

THE FUNCTIONAL-LEXEMATIC MODEL OF WORD-FORMATION¹



Francisco J. Cortés Rodríguez
University of La Laguna

This paper provides a description of an autonomous Word Formation Component to be integrated in Dik's (1989) Functional Grammar. The integration of this component implies a separation between word formation and predicate formation in the context of this grammar. Word formation is regarded as both a process of lexicalization of analytical structures and as a grammaticalization of the lexicon as well. The outline of the model was proposed in Martín Mingorance (1985), but this paper presents a fuller description and adds a new perspective by integrating a cognitive dimension: word formation must also be regarded as an act of linguistic categorization.

0. INTRODUCTION

In Dik's Functional Grammar (FG) (Dik 1978, 1989; Siewierska 1991), like other grammatical models such as Lexical Functional Grammar (Bresnan 1982) Government & Binding Theory (Chomsky 1981, 1982, 1986) or Categorical Grammar (Reichl 1982), the lexicon occupies a central role as the initial component for the generation of messages. In fact, linguistic expressions are the result of projecting or expanding lexical information. Apart from a store for basic lexical units (or *predicates*, as they are labelled in the model) there is a predicate formation component that accounts for derived predicates.

This predicate formation component includes rules that would produce, among other types of constructions, derived words. But predicate formation rules are not an adequate instrument to explain word formation: the bivalent nature of word formation (as a lexical phenomenon that is based on the combination of linguistic signs and as a hinge between lexicon and grammar) and its relation to all levels of grammatical description would point up the need for the establishment of an autonomous word formation component (WFC) within, in this case, Functional Grammar.

The first section of this paper is a description of the standard treatment of lexical creation in FG and an explanation of its inadequacies; the second section explains the theoretical foundations of the WFC; the third section is a detailed description of the analytical methodology for the study of derived lexical units; the results of this study constitute the two initial subcomponents of the WFC: the affix lexicon, where all information on the restrictions for word-formation patterns are codified, and the set of underlying formulae that correspond to the functional and semantic meanings of those patterns; the fourth section provides a new perspective on the model by integrating a third cognitive axis of description; this is achieved by considering underlying formulae as the formal representation of

¹ This paper is based on a project funded by the Spanish Government's Bureau for Universities and Research (DGICYT), grant no. PB94-0437. This grant is hereby gratefully acknowledged.

cognitive schemata and, consequently, viewing the production of a new lexeme as an instantiation of an act of categorization by the speaker; the explanatory potential of the WFC is thus increased: section five articulates briefly its potential to account for non-canonical creations with a given procedure and also for the diachrony of word-formation patterns.

1. PREDICATE FORMATION RULES

In FG word-formation is located in the Predicate Formation Component, and its effects are the result of applying predicate formation rules. The tasks of these rules is to perform «mappings of predicate frames onto (derived) predicate frames» (Dik 1989, 55). Consequently, any process that would change some feature of a predicate frame (category of the predicate, quantitative or qualitative valency, change in the properties of the SoA or meaning) should be explained by means of these rules. For instance, deadjectival causative verb formation in Hungarian is explained by the following predicate formation rule (De Groot 1987, 21):

DEADJECTIVAL PREDICATE FORMATION IN HUNGARIAN

input: [-dyn] pred_A ([-con]x₁)_ø

output 1: [+dyn] pred-D_V ([-con], [± tel] x₁)_{Proc}

D = -ul/-ül

meaning: “the property expressed by pred_A is presented as coming about through a process”.

output 2: [+dyn] pred-D_V ([+con] x₂)_{Ag} ([±tel] x₁)_{Go}

D = -it

meaning: “x₂ brings it about that the property expressed by pred_A applies to x₁”

E.g.- rövid “short” - rövidül “shorten” - rövidít “shorten”

In this context, word formation doesn't occupy a separate position, but rather it appears intermingled with other kinds of processes, such as «inchoativization» (which relates the two constructions of the verb OPEN: *Jane opened the door* → *the door opened*), de-transitivization (compare *John was drinking a glass of milk* → *John drinks*), periphrastic causativization (f.i. Spanish *Juan hizo abrir la puerta a su hermano*, which is the result of a valency increase rule on the predicate *abrir*), or comparative degree formation for adjectives.

There are two important drawbacks to predicate formation rules as a mechanism for producing complex lexical units: First, these devices are too powerful for an adequate explanation of the nature of word formation processes. The function of predicate formation rules is to establish the systematic relations between two types of predicate frames; that is, between two structural configurations. But, to put the processes of word formation on a par with other processes that express such structural systematic connections overstates the regular nature of lexical creation. In this sense, it can be said that FG's approach is comparable to the transformational view of Generative Semantics or Case Grammar, where there is no separate treatment for word formation: complex and simple words, and even sentential expressions, are generated from a common underlying structure, the only difference being the surface lexical expression of the abstract nodes of such a structure. In the context of FG, predicate formation rules are very similar to these lexical generative transformational rules. And one type of these rules, say valency reduction, with the general format:

VALENCY REDUCTION

input: $\text{pred}_X(x_1) \dots (x_n)$ output: $\text{pred}_{X/Y}(x_1) \dots (x_{n-1})$

would apply to the following phenomena:

1. INCHOATIVIZATION (without affixation, English)
input: $\text{pred}_v(x_1)_{\text{Ag}}(x_2)_{\text{Go}}$
output: $\text{pred}_v(x_2)_{\text{Pro}}$
meaning: "the predicate pred_v is relevant only to (x_2) "
Ex.- *John closes the door* → *the door closes*
2. INCHOATIVIZATION (with affixation, Hungarian. De Groot 1987, 16-17 «Intransitive Predicate Formation»)
input: $\text{pred}_v([\pm\text{con}] x_1)_{\text{Ag}}([\pm\text{tel}] x_2)_{\text{Go}}$
output: $\text{pred-R}_v([\text{-con } \pm\text{tel}] x_2)_{\text{Pro}}$
R= -ódik/-ödik
meaning: "the predicate pred_v is relevant only to (x_2) "
Ex.- *János zárja az ajtó-t* → *az ajtó zár-ódik* (**János által*)
(John close the door-acc → the door close-R John by)
3. DETRANSITIVIZATION (GOAL SUPPRESSION, English. Adapted from Mackenzie 1986, 22)
input: $\text{pred}_v(x_1)_{\text{Ag}}(x_2)_{\text{Go}}$
output: $\text{pred}_v(x_1)_{\text{Ag}}$
meaning: The relation expressed by pred_v applies to Ag
Ex.- *My horse wins all races* → *My horse always wins*
4. NOMINALIZATION (VERBNOUN FORMATION. English, adapted from Mackenzie 1986, 14)
input: $\text{pred}_v(x_1)_{\text{Ag}}(x_2)_{\text{Go}}$
output: $\text{pred-ing}_{v/n}(x_2)_{\text{Go}}$
meaning: "the SoA designated by the Input is presented as applying to the Go of that SoA"
Ex.- (*My horse's*) *winning the race*

Many of the arguments given against the treatment of derivational morphology in a transformational approach apply equally to the use of predicate formation rules for these processes: One essential feature of word formation is its restricted productivity on all levels. Whereas productivity in syntax, and even in inflection, is quite a straightforward notion, productivity in derivational morphology is the result of the interplay of a number of factors which concerns, among other things, the restrictions that must be imposed on the input structures. In other words, it is not enough to define an input predicate frame to ensure the generation of all and only the adequate complex units². In my opinion, Predicate

² For a more detailed discussion on the inadequacies of Predicate Formation Rules to account for word formation phenomena see Cortés Rodríguez, 1996.

Formation is not the adequate *locus* for derivational morphology, since, on the one hand, it tends to identify word formation with other processes of a more regular nature and connected with what could be considered a «syntacticist» view (as far as predicate frames can be considered the equivalent of analytical constituent structures). This results in a loss of adequacy, as no attention is paid to many aspects that restrict the generation of morphologically complex lexical units. .

Pertinent to this discussion is a recent distinction that has been established in FG and which helps us to distinguish between what pertains to predicate formation and what pertains to word formation. Throughout this paper I have been using the terms «predicate» and «lexical unit» with different meanings. One of the factors that has produced this confusion/masking of word formation with predicate formation processes is precisely the conflation of the concepts «lexeme» and «predicate» which, as Hengeveld (1992, 51) points out, is a conflation of lexical and syntactic units. This distinction has several consequences. The most important for us is that the syntactic unit, the predicate, (not the lexical unit), is defined in functional syntactic terms. The following example and its subsequent explanation by the author illustrates this:

5. (Mandarin Chinese; From Hengeveld 1992, 62)

Tā zài màn-màn-de pǎo.

3.SG DUR slow-RDP-NR run

«S/He is running slowly»

The constituent màn-màn-de «slowly» is a manner adverbial but not a manner adverb. It is derived from the probably verbal predicate man «slow», which is reduplicated and nominalized, the resulting construction meaning something like «slow manner».... Although the word màn-màn-de «slowly» forms part of the predicate phrase, and, being optional, must be a modifier within that predicate phrase, it is not an adverb but a nominalized verb.

As we can observe, Predicate Formation creates derived predicates, but this does not mean that they are derived lexical units. The fact that one word can appear in different syntactic environments does not imply a change in lexical category. Word formation, on the other hand, is a lexical phenomenon, concerned with the creation of new lexical units. This difference can also be applied to other terms, most outstandingly to words such as «nominalization» and «adjectivalization»: we can talk now of lexical nominalization, which encompasses all the processes for the creation of nouns, as opposed to syntactic or functional nominalization, which will account for the creation of nominal predicates (as e.g. the subject predicate in the Spanish construction *el amar es todo un arte* = the to-love_{Subj} is all an art).

2. THE WORD FORMATION COMPONENT

With the two domains of Predicate Formation and Word Formation now differentiated, I will describe, in the remainder of this paper, the structure of the Word Formation Component that Martin Mingorance (1985) proposed for FG, and will apply it to the description of an affixal unit:

The Word Formation Component is conceived as autonomous, running parallel to the grammatical component since it consists of the same subcomponents. The reason for this separation lies in the specific nature of word formation, which as the title of Kastovsky's (1977) paper explicitly expresses, lies «at the crossroads of morphology, syntax, semantics and the lexicon.» I have already indicated deficiencies of a view of word formation as

a regular grammatical process. Further problems would arise if word formation were included in one specific component. The initial hypothesis for the study of word formation is to consider it as a process involving two perspectives: both a grammaticalization of the lexicon (as in Coseriu 1978) and a lexicalization of predicational structures. This concept of «lexicalization» encompasses the perspective adopted by the generative tradition that started with Lees (1960) and continued with the Generative Semanticists and the works of Brekle (1976² [1970¹]) and Pottier (1974), to be taken up again -after the lexicist reaction- by Fabb (1984), Sproat (1985, 1988) and Shibatani and Kageyama (1988), among others.

In general terms, this conception starts from the fact that complex lexical units show the same behaviour as analytical constructions with regard to different grammatical phenomena. From a productive point of view, this *lexicalization* is understood as a process of successive reduction from an analytical underlying structure to the complex lexical unit. In this way, the derived word becomes the morphosyntactic expression of a (series of) grammatical relation(s); in other words, the complex lexical unit is a «syntagm» based on a functional syntactico-semantic relation between a determinant and a determinatum (see Marchand 1969² [1960¹],3). Despite this, studies in the lexicist tradition in the seventies and eighties³ have proved that it is not enough to consider derived words as the result of syntactic operations. The productivity and predictability of word formation rules make it impossible to accommodate word formation within syntax. At this time word formation was located in the lexicon, which was conceived as a repository for anything «idiosyncratic» or impossible to systematize. In fact, the models devised in this period were concerned with the discovery of the conditions that characterize word formation processes, conditions which differentiate them from the more regular processes of syntax and which, at the same time, brought to light semi-regularities that separate them from the primary lexicon.

These two approaches are not incompatible: word formation must be understood as a systematic process of lexicalization of analytical structures subject to a set of particular conditions which pertain to all the levels or strata of a grammatical model: phonology, morphology, syntax, semantics and pragmatics. Therefore, a word formation component cannot be integrated in any one of these levels, but must be seen as autonomous though parallel to the grammar (as in Coseriu 1978, 256-257). Consequently, apart from being a process of lexicalization -understood in the terms explained above- word formation is a lexicological phenomenon by which speakers create labels for new conceptual categories of their universe of discourse taking as basic material already existing lexical units (lexemes and affixes); that is, word formation is also a grammaticalization of the lexicon. The conception of word formation as this double and complementary process is the starting hypothesis in Martín Mingorance (1985) for the articulation of a Word Formation Component (WFC) in the FG model.

3. METHODOLOGY

Given its janus-like nature, the study of word formation consists of two phases: analytic and synthetic. In this paper I will be mainly concerned with the analytic phase, which is methodologically prior to the other. The analytic phase is composed of the following steps:

³ A tradition initiated by Chomsky's (1970) attack to Generative Semantics' treatment of nominalization, and continued —as regards word formation— by Halle (1973), Aronoff (1979² [1976¹]), Hammond (1980), Roeper and Siegel (1978), Szymanek (1980, 1985), Scalise (1987), etcétera.

- (1) Delimitation of the set of types of complex lexical units. Ideally, this should be done following Coseriu's (1962) conditions of syntopic, synstratic and synphasic levels for both languages.
- (2) Analysis of the structure of the complex lexical units. This analysis will examine, at least, their phonological, morphological, syntactic and lexico-semantic structure, including the relational component (that is, the specific type of semantic relation that holds between the elements coded as determinans and determinatum of the complex lexical unit) since, as Coseriu (1978) states, the meaning of a complex lexical unit consists of: (a) the sum of the meanings of its components, which is predictable by the application of the word formation rules (*Be-deutung*) plus (b) the relation of the linguistic signs with the referents from the extralinguistic reality (or *Bezeichnung*). Motivation, coinage and usage conditions should, ideally, be expressed too.
- (3) Construction of the whole set of fully specified predications (as understood in FG) underlying each type of complex lexical unit.
- (4) Finally, by inductive generalization, and taking into account the existence of such «lexical gaps», the whole set of basic predicational schemata, underlying the complex lexical units of each language, should be specified.

All the information obtained in the analytic phase is encoded in the two subcomponents of the Base of the Word Formation component; these subcomponents are:

- (A) A word formation lexicon, where all affixes have a lexical entry, with a full specification of its morpho-phonological and lexico-semantic structure, its combination rules and restrictions, and the types of predications underlying the complex lexical units.
- (B) The whole set of basic predicational schemata underlying all the types of complex lexical units.

An example of the outcome of the analytical phase, the different parts of a lexical entry for the English affix *-ician* are provided below (for a view of the whole entry, see appendix):

– The first section of the entry is devoted to graphemic information, in which a description of spelling variants is expressed. In the case of *-ician*, there is a basic form «-ician» with a «c», though a form with a «t»: «-itian» also occurs, but is not subject to any specific or systematic condition:

(A) *SPELLING ALTERNANTS*

– ICIAN

– ITIAN. Asystematic and idiosyncratic in *dietitian/dietician*.

There are cases that are subject to more systematic and predictable conditions; for examples the Spanish suffix *-nte*, where the stem vowel for nouns derived from verbs of the third conjugation is, with some exceptions, the diphthong [je], graphically expressed as «ie» (Ex.- *hiriente, mordiente*); however, when the base stem ends in a vowel, the spelling for this diphthong is «ye»: *influyente, reconstituyente*.

Another example concerns the subsidiary spelling of the agentive suffix *-er*: «-or», «-Or» seems to be preferred to the most frequent spelling «-er» when the base verb ends in *-ate* or when the derived unit belongs to a technical register (as in *demonstrator, incubator, generator, acceptor*). All these variants can be described in the following way:

SPELLING ALTERNANTS

-ER

-AR. *Features*: Idiosyncratic. Unconditioned. E.g.- *liar, burglar, pedlar*

Non-productive

-OR. *Features*: Restricted productivity, to:

- a) Verbal inputs of latinize sustratum. E.g.- *inspector, supervisor, actor, survivor, editor, conqueror*.
- a.1) Verbal formations in -ate as inputs. E.g.- *demonstrator, incubator, generator*.
- b) Register: technical. E.g.- *acceptor* (vs.unmarked *accepter*)
- b.1) Outputs are items of legal word-stock. E.g.- *vendor, donor*.
- c) Latinate words not fully naturalized
- d) Otherwise, idiosyncratic.

– In section (B) the phonological description of the affix is given. It includes the segmental and suprasegmental features of the affix, and the allomorphic variants. A description of the type of boundary attached to every affix is also given. I adopt the distinction, given originally by Siegel (1979) and maintained in generative phonology, between two types of affixal boundaries. The prephonological boundary «+» is attached to affixes that have an influence on the phonological and syllabic structure of the derived unit, as in the case of *-ician*: this implies a shift of word stress to the penultimate syllable of the derived word, and also a redistribution of the syllabic structure if compared to the base:

(B) PHONOLOGICAL DESCRIPTION

Segmental form [i] ∂n]*Allomorphs* Vacuous*Affixal Type* +X##_{Suffix}*Stress Location*[¹] ∂n] / [¹S₁ S₁]_{Base} _____ → [³S₁ ¹S₁ /] ∂n /₁]E.g.- {¹∂} {dξ/k} → {³∂} {¹dξ/}, {¹mju:} {¹s/k} → {³mju:} {¹s/}[¹] ∂n] / [_{1-n} ¹S₁ S₁]_{Base} _____ → [(_{0-n}) ²S₁ ³S₁ ¹S₁ /] ∂n]E.g.- {∂} {¹dξ/s} {¹t/ks} ({²∂} {³dξ/s} {¹t/})*Syllable Distribution* Disyllabic suffix. Suffix initial [l] → heterosyllabic

(CODES: S = syllable

{ } = syllable boundaries)

The affix *-er*, on the other hand, would have a postphonological boundary: #, since it is neutral with regard to stress assignment and syllabic division; therefore the description in its entry would be:

PHONOLOGICAL DESCRIPTION

Segmental Form [∂(r)]*Allomorphs* [o(r)] *Iff*: a) Spelling = -or
and b) Legal terms*Affixal Type* #x##_{Suffix}*Stress Location* [²∂(r)] / [¹S₀]_{Base} _____[³∂(r)] / [²S₀ ¹S₀]_{Base} _____*Syllable Distribution* Tautosyllabic

– In section (C) there is a description of the features of the lexical units that act as bases in the derived word. These features constitute the set of Input Conditions. There are the following types:

– Phonological: in the case of *-ician*, there are no special restrictions of this type. A typical example of phonological conditioning is the suffix *-en* which admits attachment only to bases ending in the cluster «vowel+(sonorant)+plosive» (frighten, dampen).

– Morphological: some affixes show preference for bases with a specific morphological structure. In the case of *-ician*, as can be seen, it is attached to bases ending with the morphemes *-ic* or *-ics*. In the corpus selected⁴ there are only two exceptions: *beautician* (formed from the word *beauty*) and *dietician* (formed from the word *diet*). In section 5 such anomalous formations will be used to explain certain theoretical issues concerning the dynamics of word formation patterns.

– Morpholexical conditions, which specify the lexical strata to which the base lexical units belong.

– Categorial and lexico-semantic information on the base units: This is expressed in the form of the frame types associated with the input lexical units. The specification of selection restrictions for the arguments of those frames is done in terms of Aarts & Calbert's (1979) High and Low Primary Features. Such features are equivalent to the concepts of *classemes* and *semes* from Lexical Semantics, and are useful to signal the lexical areas to which the base units belong. The entry for *-ician* makes clear that the bases for this suffix belong primarily to the area of adjectives and nouns that refer to sciences and technical skills, with a larger frequency of terms from the disciplines of philosophy, logic and medicine:

(C) INPUT CONDITIONS

Phonological Vacuous

Morphological Base Final Cluster: [l(k)s)] . Except: *beautician*, *dietician*.

Substratum Latinate technical terms

Frame Types

NOMINAL PREDICATES

TYPE 1

$\phi_n(x_1: NP <-Con -Perc -Attr \varepsilon sciences/crafts (:: philosophy/logic/medicine)> (x_1))_\emptyset$

E.g.- magician, phonetician, mathematician, obstetrician

TYPE 2

$\phi_n(x_1: NP <-Con -Perc +Attr +Ev \varepsilon aesthetics> (x_1))_\emptyset$

E.g.- beautician, cosmetician, dietician

ADJECTIVAL PREDICATES

TYPE 1

$\phi_{adj}(x_1: Adj P <+Sta -Phy \varepsilon sciences/crafts (:: medicine)> (x_1))_{Ads}$

E.g.-technician, clinician

⁴ The corpus has been extracted from the Collins COBUILD English Language Dictionary, 1987. London & Glasgow: Collins, and from Marchand (1969² [1960¹]: 297-298).

– Section (D) of the lexical entry for affixes shows the morpho-lexical operations applied to the base lexical units. The format of this operation, represented as a rule, provides the bracketed structure of the morpho-lexical constituents of the derived lexemes.

(D) MORPHOLEXICAL OPERATIONS (Immediate Constituents Structure)

$\phi_n \rightarrow [[\phi_n]_{\text{Base}} + \text{ician}]_{\text{Suf}_n}$ Denominal nouns

$\phi_{\text{adj}} \rightarrow [[(\text{adj})_{\text{Base}} + \text{ician}]_{\text{Suf}_n}$ Deadjectival nouns

– Section (E) is a specification of the output conditions. These conditions express the morphophonological readjustments that may have to be made after the morpho-lexical operation has been carried out. Following Aronoff's (1979² [1976¹]) model, there are two major types of readjustments, truncations and allomorphic variations:

(E) OUTPUT CONDITIONS

Truncations

(1) $[[\text{___} \{ \mathbf{Ck} \}]_{\text{Base}} / \int \partial n]_n \rightarrow [\text{___} \{ \mathbf{C} \} \{ \int \partial n \}]_n$
mu • si • cian, *a • rith • me • ti • cian*, *phy • si • cian*

(2) $[[\text{___} \{ \mathbf{Cks} \}]_{\text{Base}} / \int \partial n]_n \rightarrow [\text{___} \{ \mathbf{C} \} \{ \int \partial n \}]_n$
logisti • cian, *mathemati • cian*

(3) $[[\text{___} \{ \mathbf{C} \}]_{\text{Base}} / \int \partial n]_n \rightarrow [\text{___} \{ \mathbf{C} \} \{ \int \partial n \}]_n$
beauti • cian

Allomorphies

(1) $[\{ \{ S_{1-n} \} \{ ^1\text{-æ-} \} \text{___} \}]_{\text{Base}} \rightarrow [\{ \{ S_{1-n} \} \{ ^3\text{-ð-} \} \{ ^1\mathbf{C} \} \{ \int \partial n \} \}]_n$
 $[^2\text{pi:} \cdot \text{d} / \cdot \text{tæt} \cdot \text{r/ks}] \rightarrow [^2\text{pi:d} / \partial \text{tr} / \int \partial n]$

(2) $[\{ \{ S_{1-n} \} \} \{ \{ \mathbf{CV} \} \{ ^1\text{-e-} \} \text{___} \}]_{\text{Base}} \rightarrow [\{ \{ S_{1-n} \} \} \{ ^2\mathbf{CV} \} \{ ^3\text{-/} \} \{ ^1\mathbf{C} \} \{ \int \partial n \} \}]_n$
 $[f \partial \text{net/ks}] \rightarrow [^2f \partial \text{u}^3 \text{n} / \text{t} / \int \partial n]$
 $[^2\Theta / \partial \text{ret/ks}] \rightarrow [^2\Theta / \partial \text{r} / \text{t} / \int \partial n]$

(3) $[\text{___} \{ ^3\partial \{ S_1 \} \{ S_1 \} \}]_{\text{Base}} \rightarrow [\text{___} \{ ^2\text{eu} / ^2\text{æ} \} \{ ^3S_1 \} \{ ^1\mathbf{C} \} \{ \int \partial n \}]_n$
 $[f \partial \text{net/ks}] \rightarrow [^2f \partial \text{u}^3 \text{n} / \text{t} / \int \partial n]$
 $[^3\text{st} \partial \text{t} / \text{st/ks}] \rightarrow [^2\text{stæ}^3 \text{t} / \text{s} / \text{t} / \int \partial n]$

- The last part of the affixal entry includes:

(a) the type of derived unit; this is expressed, in line with FG notation, with the format of a predicate frame:

Output Frame Types

$\phi_n (x_j; \text{NP} \langle +\text{Hum} \rangle (x_j))_a$ E.g.- *technician*

(b) the derivational schemata that underlie all the lexical units derived by means of this affix. This schemata constitutes the functional and semantic representation of a word formation pattern. Such representations are obtained from the analysis of the functional structure and meanings of the derived units by means of a method of lexical decomposition. I will illustrate this procedure in a simplified manner to specify the formula underlying the Spanish agentive noun *recaudador*:

(1) Starting from the following definition: *recaudador* («tax collector»), «*persona que se dedica a recaudar impuestos*» («someone whose job is to collect taxes») the first step

would be to convert the definiton components (above underlined) into argument, predicate and satellite variables (according to the FG notation):

$$\text{recaudador}_n^i: [(\underline{x_1}: <+\text{Human}> (x_1))_{\text{Agent}} (x_2: [e: \text{RECAUDAR } (x_1)]_{\text{Agent}} (x_3: \text{impuestos } (x_3))_{\text{Goal}} (e)]_{\text{Action}} (\sigma_1: \text{“repeatedly”} (\sigma_1))_{\text{Manner/Frequency}} (\sigma_2: \text{“as an occupation”} (\sigma_2))_{\text{Manner}}$$

The literal reading of the formula would be: «*recaudador* is a human entity (=x₁) that carries out the action (=e) of collecting taxes (=x₃) repeatedly (=σ₁) and as an occupation (=σ₂)». Notice that the superscript *i* signals the entity of the formula towards which the nominalization is directed.

(2) Martín Mingorance (1985, 44) signals the necessity of reconstructing the predicate that indicates the kind of semantic relation that holds between the components of a derived or composed word in which there is no overt verbal nexus. In this case, the predicate to be introduced in the formula will be the verb «HACER» («perform»), by which the relation of agentivity that exists between the first argument (x₁) (variable position that later on in the process of generating the word would be lexically saturated by the suffix *-or*) and the action expressed as a Goal in the second argument (lexically saturated by the base word *recaudar*) is explicitly expressed:

$$\text{recaudador}_n^i: [\text{HACER } (x_1: <+\text{Human}> (x_1))_{\text{Agent}} (x_2: [e: \text{RECAUDAR } (x_1)]_{\text{Agent}} (x_3: \text{impuestos } (x_3))_{\text{Goal}} (e)]_{\text{Action}} (\sigma_1: \text{“repeatedly”} (\sigma_1))_{\text{Manner/Frequency}} (\sigma_2: \text{“as an occupation”} (\sigma_2))_{\text{Manner}}$$

(3) By comparison with the underlying structures of other derived units of a similar type (f.i. *deshollinador*, *editor*, *expendedor*, *pintor*, etc) it is possible to deduce a generalized formula by substituting the lexical variables ((s) for constants:

$$\phi_n^i: [\text{PERFORM } (x_1: <+\text{Human}> (x_1))_{\text{Agent}} (x_2: [e: \phi_v \dots (e)]_{\text{Action}} (\sigma_1: \text{“repeatedly”} (\sigma_1))_{\text{Manner/Frequency}} (\sigma_2: \text{“as an occupation”} (\sigma_2))_{\text{Manner}}$$

(4) The next step will be to assign the appropriate pragmatic functions to some of the constituents; these functions specify which elements are selected for lexical saturation due to their communicative/pragmatic salience. The pragmatic functions used in this formula would be the following: (i) *Theme*, which specifies the entity that is categorized in the universe of discourse where this formula is relevant; therefore the theme would be the derived word itself and the whole formula would be a restrictive specification on that lexical unit; (ii) The *Topic* function, assigned to the agent argument, which specifies the entity about which information is given and which later will become the determinatum of the «syntagma» (cf. Marchand 1969² [1960¹]) that corresponds to the derived word; (iii) The *Focus* function is assigned to the variable for base words, given that they signal the *differentia specifica* (and, therefore, convey modifications in the information of the addressee) of the new unit with regard to other agentive formations; that is, the base word is the determinant of the corresponding syntagma:

$$\phi_{n\text{THEME}}^i: [\text{PERFORM } (x_1: <+\text{Human}> (x_1))_{\text{Agent}} \text{TOPIC } (x_2: [e: \phi_{v\text{FOCUS}} \dots (e)]_{\text{Action}} (\sigma_1: \text{“repeatedly”} (\sigma_1))_{\text{Manner/Frequency}} (\sigma_2: \text{“as an occupation”} (\sigma_2))_{\text{Manner}}$$

The elaboration of the formulae underlying complex words is at the same time a process of delimitation of the *Bedeutung* (functional meaning) and the «inventory meanings», or, in other words, fixations of the derived word's system meaning at the norm level (cf. Laca 1993, 183) given that the designational aspect (which cannot be motivated nor is functionally relevant in the system or the norm) must be discarded. A second aspect of the-

se formulae, once they are fully specified by the assignment of pragmatic functions, is that they are also the grammatical and discourse representation of the complex words: semantic and pragmatic function assignment enables us to predict which elements will be lexically saturated (or, in other words, endowed with lexical material) for the generation of the derived units; these functions signal the elements that become communicatively relevant and, therefore, they will constitute the syntagma that binds the lexical elements of a derived lexeme in terms of a functional relation of determination. The morpho-syntagmatic structure corresponding to this general formula would be:

$$\phi_{\text{Theme}}^i: \text{-OR}_{\text{AgentTopic}=\text{Determinatum}} + \phi_{\text{Focus}=\text{Determinant}}$$

Notice that this permits comparisons with other units derived from the same formula in which function assignment varies; compare, for instance, *recaudador* with *recaudación*, where topic assignment is on the predicate; or English *payer* vs. *payee* (patient topicalization) and *payment* (predicate topicalization)

4. COGNITIVE SCHEMATA

As can be seen, the configuration of the fully specified underlying formulae is a reflection of the concept of word formation as a grammar of the lexicon; the fact that they are at the same time the starting point for the generation of complex units via successive reduction from more analytical structures make them also an adequate representation of the processes of word formation as a lexicalization of syntagmatic constructions. These were the two theoretical standpoints of this model, as was mentioned before. But there is a third and new perspective that enriches the explanatory power of the model: step 3 in the elaboration of the formulae is a process of generalization that converted a formula into a pattern for the generation of a set of derived units. The result is an underlying structure that neatly fits the concept of «schema» as defined by Langacker (1987, 371):

A schema, . . . is an abstract characterization that is fully compatible with all the members of the category it defines (so membership is not a matter of degree); it is an integrated structure that embodies the commonality of its members, which are conceptions of greater specificity and detail that elaborate the schema in contrasting ways.

It follows, therefore, that the formulae underlying derived words are a formal representation of cognitive schemata which become lexically expressed. This cognitive view illuminates a number of issues in word formation and increases the explanatory power of the Functional Lexematic WFC. One of these issues is the polisemy at the norm level of derivational patterns; if we consider the following formations with the Spanish suffix *-ero/a*:

6- *tapicero*:

$$\phi_{\text{Theme}}^i [\Sigma_1: \text{PERFORM}_v (x_1: \langle \text{+Hum} \rangle (x_1))_{\text{AgentTopic}} (x_2: [\phi_{\text{Focus}} (x_1)_{\text{Agent}} (x_3: \langle \text{+Con} \rangle (x_3))_{\text{Meta}}]_{\text{ACTION}} (x_2))_{\text{Goal}} (\sigma_1: \text{repeatedly} \langle (\sigma_1)_{\text{Manner}} \rangle (\sigma_2: \langle \text{as an occupation} \rangle (\sigma_2))_{\text{Manner}} [\Sigma_1]_{\text{ACTION}}]$$

7- *bombardero*

$$\phi_{\text{Theme}}^i [\Sigma_1: \text{PERFORM}_v (x_1: \langle \text{+Art} :: \text{Imachine/vehicle} \rangle (x_1))_{\text{Force/Topic1}} (x_2: [\phi_{\text{Focus}} (x_1)_{\text{Force/Top.2}} (x_3))_{\text{Goal}}]_{\text{PROCCES}} (\sigma_1: \langle \text{as its typical function} \rangle (\sigma_1))_{\text{Manner}} [\Sigma_1]_{\text{PROCCES}}]$$

8-regadera

$\Phi_{nTheme}^i [\Sigma_1: \text{PERFORM}_v (x_1: \langle +\text{Art} \rangle (x_1^i))_{\text{Force/Topic1}} (x_2: [\Phi_{vFocus} (x_3))_{\text{Agent}} (x_4: \langle +\text{Con} \rangle (x_4))_{\text{Goal-Aff}} (\sigma_1^i: \text{with } \langle +\text{Art} \rangle (\sigma_1^i))_{\text{Instr/Topic2}}]_{\text{ACTION}} (x_2))_{\text{Goal-Eff}} (\sigma_2^i: \text{«as its typical function»} (\sigma_2^i))_{\text{Manner}} [\Sigma_1]]_{\text{PROCES}}$

the first conclusion that can be drawn is that the term «agent nominalization» is far too restrictive: no. 6 is the only case of a proper agent nominalization; no. 7 is a case of nominalization of a Force entity. What the two lexical units have in common is that they designate the first argument of both the underlying schema and the predicate frame of the (embedded) base unit⁵. So, it seems more appropriate to consider the two derived lexemes as instances of First Argument Nominalizations.

Example no. 8 is a case of instrument nominalization. There are several factors that link this kind of process with proper agent nominalizations; in fact, both agent and instrument nominalizations, in the case of Spanish and English, are generated by the same suffixes. Furthermore, the semantic connection between agent, force and instrument nouns is so close that there are numerous instances of a lexical unit having at least two of these meanings.

Studies in Cognitive Semantics provide us with tools to articulate in full the relations holding among all these different types of nominalization: Instrument and Force nominalizations are instances of a process of metaphorization of the notion of agency, by which inanimate entities are perceived as instigators of the SoA expressed in the underlying derivational schema. The distinction between force and instrument nominalization depends on the degree of metaphorization: in some cases the inanimate object is seen as the only instigator of the event, without any participation on the part of a controller (typically an agent) being perceived; for example:

cosechadora = «maquina que cosecha» («harvester»)

lavadora = «maquina que lava» («washing-machine»)

as can be seen, the semantic definition is expressed with a formula that says «something that Vs/something that performs X». Both are instances of total metaphorization, and both designate force entities. Instrument nominalizations are cases of partial metaphorization: the controller entity is not «supplanted» by the inanimate entity; it is only «downgraded» in terms of its relevance in the cognitive structure that the underlying derivational schema represents. This «degrading», which does not imply total disappearance, is reflected in the semantic definition for instrument nouns:

cafetera = «máquina para hacer café» (= «machine for making coffee»/machine with which *one* makes coffee«)

regadera = «objeto con el que se riega/ objeto para regar» (=«object (used) to water (plants)»):

These facts are explicitly marked by the co-reference index mentioned above, as can be seen in the derivational schemata for examples 12 and 14: in the case of the word *bombardero*, the first argument of the schema is coreferential with the first argument of the

⁵ This is expressed in the formulae by means of the superscript *i* that binds the lexical variable Φ_{nTheme} with one of the arguments of the schema; that is, agentive formations are instances of what in Laca (1993) is called «nominalizaciones orientadas».

embedded base frame (there being no argument with the semantic function of Agent); in the schema for *regadera*, however, the first argument (Force) is co-indexed with an instrument satellite, and not with the first argument of the embedded frame, the Agent which, nevertheless, is present in the conceptual structure.

Relations⁶ among schemata such as those of metaphorization can be considered horizontal, rather than hierarchical. But there are also other types of relations that do show a structured semantic hierarchy among various schemata. The derivational schemata of a given affixal unit are, in fact, vertically ordered in what can be called Conceptual Domains; one case of hierarchical relation, based on a routine of specialization relates the patterns underlying «pure» or «occasional» agent formations such as *doer*, *organizer*, «iterative» agent nouns like *drinker* («someone who drinks habitually/repeatedly») and designations for occupations like *researcher* or *publisher*.

The set of underlying schemata constitutes both the final stage of the analytic phase and the starting point for the synthetic component of the WFC. The synthetic module must be capable of accounting both for the production of complex lexical units from the underlying schemata, and for the relation of these units with different types of syntactic structures, at several stages. This is done by means of a set of expression rules. Given that I have been mainly concerned with the analytic phase of the word formation component, there is only a simplified version of the application of these rules to create the lexical unit *statistician* in the Appendix..

5. SCHEMATA AND WORD FORMATION DYNAMICS

We have seen before how some of the contributions from Cognitive Semantics can be integrated within the study of word formation. This adds a new perspective to the model proposed by Martín Mingorance, and opens the door for the addition of new «tools» not only to account for systematic relations between several derivational patterns, as has been shown, but also to explain factors concerning the dynamics of those patterns.

The identification of derivational patterns with cognitive schemata, in the sense of Langacker (1987), fits nicely with the conception of a complex lexical unit as the result of a categorization process motivated by the expressive needs of the speakers. Such needs are many and various: it may be the need to label entities in a discourse universe that are perceived as new, or the desire to present processes or events encapsulated as objects (f.i. *calculation* or *reduction*), or the constructions of events from some of the entities involved (as in *enslave*), etc.

But, given the synthetic character of these units, there hangs a basic communicative condition on them: the speaker that creates a novel lexical unit must make sure that this can be successfully decoded by the hearer for the referential act to be valid. This imposes the need for a word formation model to be as exhaustive as possible: a detailed description of the conditions that surround any derivational structure and of the underlying schemata will guarantee that any new lexical unit can be used in an effective way in a given

⁶ Langacker (1987, 377 and ff.) states that a complex category can be described as a schematic network, in which all nodes (= schemata) are interconnected by one of two types of categorization relations or cognitive routines: specialization and extension, or in other terms, total or partial schematization (schema inclusion), respectively. The degree of extension (discrepancy between schemata) varies, with instances of metaphorization being the most extreme cases.

communicative situation. However, any functional study of language must confront the paradox on which the dynamics of communication are based: on the one hand, speakers need to establish restrictions in order to develop and interpret all acts of communication; on the other hand, and because of the constant expressive needs, those restrictions must sometimes be broken. In this interplay, word formation has a crucial role: the ever present need to «label» compels speakers to extend the rules, conditions and schemata in order to create lexical units that one way or another will distance themselves from prototypical derived units. But, even so, they must show that there are enough connections with these prototypical constructions for them to be understood and, in turn, decoded.

These novel creations may stimulate the creation of others following the same conditions; in such cases, it is possible that these new lexical units become recategorized as a new derivational pattern. For example: over the last twenty years or so we have witnessed the creation of a new English and relatively productive morpheme, namely *-gram*. Based on the neo-classical compound *telegram* there have been recently a number of *-gram* words coined recently, e.g. *kiss-o-gram*, *Tarzan-o-gram*, *gorilla-gram* and even *potato-gram*.

In this context the terms *beautician* and *dietician*, which were encoded as exceptions in the entry for *-ician*, are worthy of comment. It was observed that the bases for these words did not satisfy the morphological conditions imposed by the suffix, by which it is attached to nouns and adjectives ending in [ik(s)]. They are also anomalous semantically since neither *beauty* nor *diet* can be considered scientific terms. Another anomalous formation from a semantic point of view would be *cosmetician*.

If those exceptions are considered as categorial extensions from prototypical constructions with the suffix *-ician*, it is possible to explain their creation by making use again of underlying schemata and restrictions, as is shown diagrammatically in the following chart:

Underlying Derivational Schema

$\phi_n^i \text{Theme} [\sum_1: \text{PERFORM}_V (x_1: \langle +\text{Hum} \rangle (x_1^i))_{\text{AgentTopic}} (x_2: \phi_n^i \text{Focus} \langle -\text{Attr} \ \varepsilon \ \text{sciences} :: \text{medicine} \rangle (x_2)_{\text{Goal Eff.}} (\sigma_1: \text{“repeatedly”} (\sigma_1)_{\text{Manner}} (\sigma_2: \text{“as an occupation”} (\sigma_2)_{\text{Manner}} (\sigma_3: \text{“as an expert”} (\sigma_3)_{\text{Manner}} [\sum_1] \text{ACTION}$

Prototype: optician

Prototype extensions:

- (1) From prototypes to *cosmetician*.

MOTIVATION: connotative level (implication of «technical term») → cultural motivation: nuance of «prestige».

PARAMETERS OF RELATION TO PROTOTYPE⁷: Functional similarity with gestalt properties:

Trajector = A₁ (+Human)

⁷ Brown (1920, 22) states:

category extension from a prototype to non-prototypes can involve relationships motivated by: (1) similarity; (2) metaphor; or (3) metonymy [...]. With respect to similarity, non-prototypes may be related to a prototype (a) if they are perceptually similar to a prototype or (b) if their propositional models specify some property or properties which are also possessed by a prototype.

In the case of **cosmetician**, **dietician** and **beautician**, their meanings are related to prototypical formations with *-ician* because there are comparable or similar features, like the implications of agentivity, iterativity, expertise, etc.

Connection of bases conceptual spheres (transfer of selection restrictions on base word variable): medicine ↔ aesthetics.

EFFECTS ON SCHEMA/CONDITIONS: extension on semantic restrictions on base word.

-(2) From prototypes (possibly via *cosmetician*) to *dietician*, *beautician*.

MOTIVATION: connotative level (implication of “technical term”) → cultural motivation: nuance of «prestige».

PARAMETERS OF RELATION TO PROTOTYPE: Functional similarity with gestalt properties:

Trajector = A₁ (+Human)

Connections of bases conceptual spheres (transfer of selection restrictions on base word variable): medicine ↔ aesthetics.

EFFECTS ON SCHEMA/CONDITIONS: cancelling of morphological condition on base (final cluster *-ic(s)*).

I have sought to demonstrate in this paper that a Functional Lexematic model of word formation has more explanatory power than the standard FG treatment of this phenomenon; such explanatory power is further increased through the integration of several elements from Cognitive Semantics, which allow for a more complete understanding of the intrinsic dynamic nature of word-formation, such that we can now place it not only at the crossroads of morphology, syntax, semantics and the lexicon, but also at the crossroads of synchrony and diachrony.

REFERENCES

- Aarts, Jan M.G. and J.P. Calbert. 1979. *Metaphor and Non-Metaphor (The Semantics of Adjective-Noun Combinations)*. Tübingen: Max Niemeyer Verlag.
- Aronoff, Mark. 1979² [1976¹]. *Word-Formation in Generative Grammar*, Cambridge, Mass.: The MIT Press.
- Brekke, Herbert. E. 1976² [1970¹]. *Generative Satzsemantik im System der englischen Nominalkomposition*. München: Wilhem Fink.
- Bresnan, Joan. (ed.). 1982. *The Mental Representation of Grammatical Relations*. Cambridge, Mass.: The MIT Press.
- Brown, Cecil. H. 1990. «A survey of category types in natural language». In S.L. Tsohatzidis (ed). *Meanings and Prototypes (Studies in Linguistic Categorization)*. London & New York: Routledge.
- Chomsky, Noam. 1970. «Remarks on nominalization». In R. Jacobs & P. Rosenbaum (eds). *Readings in English Transformational Grammar*. Waltham, Mass.: Ginn & Co.
- 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass.: The MIT Press.
- 1986. *Barriers*. Cambridge, Mass.: The MIT Press.
- Cortés Rodríguez, Francisco J. 1996. «La morfología derivativa en la Gramática Funcional de Dik. ¿Formación de palabras o formación de predicados? *Alfinge* 9. (in press).

- Coseriu, Eugenio. 1962. *Teoría del lenguaje y lingüística general*. Madrid: Gredos.
- 1978. *Gramática, Semántica, Universales*. Madrid: Gredos.
- De Groot, Casper. 1987. «Predicate formation in Functional Grammar». *Working Papers in Functional Grammar 20*. Univ. of Amsterdam.
- Dik, Simon C. 1978. *Functional Grammar*. Dordrecht-Holland/Cinnamison-USA: Foris.
- 1989. *The Theory of Functional Grammar. Part I: The Structure of the Clause*. Dordrecht: Foris.
- Fabb, Nigel. 1984. *Syntactic Affixation*. MIT Dissertation. Cambridge, Mass. The MIT Press.
- Halle, Morris. 1973. «Prolegomena to a theory of word-formation». *Linguistic Inquiry* 4,1: 3-16.
- Hammond, Deanna L. 1980. *Word Formation in Generative Grammar: Spanish Derivational Morphology*. Doctoral dissertation. Ann Arbor: UMI.
- Hengeveld, Kees. 1992. *Non-verbal Predication. Theory, Typology, Diachrony*. Berlin/New York: Mouton de Gruyter.
- Kastovsky, Dieter. 1977. «Word-formation, or: At the crossroads of morphology, syntax, semantics, and the lexicon». *Folia Linguistica* 10: 1-33.
- Laca, Brenda. 1993. «Las nominalizaciones orientadas y los derivados españoles en *-dor* y *-nte*». In S. Varela (ed.). pp. 180-204.
- Langacker, Ronald W. 1987 *Foundations of Cognitive Grammar vol. 1*. Stanford, Cal.: Stanford UP.
- Lees, Robert B. 1960. *The Grammar of English Nominalizations*. Bloomington, Ind.: Research Centre in Anthropology, Folklore and Linguistics.
- Mackenzie, John L. 1986. «Aspects of nominalization in English and Dutch». *Working Papers in Functional Grammar 15*. Univ. of Amsterdam.
- 1993. «Nominalization in a Functional Grammar of English». Ms. of seminar taught at Univ. of La Laguna (Nov-December 1993).
- Marchand, Hans. 1969² [1960,]. *The Categories and Types of Present-Day English Word-Formation. A Synchronic-Diachronic Approach*. München: C.H. Beck'sche Verlagsbuchhandlung.
- Martín Mingorance, Leocadio. 1985. «Bases metodológicas para un estudio contrastivo del léxico derivado». *Revista Española de Lingüística Aplicada* 1: 37-54.
- Pottier, Bernard. 1974. *Linguistique générale*. Paris: Klincksieck.
- Reichl, Karl. 1982. *Categorial Grammar and Word-Formation: The Deadjectival Abstract Noun in English*. Tübingen: Max Niemeyer Verlag.
- Roeper, Thomas & M. Siegel. 1978. «A lexical transformation for verbal compounds». *Linguistic Inquiry* 9: 199-260.
- Scalise, Sergio. 1987. *Morfología generativa*. Madrid: Alianza.
- Shibatani, Masayoshi & T. Kageyama. 1988. «Word formation in a modular theory of grammar: postsyntactic compounds in Japanese». *Language* 64, 3: 451-482.
- Siegel, Dorothy. 1979. *Topics in English Morphology*. New York & London: Garland.
- Siewierska, Anna. 1991. *Functional Grammar*. London/New York: Routledge.
- Sproat, Richard. 1985. *On Deriving the Lexicon*. Doctoral Diss. MIT.

- 1988. «Bracketing paradoxes, cliticization and other topics: The mapping between syntactic and phonological structure». In M. Everaert & A.L. (eds) *Morphology and Modularity: In Honour of Henk Schultnik*. Dordrecht: Foris. pp. 339-360.
- Szymanek, Bogdan. 1980. «Phonological conditioning of word formation rules» *Folia Linguistica* XIV,3: 413-425.
- 1985. «Disjunctive rule ordering in word formation». *Papers and Studies in Contrastive Linguistics* 20: 45-64.

APPENDIX

LEXICAL ITEM (FROM LEXICON OF AFFIXES)

AFFIX: /+ / [] ∂n/

(A) SPELLING ALTERNANTS

– ICIAN

– ITIAN. Asystematic and idiosyncratic in *dietitian/dietician*.

(B) PHONOLOGICAL DESCRIPTION

Segmental form [/] ∂n]

Allomorphs Vacuous

Affixal Type +X##_{Suffix}

Stress Location

[/] ∂n] / [¹S₁ S₁]_{Base} _____ ([³S₁ ¹S₁ / [] ∂n]₁)E.g.- β/θ] { ¹dξ/ } [] ∂n} , βmju: } { ¹s/ } [] ∂n}[/] ∂n] / [S_{1-n} ¹S₁ S₁]_{Base} _____ → [(S_{0-n}) ²S₁ ³S₁ ¹S₁ / [] ∂n]E.g.- { ²kos } { ³me } { ¹t/ } [] ∂n} , { ²θ } { ³dξ/s } { ¹t/ } [] ∂n}

Syllable Distribution Disyllabic suffix. Suffix initial [/] → heterosyllabic

(C) INPUT CONDITIONS

Phonological Vacuous

Morphological Base Final Cluster: [/k(s)] . Except: *beautician, dietician*.

Substratum Latinate technical terms

Frame Types

NOMINAL PREDICATES

TYPE 1

φ_n (x₁: NP <-Attr (sciences/crafts (: philosophy/logic/medicine)> (x₁))_δ

E.g.- magician, phonetician, mathematician, obstetrician

TYPE 2

φ_n (x₁: NP <+Attr +Ev (aesthetics> (x₁))_δ

E.g.- beautician, cosmetician, dietician

ADJECTIVAL PREDICATES

TYPE 1

φ_{adj} (x₁: Adj P <+Sta -Phy (sciences/crafts (: medicine)> (x₁))_{Ads}

E.g.-technician, clinician

(D) MORPHOLEXICAL OPERATIONS (Immediate Constituents Structure)

φ_n → [[φ_n]_{Base} +ician]_{Suf_n} Denominal nounsφ_{adj} → [[φ_{adj}]_{Base} +ician]_{Suf_n} Deadjectival nouns

(E) OUTPUT CONDITIONS Truncations

(1) [[_____ { **C**k }]_{Base} / [] ∂n]_n → [_____ { **C** } / [] ∂n]_n

mu • si • cian , a • riñh • me • ti • cian , phy • si • cian

E.g.- clinician, diagnostician

SYNTHETIC PHASE

ITEM: statistician

CYCLE I:

Underlying Derivational Schema

$\phi_{n}^{iTheme} [\sum_{1}: PERFORM_{V} (x_{1}^{i}: <+Hum> (x_{1}^{i}))_{AgentTopic} (x_{2}: \phi_{n}^{iFocus} <-Attr \varepsilon sciences/crafts> (x_{2}))_{Go-}$
 $al\ Eff. (\sigma_{1}: "repeatedly" (\sigma_{1}))_{Manner} (\sigma_{2}: "as\ an\ occupation" (\sigma_{2}))_{Manner} (\sigma_{3}: "as\ an\ expert" (\sigma_{3}))_{Manner}$
 $[\sum_{1}]]_{ACTION}$

Arguments Specification

(x₁: "somebody"(x₁))

(x₂: "statistics"(x₂))

CYCLE II:

Reduction to Sintagmatic Structure

(A) Omission of argument variables:

$\phi_{n}^{i} \rightarrow \langle \text{Perform somebody}_{Top} \text{ statistics}_{Foc} \text{ (repeatedly, as an occupation, as an expert)} \rangle$

(B) Overt expression of co-reference variable + Linearization:

$\langle \text{Somebody}_{Top} \text{ that performs statistics}_{Foc} \text{ (repeatedly, as an occupation, as an expert)} \rangle$

Reduction to Morpho-syntagma:

(A) Assignment of morpho-syntagmatic functions:

Topic \rightarrow *Determinatum (DM)*

Focus \rightarrow *Determinant (DT)*

ϕ_{n}^{i} : somebody(DM)+ statistics(DT)

(B) (B) Affix search:

Retrieve : *-er* \rightarrow Discard: non-congruence with arguments restrictions

-ant/ent \rightarrow Discard: non-congruence with base category

-ist \rightarrow Discard: non-congruence with base morphology

-ician \rightarrow SELECTED

(C) Morpho-syntagma:

ϕ_{n}^{i} : *-ician*(DM)+ statistics(DT)

CYCLE III: Reduction to lexical unit

(A) Word-Formation Rule:

$\phi_{n}^{i} ([[\text{statistics}_{n}]_{Base} + \text{ician}]_{Surf_{n}} \text{Denominal nouns}$

(B) Adjustment rules:

Truncation:

$[[\text{st}\partial\text{t}/\text{s} \{ \text{t}/\text{ks} \}]_{Base} / \int \partial n]_{n} \rightarrow [\text{st}\partial\text{t}/\text{s} \{ \text{t} \} \{ \int \partial n \}]_{n}$

Allomorphy:

$[\text{st}^{\partial}\partial\text{t}/\text{st}/\text{ks}]_{Base} \rightarrow [\text{st}^{\partial}\partial\text{t}^{\partial}/\text{st}^{\partial}/ \int \partial n]_{n}$

OUTPUT (LEXICAL ITEM):

statistician $[\text{st}^{\partial}\partial\text{t}^{\partial}/\text{st}^{\partial} / \int \partial n]$