Words, we can move lexical categories from derivational morphemes into syntax. This is in agreement with the kind of independent evidence briefly reviewed in section 2. Due to limitations of space, I have to leave many interesting issues untouched; what follows is rather a collection of what I take to be the most revealing aspects of Finnish word formation.

4.2 Nominalization and adjectivization

In this section I will look more closely at Finnish nominalization. In Finnish, any verb can be nominalized by suffixing it with *-minen*

ihmisestä, vaan vaikkapa hammaspyörästä tms. kuvaannollisesti myös ihmisestä, siis antaa aseet käteen jossain tilanteessa, jolloin joku on hampaaton, ei osaa puolustautua sanallista hyökkäystä vastaan.

(Hakulinen & Karlsson 1979: § 14.9, Koski 1978). This is part of a larger group of deverbal nominalizers, of which there are about 15 in Finnish, and it is one of the most productive; most of the others are conditioned by morphological and phonological properties of the stem they attach to (Hakulinen et al. 2004: § 222, 227). Interestingly, *-(i)nen*, which seems to be a part of *-minen*, is nevertheless a very common suffix that appears either alone as a nominal affix in a wide range of both nouns and adjectives, or in combination with other material whose status has been so far unclear. Some examples include *-(h)inen*, *-iainen*, *-i(m)mainen*, *-kainen*, *-kalainen*, *-kkainen*, *-(k)ko(i)nen*, *-lainen*, *-llinen*, *-lloinen*, *-mainen*, *-moinen*, *-nainen*, *-nkainen*, *-noinen*, *-rainen*, *-ttainen*, *-tuinen*, *-uainen*, *-uinen*, *-jainen*, and so forth (see Hakulinen 2000: § 3, Hakulinen et al. 2004: § 261–283, Karlsson 1983: 232–243). Many seemingly simple nouns end with *-nen*, such as *hevonen*, *ihminen*, *työläinen* ‘horse, human, worker.’ The suffix *-nen* can also be attached productively to a noun. In this case its meaning can be best described as diminutive. Thus, *kirja-nen* (‘book-let’) means ‘little book’. The distribution of *-i-* in the affix seems to be conditioned by morphological and phonological properties of the stem (Hakulinen et al. 2004: § 263).

This data raises a series of questions: Are all these affixes distinct atoms? Why is *-inen* so common in the Finnish lexicon? As a way into what I think is the correct answer, consider the suffix *-llinen* (11a) and a closely related variation (11b):

- (11) a. *hampa-(a)llinen*
 teeth-A
 ‘something that has teeth’
- b. *hampa-(a)llista-minen*¹²
 teeth-R-N
 ‘the property of causing to have teeth’

The suffix *-llista* in (11b) carries the same meaning ‘something that has –’ with (11a), plus an identical phonological shape *-ll-*.¹³ The nominal form contains *-(i)nen*, the V form causative *-sta-*:

¹² Other examples are: *teollistaminen* ‘industrialization,’ *kansallistaminen* ‘nationalization,’ *ennallistaminen* ‘restoration,’ *kaupallistaminen* ‘commercialization,’ *koneellistaminen* ‘mechanicalization.’

¹³ According to Rintala (1980a, b), the meaning associated with *-llinen* is heterogenous. She cited the following possible relations: ‘x which is y,’ ‘x which belongs to y,’ ‘x

- (12) a. *-ll-inen*
 ‘something that has –’ (nominal)
- b. *-ll-ista*
 ‘to cause something to have –’ (neutral: can be nominalized or verbalized)

It thus looks as if *-llinen* contains actually two components, *-ll-* (or *-lli-*), expressing the meaning ‘something that has–,’ which can be nominalized and verbalized, plus the nominalization or adjectivization *-(i)nen*. Furthermore, *-(s)ta* is a neutral morpheme piece that can be verbalized and nominalized.¹⁴ Yet there is one problem in dissolving *-llinen* and *-llista* into two morphemes: *-ll-* is an affix that cannot alone produce a word. Thus, according to a traditional word-based theory, it would be a problem to separate *-ll-* from *-llinen*. Assuming that word formation is based on morphemes, not words, provides a solution. Suppose that *-ll(i)-* is a categorially indeterminate morpheme piece with its own meaning ‘something that has –’. Because it is a categorially indeterminate morpheme, it does not, by itself, constitute a Word. Because it is not a Word, it cannot be pronounced alone. In this way, it is possible in principle to divide many of the *-(i)nen* forms listed above into two or more morpheme pieces, and thereby simplify Finnish word formation. For instance, there is only little reason to assume that *-lliste*, *-llistin*, *-llisto*, *-llista*, *llistum*, and so forth, are unrelated atomistic affixes, and no need to stipulate that some of them are terminating stratum 2 affixes.¹⁵

To test this hypothesis, we can try to combine *-ll-* with other Root affixes according to (9a). Here I consider some of the affixes present in (7). Example (13a) shows how to combine other Root affixes to *-ll-* and (13b)

which is in y’s possession,’ ‘x which is like y,’ ‘x where y is,’ ‘x which produces y,’ among others.

¹⁴ The affix *-sta-* is composed from two elements, *-s-* and causative *-ta-*. This is related to the fact that *-inen* is replaced with *-s-* in certain contexts; see below. I thank an anonymous reviewer for pointing this out.

¹⁵ When the *ll-* material is separated from the suffixes, we get more fine-grained analysis, e.g. *-ll-is-ta-e* poss-is-cau-res = *-lliste*. Other *-(i)nen* nominalization affixes that are, according to Hakulinen (2000), similarly formed from two affixes are the following: *-kainen* (*-kka-(i)nen*, *-nkainen* (*-nka-(i)nen*), *-noinen* (pronominal adverb *-noin-(i)nen*)); likewise affixes such as *-rainen*, *-ttainen*, *-uinen*, *-tuinen*. The analysis is here likewise simplified if *-(i)nen* is separated from the other material. Space limitation prevents me from going into the details of these affixes and analyses; they have to be studied one-by-one. What matters for present concerns that the category-free theory leads potentially into a more fine-grained and, I think, more simple analysis.

shows the vice versa. Word forms which are phonologically impossible in Finnish, creating ungrammatical CCC combinations /nll/ and /sll/ (Karlsson 1983: § 4.2.2.2), are of course impossible.

- (13) a. hampaa-lli-ste (työlliste, lailliste) hampaa-lli-stin (>paika-lli-stin), hampaa-lli-sto, hampaa-lli-staminen (>kansallistaminen, kaupallistaminen), hampaa-lli-sta-ma (>ehdo-lli-sta-ma, lai-lli-sta-ma), hampaa-lli-stus (>enna-lli-stus, havainno-lli-stus), hampaa-lli-sti, hampaa-lli-staa (>paikallistaa, kohtuullistaa, kansallistaa, kaupallistaa, alueellistaa), hampaa-lli-stella, hampaa-lli-stua (>osallistaa, kaupallistaa, pinnallistaa, koneellistaa, liikunnallistaa, ammatillistaa, taiteellistaa)
- b. hammas-ste-lla (>arvostella, perustella, varmistella, kummastella, kauhistella, ujustella), *hammas-tin-lla, hamma-sto-ll-inen (>yliopistollinen, laivastollinen, osastollinen, linnustollinen), ??hamma-sta-ll-inen, *hammast-us-ll-inen, hamma-st-e-ll-inen (>ennusteellinen), ??hamma-st-u-ll-inen.¹⁶

But what kind of affix is *-inen*? Aside from the diminutive meaning in some cases, it has no clear meaning in itself, and it is connected only to nominal forms (including adjectives, which are discussed below). The category neutral theory provides a straightforward answer: *-(i)nen* is one exponent of the syntactic marker for nounhood (Marantz 1997) and adjectivehood.

The same reasoning can be applied to monomorphemic words which end with *-(i)nen*. Many seemingly monomorphemic nouns are affixed with *-(i)nen*, among them *hevonen* ‘horse’ and *ihminen* ‘human.’ Separating *-(i)nen* from these words creates stems *ihmi-* and *hevo-*, which cannot be used as Words. This is not a problem for the morpheme-based theory, since these forms are Roots: morphemes which, by themselves, cannot form Words. We can avoid repetition in the Finnish lexicon by assuming that *-(i)nen* is the exponent for *n* (see Koskeniemi 1983: § 1.8, for the same

¹⁶ The words *hammastallinen* and *hammastullinen* are impossible, because *-ll-* only attaches to N affixes; I return to categorial selection in section 4.3. Note that if (13) is the correct analysis of words such as *perustella*, *perustele-n*, then the verbal reflexive morpheme *-ele-*, discussed in section 4.3, is composed out of *arvo-s-ta-e-ll-a* ‘value-S-CAU-E-LL’ and the marker *-a* for the infinitival verb. Similarly, we have *tule-o-s-ta-e-lla* ‘come-n-s-cau-e-ll-inf’. If we produce a separate Word for each of these phases, we get *tule* ‘come!’, *tulo* ‘coming’, *tulos* ‘the result of coming, i.e., a result’, *tulos-taa* ‘to print’, *tulos-te* ‘a printout’, *tulos-tee-ll-inen* ‘something that has printouts’ and *tulos-tell-a* ‘to do printing in a specific, casual manner’. It also follows that *perustella* is related to *perustee-ll-inen*. Needless to say, this hypothesis is controversial.

argument). For instance, the morphological analysis of *ihminen* is $\sqrt{ihmi}+n$, *hevonen* = $\sqrt{hevo}+n$. But then phonological material such as *ihmi-* and *hevo* cannot constitute Words, hence there is little reason to say that they belong to any of the lexical categories. In other words, they are Roots.

There is morphosyntactic evidence that bare nouns ending with *-(i)nen* should be treated similarly to complex nouns having the *-(i)nen* ending. To begin with, all nominals ending with *-(i)nen* have exceptional case-marking properties: *-(i)nen* is turned into *-s(e) + case*. This is not the typical case-marking pattern for Finnish nominals (Nelson 1998). The following data illustrates the difference with the words *talo* ‘house’ and *hevonen* ‘horse’:

CASE	bare noun	<i>(i)nen</i> -noun
NOMINATIVE	talo	hevo-nen
ACCUSATIVE	talo-n	hevo-se-n
PARTITIVE	talo-a	hevo-s-ta
INESSIVE	talo-ssa	hevo-se-ssa

Table 2. Case marking for two kinds of Finnish nominals, bare nominals and *(i)nen* nominals.

Table 2 shows that the case-marking properties of *-inen* nouns and bare nouns are different. The original *-nen* is replaced with *-s-* or *-se-*, and then the regular case affix rules are applied.¹⁷ If so, lexicalized nominals ending with *-(i)nen* must be marked overtly for *n*, since the same rule applies to these cases as well, as shown in (14). More generally, this suggests that there is some syntactic structure inside of *hevonen* comparable to *juokseminen*, namely, Marantz’s (1997) *n*(ominalization) head.

One alternative hypothesis would be that the rule is phonological, applying to all words ending with *-(i)nen*. There is evidence that this is not the case. Note, first, that complex nominals ending with *-(i)nen* cannot occur as the first constituent in a compound, but nouns containing covert *n* can. Thus *liitu#taulu* ‘blackboard’ is an acceptable compound in Finnish, but **liitu+mainen#taulu*, **työ+läinen#mies* are not. Instead of a compound, a word boundary must be used here (*liitumainen taulu*). The

¹⁷ Affixes *-s-* and *-se-* are layer 1 affixes, since they occur in causatives as well. Thus, this is different affix than *-(i)nen* (See Hakulinen 1979: 124–5). The *-e-* material seems to be there to prevent outputs such as **hevosn*.

same is true of lexicalized forms ending with *-(i)nen*, thus **hevonenmies* ‘horseman’ is not a possible word in Finnish. Suffix *-nen* must be replaced with *-s* generating *hevosmies*. The same truncation occurs in the case of any suffix; thus, a causative of *hevonen* is *hevo-s-taa* not **hevonentaa*. Barring for now the analysis of the emerging *-s-*, the general rule is to ban further derivation after overt *n*, as *n* behaves like a closing affix. This explanation is in line with the category-free model. This rule then automatically explains the behavior of lexicalized *-(i)nen* nominals if it is assumed that they contain the syntactic constituent *n*. Now consider the word *onnen* ‘of luck.’ This genitive word can be used as a first member of a compound, as in *onnenpäivät* ‘lucky days.’ The word ends with *-(i)nen*, yet a compound is possible. If the rules regulating *-(i)nen* nominals were phonological, then *onnenpäivät* should not be a possible word. There is thus evidence that Finnish has an overt marker for a functional head *n*. There are at least two alternative hypotheses. According to the first hypothesis, the distribution of *-(i)nen* throughout the Finnish lexicon is a matter of diachronic facts only. This hypothesis makes it hard to explain why the suffix is still in productive use. But note that in some cases the *-inen* could well be frozen inside of a larger suffix. Another problem with this hypothesis is how the exceptional properties of *-(i)nen* words are to be explained.¹⁸ Another hypothesis, to my mind much more plausible, is to claim that many nominals end with *-(i)nen* because nominals have “prototypical *-(i)nen* instances” which are used when new words are coined. Here *-(i)nen* would not be its own morpheme piece, but part of a phonological template which is used to coin new words, perhaps by relying on “analogy”. This leaves unexplained why *-(i)nen* could nevertheless be separated from the stem, leaving semantical and formal material (‘morphemes’) that can enter into other contexts as well. Furthermore, there is direct evidence (case suffixation etc.) that the behavior of *(i)nen* nominals cannot be accounted for in terms of phonological rules only. I conclude that *-(i)nen* seems to be a morpheme piece on its own. If this is so, there is a reason to believe in the existence of Roots, since what is left of many morphemes when *-(i)nen* has

¹⁸ Finally, if the distribution of *-(i)nen* would be a matter of diachrony, being completely lost in modern Finnish as a separate morpheme, it would not remove the explanatory burden: at least *-(i)nen* had been some kind of morpheme. If so, what kind? Why does it have such a distribution?

been separated does not, in itself, create a Word but a unpronounceable Root.¹⁹

The affix *-(i)nen* can be used to derive adjectives as well; indeed many *-(i)nen* forms are adjectives. There are two possible hypotheses with respect to this data. First, one could say that *-(i)nen* is homonymous between two possible underlying syntactic representations, being the exponent of *n*(ominals) or *a*(djectives). Another hypothesis says that adjectives are derived from nouns in Finnish. The latter hypothesis is more interesting, given the cross-linguistic observation that adjectives tend to cluster either with nouns or verbs in languages around the world (Dixon & Aikhenvald 2004), suggesting that adjectives can be either ‘noun-like’ or ‘verb-like’. Not surprisingly, Finnish adjectives are noun-like: they inflect like nouns, they require the supporting copula in the position of a predicate, and like nouns, they take genitive modifiers (*valtava-n iso* ‘extremely-GEN big-A’). This pattern becomes intelligible at once if we regard adjectives as being constructed out of *n*. Yet apart from their noun-like properties, adjectives clearly have some properties that nouns do not have. For instance, they occur in the modifier position inside of the NP in their bare forms, whereas nouns occur in genitive forms (14a–b), but note that both nouns and adjectives are capable of occurring in the argument position (14b–c). Adjectives and nouns differ in some of the syntactic contexts they may appear (e.g., 14d–e).

- (14) a. *hevo-sen satula* (nominal)
 horse-N.GEN saddle
 ‘a/the saddle of a horse’.
- b. *puna-inen satula* (adjective)
 red-N.NOM saddle
 ‘a/the red saddle’.
- c. *puna-inen on väri* (adjective)
 red-N.NOM is color
 ‘red is a color’.

¹⁹ As I said before, all this leaves room for the possibility that *-(i)nen* is lexicalized in *some* cases. From the fact that the causative morpheme *-tta* is frozen inside of a word such as *roko-ttaa* ‘to vaccinate, lit. to cause to have pox’ we cannot conclude that it is not in productive use at all.

d. *hyvin punainen*
 very red

e. **hyvin satula*
 very saddle

The difference between adjectives and nouns in Finnish is grammatically very subtle. There is nonetheless a quite dramatic *semantic* difference between words which occur in a modifier and head position inside of an NP. The modifier attributes a general property to the head, whereas the head, perhaps together with an overt or covert determiner, is capable of picking a concrete or abstract referent from the world (Baker 2003). The difference between adjective-like *punainen* ‘red’ and noun-like *punainen* ‘redness’ is thus that in the nominal use, it refers to the redness as a whole concept/property (i.e., that the redness as such is a color), whereas in the attributive use it attributes one particular shade or one particular piece of redness to the saddle (Salo 2003: 69–73). We can then say that an *nP* bears a “referential index,” to use Baker’s (2003) terminology. This has the semantic consequence that it refers to a whole entity in the world (and in turn explains how it interacts with quantifiers, determiners, binding theory, number, and the rest of it) and cannot occur in the modifier position without undergoing the grammatical alteration from NOM to GEN. Other morphemes have the same effect, for example, the *-ll-* affix discussed above. The fact that those *nPs* which occur in the modifier position in their zero-derived forms possess comparative forms can be explained on the same semantic basis, since comparatives produce gibberish when combined with words possessing a referential index (*??talompi*, ‘house-COMP’). This hypothesis, which I assume tentatively here, is in agreement with Baker’s claim that adjectives are characterized by the lack of referential capacity belonging to nouns. In a language where adjectives behave in a verb-like manner, they are then created from *vPs* and thus they bear many verbal properties, such as verbal inflection and a lack of copula in predicative construction (Baker 2003: 249–263). I will summarize these ideas in the next section, after first discussing verbs and eventiveness.

4.3 Verbal affixes

In this section I will look at the properties of “verbal affixes” (as traditionally understood) of Finnish derivational morphology and argue that

they are best described as category neutral Root suffixes which are *eventive* in their semantics. Furthermore, this section presents more detailed data concerning verbal affixes, causatives included, which I discussed in a more preliminary sense in section 2.

According to the model presented here, eventive ‘verbal affixes’ (V affixes from here on) in Finnish morphology would belong to category neutral affixes as they can be verbalized, nominalized and causativized. Recall from section 2 that causatives were seen as a core example of this phenomenon. What sets these ‘verbal’ affixes apart from the nominal ones is that they are semantically *eventive*. Nominals are anchored to the spatio-temporal word by means of spatially bounded ‘things’; verbs are anchored by means of events (e.g., Rijkhoff 2002). But, as I will argue presently, eventiveness does not imply syntactic verbhood. In fact, the verbal category label +V is completely redundant in the case of Finnish V affixes.

To begin with, Finnish has four types of simple V affixes: causatives CAU ‘to cause,’ reflexives REF ‘to do by oneself,’ frequentives/continuatives FRE ‘to do continuously, frequently’ and momentives MOM ‘to do momentarily, fast’ (Hakulinen et al. 2004: § 303, Karlsson 1983: 237ff., Wiik 1975). These semantically aspectual morphemes are derived from the repertoire of the universal grammar, as they are in no way limited to Finnish (Cinque 1999). Causatives and reflexives are valency changing morphemes, and frequentives and momentives express the manner of performing the action. Several morphological forms in Finnish can correspond to one or another of these eventive semantic attributes, which means that the classification is not based on form but follows current tradition by using semantic classification. The allomorphy of FRE includes at least *skele, ele, ile*; MOM *ahta, aise, alta*; CAU *ista, sta, ta, tta*; REF *istu, itu, stu, tu, u, utu, Vntu*; see Karlsson (1983: 276). This allomorphy seems to simplify under the category-free theory, as these endings dissolve further (see § 4.2 and fn. 16), but I will put that aside for now. Importantly, each of these affixes carries an eventive or aspectual meaning. There are other nonproductive affixes that can belong to these groups, but which are not addressed here. Combinatorial possibilities of these affixes are shown in Table 3 (adopted from Karlsson 1983) with concrete examples in (15).

1	2	3
CAU	CAU	CAU
REF	REF	FRE
FRE	FRE	
MOM		

Table 3. Combinatorial possibilities of Finnish V affixes according to Karlsson (1983).

- (15) *tee-tä-ttä-ä*, *laula-tt-ele*, *lanno-itt-u*, *juo-vu-tta*,
do-CAU-CAU, sing-CAU-FRE, fertilizer-CAU-REF, drink-REF-CAU,
‘to cause to cause to do’, ‘to cause to sing frequently’, ‘to be fertilized’, ‘to get drunk’,
vaiht-u-ile, *heitt-el-yttä*, *suoj-el-tu*
change-REF-FRE, throw-REF- CAU, protect-FRE-REF
‘to change oneself frequently’, ‘to cause to throw frequently’, ‘to become protected frequently’.

The table shows only three positions, but as many as five of these affixes can be combined together to form such sequences as (*luetuttelutella* ($\sqrt{\text{READ-CAU-CAU-FRE-CAU-FRE}}$), *luettelutella* ($\sqrt{\text{READ-CAU-FRE-CAU-FRE}}$), *lueskeluttelutella* ($\sqrt{\text{READ-FRE-CAU-FRE-CAU-FRE}}$) and *lukaisuttelutella* ($\sqrt{\text{READ-MOM-CAU-FRE-CAU-FRE}}$) (see Ojanen & Uotila–Archelli 1979, Karlsson 1983). Although these forms are quite unusable, they are possible words in Finnish.

Category neutral V affixes combine rather freely, to a point where it looks clear that the process is recursive and turns a stem into another, more complex stem. Much like the causative morpheme, I would like to suggest that all V affixes are part of the categorially neutral level 1 derivation. This is because (i) they cannot be pronounced alone as words, and because (ii) all stems formed by using any of these affixes can be again verbalized, adjectivized and nominalized. In short, they sit in the category of level 1 Roots.

Verbal affixes CAU, MOM, FRE and REF, in the traditional sense, are called “verbal” because when merged with appropriate verbal inflectional elements they possess properties of verbs: the ability to transitive, to bear tense/aspect/mood, and so forth. Without these inflectional elements they are verbal in the sense that they are *eventive* in their meaning. However, there is evidence that semantic eventiveness and the syntactic category of verbs should be separated from each other. Many nominals are also

eventive and have a ‘process reading’ (Alexiadou 1999, Fu, Roeper & Borer 2001, Grimshaw 1990, Vendler 1967, 1968). Nominalized verbal stems can behave in this way in Finnish. For instance, *juokseminen* ‘running’ or *juoksu* ‘a run’ are more eventive in their meaning than *talo* ‘house’ or *Pekka* (proper name). The former allows modification by aspectual adverbs much better than the noneventive nominals:

- (16) a. *isä-n juokse-minen nopeasti hämmensi muitakilpailijoita*
 father-GEN run-N.NOM fast confused other competitors
 ‘Father’s running fast confused other competitors-’
- b. **isä-n talo nopeasti oli punainen*
 father-GEN house fast was red

In Finnish, deverbal adjectives have the same property; they can be combined with eventive adverbs, as shown below:

- (17) a. *nopeast juokse-va isä*
 fast run-A father
 ‘a/the fast running father’
- b. *?*nopeasti puna-inen talo*
 fast red-A house
 ‘a/the fast red house’

Finally, as pointed out by Peter Bosch (p.c.), certain verbal contexts probe eventive readings of simple nouns:

- (18) a. *Pekka began the book*
 ‘Pekka began to read/write the book’
- b. *Pekka began the tunnel*
 ‘Pekka began to drive/build/walk the tunnel’

Nouns *book* and *tunnel* can be seen as referring to an event, not a concrete thing. Thus, eventiveness is not the same thing as the syntactic category of verbs. Furthermore, if we follow Marantz (1997) and assume that verbhood is associated with words only in syntax, i.e., in a component which generated phrases, we need to look at verb phrases as well. Here we find examples of noneventive verb phrases, which are generated from non-eventive words:

- (19) a. ?**Pekka on ihminen nopeasti*
 Pekka is human fast
- b. *Pekka juoksee nopeasti*
 Pekka run fast

The VP *on ihminen* ‘be a human’ functions in the role of the predicate, exactly as the verb *run* ‘to run’. Depending on the theory, the copula can be taken as an inflectional element which realizes verbal inflection (Salo 2003: 247–256) or the phonological exponent of a functional head signaling predicatehood (see Baker 2003). I will thus assume that verbal affixes, so called in the traditional theory, are categorially underdetermined, but marked with a feature [+eventive] (or with a more complex eventive structure, if necessary).

Consider the fact that Finnish nominalization affixes typically select for ‘verbal affixes,’ so that there are very rarely two consecutive N affixes (Karlsson 1983: 241).²⁰ We do not need to assume that they select words which are syntactically verbs; rather, these affixes are either closing affixes, or they select words which are marked as being eventive. In this way, selectional restrictions that were originally explained by relying upon an ultimately redundant syntactic category can be explained in a functionalist fashion by using a feature [\pm eventive]. A word-internal verbal category is thus redundant; rather, it obscures the fact that syntactic verbhood and eventiveness do not always occur in tandem.

In sum, Finnish verbal affixes are not yet syntactically verbal: they cannot be used as verbs without certain further inflectional markers. Moreover, they can be used as nouns as well. They are verbal in the sense that they are semantically eventive, yet many nominals and adjectives are also eventive in their meaning, namely, when they contain an eventive Root. The following table lists all possible feature combinations of *nP*–*vP* and referential–eventive in Finnish, as discussed so far.

²⁰ In Finnish and likewise in English (Fabb 1988), few nominalizers can apply to already nominalized forms. One of them is *-lainen* (Rintala 1972: 4.3.2). The affixes *-la* (‘place of-’), *-tar* (‘female member of -’) and *-kas* (semantics more obscure) can be affixed to certain nominalized forms.

NOMINALS	
nP , +EVENTIVE+REFERENTIAL	<i>juokse-minen</i>
nP , -EVENTIVE-REFERENTIAL	<i>talon -</i> , <i>Apunainen -</i>
nP , -EVENTIVE+REFERENTIAL	<i>talo</i>
nP , +EVENTIVE-REFERENTIAL	<i>Ajuokse-va-</i>
VERBS	
vP , +EVENTIVE+REFERENTIAL	–
vP , -EVENTIVE-REFERENTIAL	<i>on talo</i>
vP , -EVENTIVE+REFERENTIAL	–
vP , +EVENTIVE-REFERENTIAL	<i>juost-a</i>

Table 4. Various combinations of the two grammatical dimensions discussed here: syntactic category (nP – vP) and semantic features eventive and referential. The syntactic category reflects the position of the Root in the syntax, nP corresponding to the position of arguments (or logical subjects), vP corresponding to the position of predicates. Semantic features \pm eventive (i.e., the existence of ‘temporal contour’), and \pm referential (whether the element refers to a complete entity) are features of the Roots. Adjectives are generated from nPs in Finnish.

The distinction between nP – vP corresponds to the syntactic position of the element (argument–predicate), whereas features \pm eventive and \pm referential are properties of the Roots. Intuitively, lexical category is determined from the context of the word, semantic features from within.

4.4 Morphophonology, roots and words

The categorially underspecified theory of the lexicon postulates the existence of abstract, completely neutralized roots that cannot be pronounced alone. In the case of complex verbs, there is no special difficulty in postulating such elements. In the case of nouns, the situation is more complex. For instance, the word *vesi* ‘water’ is derived from an abstract Root $\sqrt{\text{WATER}}$, for which there does not exist any kind of vocabulary entry. This means that $\sqrt{\text{WATER}}$ is a completely neutralized item: the minimal pronounceable word is the nominal, nominative singular form *vesi*. One could argue that the postulation of completely neutralized items is too risky and perhaps too speculative.

However, there is evidence for the existence of the abstract Root. Consider example (22), which illustrates complex words derived from the

Root $\sqrt{\text{WATER}}$ (the $d-t$ alteration is predicable by a general consonant gradation rule):

- (20) *vede-ssä*, *vete-en*, *vede-ttä*, *vede-llä*, *vede-n*
 water-INE, water-ILL, water-ABE, water-ADE, water-ACC
 ‘in the water, to the water, without water, at water, water’

This is an instance of well-known stem alteration in Finnish. Affixation is based on the word *vede-*, which is never pronounced as such, whereas the nominative singular form is *vesi*. Thus we have to regard *vede-* as a Root. Accordingly, *vesi* is generated by a suppletion-like process for nominative singular nominals. This model predicts that all syntactic derivations operating with a nominative nP should attach to the exceptional nominal form *vesi*, not *vede-*. This prediction is borne out in the case of clitics, which indeed are adjoined to the nominal singular form *vesi+hän* (**vedehän*). Clitics are clearest examples of morphemes which are not part of the derivational morphology, as they are suffixed to phonological words even after inflectional markers. Furthermore, Root affixes trigger the regular form:

- (21) a. *vesi* ‘water’
vede-ll-inen, *vede-ttää*, *vede-stää*, *vede-stö*, *vede-tär*
 water-LL-INEN, water-CAU, water-CAU, water-COL, water-TAR
 ‘something containing water, to cause to contain water, to cause to contain water, a collection of waters, a female related to water, such as a mermaid.’
- b. *susi* ‘wolf’
sude-ll-inen, *sude-ttaa*, *sude-staa*, *sude-sto*, *sude-tar*
 wolf-LL-INEN, wolf-cau, wolf-cau, wolf-col, wolf-tar
 ‘something containing a wolf, to cause to have/be a wolf, to cause to have/be a wolf, a collection of wolfs, a female wolf.’
- c. *käsi* ‘hand’
käde-ll-inen, *käde-ttää*, *käde-stää*, *käde-stö*, *käde-tär*
 hand-LL-INEN, hand-CAU, hand-CAU, hand-COL, hand-TAR
 ‘something having a hand, to cause to have a hand, to cause to have a hand, a collection of hands, a female related to a hand (such as a female having a hand).’

Root affixes thus select the Root form since they are merged to the Root before n , whereas merging of the n with features [+nominative, singular]

triggers suppletion.²¹ Kiparsky (2004) notes that stems in Finnish must end with a vowel before affixation, whereas this constraint does not apply to Words. Thus we have alterations such as:

- (22) a. *tuhanne-ssa tuhat*
 thousand-INE thousand-SG.NOM
 ‘in thousand, thousand’.
- b. *avaimessa avain*
 key-INE key-SG.NOM
 ‘in the key, a/the key’.
- c. *miehenessä mies*
 man-INE man-SG.NOM
 ‘in the man, a/the man’.

The word form on the right occurs only in the nominative and partitive singular form, so that morphologically the regular form ending with *-e* occurs at the Root level, whereas the exceptional form is supplied at the Word level. This makes sense if we identify Stem morphology with Roots, and assume that the nominative singular is provided by suppletion, as argued in the case of *vesi-vede*. Again, clitics and other closing affixes attach to the nominal singular forms which suggest that these forms are Words, thus categorially nominals:

- (23) a. *tuhat-han, avain-pa, mies-kö*
 thousand-CL, key-CL, man-CL
 ‘thousand, (it is) a key, is it a man?’
- b. *tuhat-mainen, avain-mainen, mies-mäinen*
 thousand-MAINEN, key-MAINEN man-MAINEN
 ‘a thousand-like, a key-like, a man-like’

We can thus explain part of the curious stem alterations in Finnish word formation by making a distinction between Roots and Words: certain stems are phonological exponents of Roots, others are phonological exponents of Words. What applies to verbs applies to many nouns, namely, there is a component inside of the phonological word that does not make a word itself.

²¹ Furthermore, *vesi-* is used for the exponent of the Root plural forms. Thus *vesi-ä, käsi-ä* ‘water-PRT, hand-PRT’ means ‘(many) waters’ or ‘many hands’, respectively.

Cross-linguistically, the idea that word formation can be dissolved into levels or strata is not anything new (Allen 1978, Kiparsky 1982, Seigel 1974; for a recent review, see McMahon 2000: 1–53). First, stratum 1 affixes (e.g. English *-y*, *-al*, *-ic*, *-ize*, *-ous*, *-ive*, *-ese*) constitute a prosodic word, whereas stratum 2 affixes do not have similar effects in that they are not part of the prosodic word they attach to (*-ness*, *-less*, *-ful*, *-hood*, *-ship*, *-ly*, *-ish*, *-dom*). Among the concrete differences between stratum 1 and stratum 2 affixes are the following:

Stratum 1 affixes can alter the syllabic and stress patterns of the stems, whereas stratum 2 affixes do not have this property.

Stratum 1 affixes are often not productive morphologically and semantically, whereas stratum 2 affixes are.

Stratum 1 affixes cannot be attached outside of stratum 2 affixes (the so-called Affix Ordering Generalization), while stratum 1 affixes are the only affixes which can be attached to bound morphemes.

Although this model itself is not uncontroversial, it is possible that the traditional distinction between stratum 1 and stratum 2 affixes coincides with the distinction between Roots and Words. Giegerich (1999), drawing on Selkirk (1982), provides such a theory and a number of arguments in favor of it.²² The Finnish *-minen* affixation is extremely productive, nearly unrestricted, whereas reflexivization, frequentivization, and likewise many N affixes are productive but not unrestricted (Karlsson 1983). Similarly, stratum 1 affixes are fused with their hosts more tightly than stratum 2 affixes (Hakulinen et al. 2004: § 159):

- (24) a. *tule+o* = *tulo* (stratum 1)
 tule+minen = *tuleminen* (stratum 2)
- b. *juoksu+tta+in* = *juoksutin* (stratum 1)
 juokse+minen = *juokseminen* (stratum 2)

²² Giegerich describes his model as follows: “Under this model, the inputs to all stratum-1 affixation are members of the category Root. Given the multiple stratum-1 affixation is possible [...] the outputs of stratum-1 affixation must again be members of the category Root. Root must therefore be a recursive category: unlike in the traditional understanding of the term [...] roots may be morphologically complex as long as such complexity is the result of the stratum-1 morphology” (73). In other words, R-affixation applies to Roots and creates new Roots as assumed here. The same idea seems to apply well to Finnish.

- c. puu+isto = puisto (stratum 1)
puu+mainen = puumainen (stratum 2)
- d. perus+ta+e+ll+inen = perusteellinen (stratum 1)
perusta+minen = perustaminen

The Root-Word hypothesis can explain some otherwise stipulated properties of the two stratum model. Siegel's (1974) Affix Ordering Generalization, which states that stratum 1 affixes are applied before stratum 2 affixes, now follows automatically: R affixes apply by default to Roots which are categorially underspecified. Stratum 1 affixes are morphologically more unproductive and semantically more non-compositional than stratum 2 affixes, which can perhaps be explained on the basis of their different functional roles: whereas level 1 affixes are more relevant to the semantics of the complex predicate, level 2 affixes encode parts of speech and hence syntactic context, either alone or in addition to some semantic function. Productivity is a necessary feature of the latter for the same reason as inflection is, whereas productivity is compromised in the case of the former. Also, one major weakness of the two-stratum model has always been the fact that the postulation of the strata has been an isolated property, not derivable from independently motivated general principles. The present model can avoid this problem, since the distinction between the two strata is principled on an independent basis, as it can be derived from the Root-Word distinction.

5. Conclusions

There is at least some evidence for the hypothesis that linguistic theory cannot do without studying category neutral Roots, along with nouns, verbs and adjectives. Somewhere in the depths of grammar there is a rule which combines category features +N, +V and +A with Roots that do not bear such features. The standard theory says that this combination is relevant neither for word formation nor for syntax. I argued that, on the contrary, it is potentially important for both. While many specific aspects of the model presented here are controversial, and many issues had to be left out due to limitations of space, I still feel that the more fine-grained analysis along these lines could provide solutions to some of the problems concerning Finnish word formation.

Because category neutral lexical Roots are almost completely devoid of syntactic information, they are best understood as “concepts” in the psychological sense, thus some kind of mental symbols which are individuated based on their semantics alone. For example, the lexical Root $\sqrt{\text{OSTA}}$ - ‘buy’ is a lexical item which is “related to or perhaps identical with the concepts that are the elements of the ‘cognoscitive powers,’ sometimes now regarded as a ‘language of thought’” (Chomsky 2005: 19). It obtains syntactic properties when the concept is used to construct a linguistic expression referring to buying. When buying is the logical subject of the proposition, it obtains the properties of nouns (*osta-minen* ‘buy-ing’), and when it is used in the role of a logical predicate, it obtains properties of verbs (*osta-a* ‘to buy’). When it is used attributively in a non-predicate position, it becomes an adjective (*osta-va*). Lexical Roots thus serve as the interface between psychology and linguistics: they are both conceptual, being associated with meaning but not syntactic features, and linguistic, serving as the basic building blocks of linguistic expressions.

References

- Adger, David (2003) *Core syntax: A minimalist approach*. Oxford: Oxford University Press
- Alexiadou, Artemis (1999) *On the syntax of nominalization and possession*. Doctoral dissertation, University of Potsdam.
- (2001) *Functional Structure in Nominals. Nominalization and Ergativity*. Amsterdam: John Benjamins.
- Allen, Margaret (1978) *Morphological Investigations*. Dissertation, University of Connecticut.
- Aronoff, Mark (1976) *Word formation in generative grammar*. Cambridge, MA.: MIT Press.
- Baayen, Harald (1993) On frequency, transparency and productivity. In Geert Booij & Jaap van Marle (eds.) *Yearbook of Morphology 1992*, pp. 181–208. Dordrecht: Kluwer.
- Baker, Mark (2003) *Lexical categories: Verbs, Nouns, and Adjectives*. Cambridge, MA.: MIT Press.
- Barner, David & Bale, Allen (2002) No nouns, no verbs: psycholinguistic arguments in favor of lexical underspecification. *Lingua* 111: 771–791
- Berman, Ruth (1999) Children’s innovative verbs versus nouns: structured elicitations and spontaneous coinages. In Lise Menn & Nan Bernstein Ratner (eds.) *Methods for Studying Language Production*, pp. 69–93. New Jersey: Erlbaum.
- Caramazza, Alfonso & Shapiro, Kevin (2004) Language categories in the brain: Evidence from Aphasia. In Adriana Belletti (ed.) *Structures and Beyond: The*

- cartography of syntactic structures*, Volume 3, pp. 15–30. New York: Oxford University Press.
- Chomsky, Noam (1970) Remarks on Nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, pp. 184–221. Waltham, Mass.: Ginn & Co.
- (2004) Beyond Explanatory Adequacy. In Adriana Belletti (ed.) *Structures and Beyond: The cartography of syntactic structures, Volume 3*, pp. 104–131. New York: Oxford University Press.
- (2005) Three Factors in Language Design. *Linguistic Inquiry* 36: 1–22.
- Cinque, Guglielmo (1999) *Adverbs and Functional Heads. A Cross-linguistic perspective*. New York: Oxford University Press.
- Dixon, Robert M. V. & Aikhenvald, Alexandra (eds.) (2004) *Adjective classes: A cross-linguistic typology*. New York: Oxford University Press.
- Fabb, Nigel (1988) English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory* 6: 527–539.
- Farrell, Patrick (2001) Functional shift as a category underspecification. *English Language and Linguistics* 5: 109–130.
- Fodor, Jerry (2003) *Hume Variations*. Oxford: Oxford University Press.
- Fu, Jingqi, Roeper, Thomas & Borer, Hagit (2001) The VP within process nominals: evidence from adverbs and the VP anaphor *Do-so*. *Natural Language & Linguistic Theory* 19: 549–582.
- Giegerich, Heinz J. (1999) *Lexical Strata in English. Morphological Causes, Phonological Effects*. Cambridge: Cambridge University Press.
- Gil, David (1995) Parts of speech in Tagalog. In M. Alvez (ed.) *Papers from the Third Annual Meeting of the Southeastern Asian Linguistics Society*, pp. 67–90. Tempe, AZ: Arizona State University Program for Southeast Asian Studies.
- Givón, Tom (1984) *Syntax: a functional-typological introduction*. Amsterdam: John Benjamins.
- Grimshaw, Jane (1990) *Argument Structure*. Cambridge, MA.: MIT Press.
- Heim, Irene & Kratzer, Angelika (1998) *Semantics in Generative Grammar*. Malden & Oxford: Blackwell.
- Karlsson, Fred (1983) *Suomen kielen äänne- ja muotorakenne*. Juva: WSOY.
- Kiparsky, Paul (1982) From Cyclic Phonology to Lexical Phonology. In Harry van der Hulst & Norval Smith (eds.) *The structure of phonological representations*. Vol 1. Dordrecht: Foris, 131–175.
- (2004) Finnish Noun Inflection. In Diane Nelson & Satu Manninen (eds.) *Generative Approaches to Finnic and Saami Linguistics*, pp. 109–162. Stanford: CSLI Publications.
- Koskeniemi, Kimmo (1983) *Two-level Morphology: A general computational model of word-form recognition and production*. Helsinki: Suomen kielen laitos, Helsingin yliopisto.
- Koski, Mauno (1978) Suomen sananjohdon perustyypit. *Papers from the conference on general linguistics*, Suomen kielitieteellisen yhdistyksen julkaisuja, pp. 103–117.
- Hale, Ken & Keyser, Samuel J. (2002) *Prolegomenon to a theory of argument structure*. Cambridge, MA.: MIT Press

- Hakulinen, Auli, Vilkuna, Maria, Korhonen, Riitta, Koivisto, Vesa, Heinonen Tarja-Riitta & Alho, Irja (2004) *Iso Suomen kielioppi*. Helsinki: SKS.
- Hakulinen, Auli & Karlsson, Fred (1979) *Nykysuomen lauseoppia*. Helsinki: SKS.
- Hakulinen, Lauri (2000). *Suomen kielen rakenne ja kehitys* (5th edition). Vantaa: Helsingin yliopiston suomen kielen laitos.
- Harley, Heidi & Noyer, Robert (1999) Licensing in the non-lexicalist lexicon: nominalization, roots and the encyclopedia. *MIT Working Papers in Linguistics* 32: 119–137.
- Hengeveld, Kees, Rijkhoff, Jan & Siewierska, Anna (2004) Parts-of-speech systems and word order. *Journal of Linguistics* 40: 527–570.
- Jeanne, LaVerne M. & Hale, Ken (1998) Transitivity in Hopi. In *Memorias del Quinto Encuentro Internacional de Lingüística en el noroeste*. Tomo 1. Hermosillo: Universidad de Sonora.
- Langacker, Ronald W. (1987) *Foundations of cognitive grammar*. Stanford, California: Stanford University Press.
- Logan, Gordon D. (1988) Towards an instance theory of automatization. *Psychological Review* 95: 492–527.
- Marantz, Alec (1997) No escape from syntax: Don't try morphological analysis in the privacy of your lexicon. In Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams (eds.), *Proceedings of the 21st Annual Penn Linguistic Colloquium*, pp. 201–225. (Pennsylvania Working Papers in Linguistics 4.2). Philadelphia: Penn Linguistics Club.
- (1999) Creating verbs above and below little v. Manuscript, MIT.
- (2000) Reconstructing the lexical domain with a single generative engine. Unpublished ms., MIT.
- Mc Laughlin, Fiona (2004) Is there an adjective class in Wolof? Robert Malcolm Ward Dixon & Alexandra Aikhenvald (eds.) (2004). *Adjective classes: A cross-linguistic typology*, pp. 242–261. New York: Oxford University Press.
- McMahon, April (2000) *Lexical phonology and the history of English*. Cambridge: Cambridge University Press.
- Mosel, Ulrike & Hovdhaugen, Even (1992) *Samoan reference grammar*. Oslo: Universitetsforlaget AS.
- Nelson, Diane C. (1998) *Grammatical Case Assignment in Finnish*. London: Garland.
- Ojanen, Jussi & Uotila-Arcelli, Eeva (1977) Suomen kielen frekventatiivi- ja momentaaniverbien derivaatio. *Virittäjä* 81: 245–264.
- Pensalfini, Robert (1997) *Jingulu grammar, dictionary and texts*. Dissertation, MIT.
- Pesetsky, David (1995) *Zero Syntax*. Cambridge, MA.: MIT Press.
- Pesetsky, David & Torrego, Esther (2004) Tense, Case, and the Nature of Syntactic Categories. In Jacqueline Guéron & Jacqueline Lecarme (eds.) *The Syntax of Time*, pp. 495–537. Cambridge, MA.: MIT Press.
- Pitkin, Harvey (1984) *Wintu grammar*. Berkeley: University of California Press.
- Rijkhoff, Jan (2002) *The Noun Phrase*. Norfolk: Oxford University Press.
- Robins, Robert H. (1989) *A Short History of Linguistics*. London: Longman.
- Saariluoma, Pertti (1995) *Chess Players' Thinking*. London: Routledge.

- Salo, Pauli (2003) *Causatives and the empty lexicon: A minimalist perspective*. PhD Dissertation, University of Helsinki.
- Sasse, Hans-Jürgen (1993) Das Nomen – eine universale Kategorie? *Sprachtypologie und Universalienforschung* 3: 187–221.
- Scalise, Sergio (1998) Inflection and derivation. *Linguistics* 26: 561–581.
- Selkirk, Elisabeth O. (1982) *The Syntax of Words*. Cambridge, MA.: MIT Press.
- Siegel, Dorothy (1974) *Topics in English morphology*, dissertation, MIT.
- Sulkala, Helena & Karjalainen, Merja (1992) *Finnish*. London: Routledge.
- Rintala, Päivi (1976) Suomen nominikantaiset *mainen*-johdokset. *Virittäjä* 80: 163–91.
- (1980a) Adjektiivisten *llinen*-johdosten semantiikkaa ja syntaksia. *Virittäjä* 84: 212–29.
- (1980b) Adjektiivisten *llinen*-johdosten semantiikkaa ja syntaksia. *Virittäjä* 84: 302–18.
- Vendler, Zeno (1967) *Linguistics in Philosophy*. New York: Cornell University Press.
- (1968) *Adjectives and Nominalizations*. The Hague: Mouton.
- Vogel, Petra M. & Comrie, Bernard (eds.) (2000) *Approaches to the Typology of Word Classes*. Berlin, New York: Mouton.
- Whitman, Neal. (2004) *Category Neutrality: A Type-Logical Investigation*. London: Routledge.
- Wiik, Kalevi (1975) Suomen frekventiivi ja kontinuatiivi. *Virittäjä* 79, 153–166.
- Williams, Marianne (1976) *A grammar of Tuscarora*. New York: Garland.

Contact information:

Pauli Brattico
Cognitive Science Unit/Department of Psychology
Siltavuorenpenger 20 C
00014 University of Helsinki
e-mail: pauli.brattico@helsinki.fi