Hans-Heinrich Lieb

Towards a general theory of word formation: the Process Model

Foreword

The present essay—longer than a paper but shorter than a book—characterizes the *Process Model of Word Formation* that represents a new approach to word formation intermediate between constructionist and generative approaches; the model will be elaborated in detail in: Lieb, Hans-Heinrich (in prep.), *The Process Model of Word Formation and Inflection*. Amsterdam/Philadelphia: Benjamins. The essay, which is independent of the book, replaces an earlier, unpublished manuscript (Lieb 2011/2012), of which it is a completely revised and enlarged version. The essay was completed in July 2013; it is an outcome of work undertaken by the author since roughly 2006 but originating from still earlier work (first presented at a Research Colloquium held at the Freie Universität Berlin in 2001, and subsequently by a lecture read at the Annual Meeting of the *Deutsche Gesellschaft für Sprachwissenschaft* in 2006: Lieb 2006).

 The Freie Universität Berlin also houses a major effort at producing book-length Open Access publications in linguistics, organized into series: Language Science Press, langsci-press.org. The present essay does not fit this framework, both for lack of a suitable series and for being shorter than a book. An Open Access format was chosen for the essay by its author for a number of reasons, most importantly, to provoke discussion (contact the author via: edocs@ub.fu-berlin.de).

Please quote this essay as follows, together with its url:

Lieb, Hans-Heinrich. 2013. Towards a general theory of word formation: the Process Model. Berlin: Freie Universität Berlin. (An Open Access publication.)

The author is Full Professor (em.) of General and German Linguistics at the Freie Universität Berlin in Berlin, Germany.

Berlin, July 2013

Hans-Heinrich Lieb

Contents

A. Introduction and background

1 Introduction

1.1 Aim, method, and theoretical background

1.1 a. Aim

Where does word formation belong? Does it really belong in morphology? Exactly what is being formed in word formation? Is word formation stem formation? Is the formation of lexical idioms—with many-word forms—word formation? Is word formation to be construed in terms of processes, or process results? How are different kinds of word formation to be defined? What exactly is it that is shared by different kinds of word formation, such as compounding, derivation, and conversion? What exactly is the difference between word formation and inflection, and what do they have in common? What is it that makes one kind of word formation, such as compounding, the same kind in different languages? How do we arrive at a consistent, disambiguated terminology for dealing with word formation? What is an appropriate format for word-formation statements?

There continues to be considerable disagreement on these questions, and no current approach appears to have answers to all of them. It is the aim of the present essay to characterize a conception of word formation and a theory—meant to be adequate for all formation types—that would allow us to answer the questions in a unified way. The theory is to be called the *Process Model of Word Formation*.

The conception and the theory have four basic aspects:

(i) they allow for a clear distinction to be drawn between what is morphological and what is syntactic in word formation;
(ii) they assume that only three basic processes are involved in word formation: compounding, derivation, and conversion, each of them morphosyntactic because of its two major subcases: a *morphological* process resulting in stem forms, and a *syntactic* process resulting in word forms;
(iii) they construe all processes involved in word formation in a unified way;
(iv) they interlink with a specific Word and Paradigm model.

The first aspect, touching upon a basic problem of long standing in word-formation theory—topical in Construction Grammar work—is left undiscussed here in its general form. Impasses and oppositions appear to prevail in interrelating the lexical, the morphological, and the syntactic, both within and between approaches (for recent discussion, see Blevins 2006, Müller 2006, Jackendoff 2011). It may be claimed, though, that the Process Model of Word Formation avoids them by adopting (ii), (iii), and (iv). Aspects (ii) and (iii) are at the centre of this essay; the fourth aspect is treated only as far as necessary for dealing with the other two.

1.1 b. Comments

Offering a new model in an area that has been researched for centuries if not millenia and where different approaches have been, and still are, competing with one another, is daring to say the least, and the model's newness may be doubted right away. I am well aware of this. The Process Model should still be important as an attempt to rethink basic questions and come up with answers that are truly comprehensive and unified. A detailed comparison of the model with existing approaches would be in order. Unfortunately, this proved impossible within the confines of the essay; in particular, no systematic representation of the literature has been attempted. A choice had to be made between characterizing the conception by means of examples and presenting part of the theory vs. providing a more detailed discussion of the literature. There are, however, references of a historical nature, and a selection was made from the more recent literature. The following remarks are to give a first impression of how to place the model in the field.

The conception and the theory are non-constructionist: word formation is not covered by using a notion of construction as developed in some version of Construction Grammar; instead, notions of process are taken to be basic. Yet, the conception is non-generative: word formation is not treated by means of rules for generating formal objects related, in one way or another, to 'language' or 'languages'. Roughly, word formation *in* languages is to be described directly by means of statements *on* languages; grammars are assumed to be 'radically declarative'. The conception and the theory share their declarative outlook with constructionism but combine it with a process orientation; in this respect, they are intermediate between constructionist and generative approaches.

The essay presupposes rather than adds to the vast amount of studies on word formation and inflection in individual languages, characterizing a theory that attempts to systematize their results. Thus, the essay addresses morphologists and theoreticians interested in such systematization. However, as it is an ultimate aim of the theory to provide an improved framework for actual word-formation studies, the essay will also be of interest to descriptive linguists willing to familiarize themselves with a new approach in their field that is descriptively relevant (see Sec. 7, below).

The framework characterized in the present essay may be used for research from a semasiological, 'form to meaning' point of view, but also for research adopting an onomasiological, 'meaning to form' perspective (prominently applied by Štekauer, cf. Štekauer 2005). The framework allows for an onomasiological perspective because 'semantic' entities (concepts and functions operating on them) are conceived as non-linguistic entities that can be precisely identified and whose role in word-formation processes may then be studied.

Concepts are construed as content-based properties of the conceptions or perceptions that actual speakers may have, which makes lexical meanings testable. This is one example of how the abstract ontological entities assumed by the word-formation theory—and ultimately, by the overall theoretical framework—can be empirically justified. The theory and the framework are sufficiently explicit for their ontological commitments to stand out; the commitments are alien to the ones that characterize Chomskyan generative grammar as critically discussed—with peer evaluation—by Stokhof & van Lambalgen (2011).

The conception and the theory aim at complete coverage: if successful, all phenomena of 'word formation in the world's languages' as outlined in the recent typological survey by Štekauer, Valera & Körtvélyessy (2012) are—in principle—covered.

1.1 c. Method

The conception of word-formation processes embodied in the Process Model is characterized in Part B (Secs 3 to 5), the Process Model itself is outlined (and confronted with a related model) in Part C. The method of characterizing the conception essentially consists in presenting a detailed analysis of some interrelated, non-trivial examples from a single language (English), leading up to a sketch of part of the theory of word formation. The examples belong to a ‘word family’, a set of lexical words interrelated by word-formation processes. Only occasionally will there be reference to other examples, either from English or from other languages.[1]

The method of characterization implies a reduction of language material actually presented in the essay. This is compensated for by an in-depth analysis of the material. Special attention is paid to *semantic* detail, both with respect to lexical and to grammatical meaning.

The presentation chooses examples (not to be confused with the empirical basis of the conception or the theory) strictly according to explanatory need, indicating how basic questions of word formation will be answered but not following up their ramifications in individual languages: the essay is to open up vistas rather than explore sceneries.

The presentation of the conception, the theory and their background is informal or semi-formal; no more is required on part of the reader than some basic knowledge of naive set theory. (An intensional logic would be needed for a more technical account, not attempted here, due to the conception of word meanings.) The essay may still be demanding, aiming at comprehensiveness and touching upon a large number of topics in a novel way.

1.1 d. Background

Any theoretical conception requires a theoretical background. The conception of word formation to be presented here and the correlated theory make some use of the framework of Integrational Linguistics (IL) as developed by Lieb and others (see Sackmann 2006, 2008), a non-generative approach sharing some of its outlook, in syntax, with constructionist and other declarative frameworks; no knowledge of the theoretical background is presupposed. Part of the background is a Word-and-Paradigm model expanded into a *Word, Lexeme and Paradigm model*: lexical words are distinguished from lexemes, and lexemes—which are to include not only stem lexemes but also affixes—are construed as entities of the same general type as lexical words. This WLP model, largely taken over from Lieb (2005) and not to be justified here, is the only part of the theoretical background that will be characterized in some detail (below, Sec. 2).

Let me emphasize that the present essay should be of interest independently of the framework it uses: it characterizes a foundational attempt in an important area of general linguistics, with consequences for descriptive work; the relation to a specific framework is secondary.

---

[1] The language used for initial orientation was German, not English, a language known for its rich word-formation potential; since, versions of the conception and the theory have been applied, in varying degrees, to German, Polish, English, French, and Chinese (no publications yet).

1.2 Word formation: the approach

1.2 a. The conception

Some major features of the conception are listed here, in a summary fashion:

(1)

a. Word formation is construed as the *formation of lexical words in their non-inflectional aspect*; the words are conceived as pairs ⟨P, b⟩ consisting of a word paradigm P and a meaning b of P that is a concept. Lexical words are conceived as *syntactic* entities.

b. Word formation is understood as, roughly, the formation of a lexical word ⟨P, b⟩ from pairs ⟨P_1, b_1⟩ and ⟨P_2, b_2⟩ through a word-formation process construed as in (e), below; the second and third pairs are, typically, lexical words or stem or affix 'lexemes', which are morphological entities. Having *at least two* positions for pairs ⟨P_1, b_1⟩ and ⟨P_2, b_2⟩ is compatible with conversion, where ⟨P_2, b_2⟩ = ⟨P_1, b_1⟩ (as it is in reduplication); having *no more than two* positions is compatible with copulative compounding (argued below, in Sec. 3.6).

c. There are just three basic word-formation processes in an idiolect system S (rather than a language, to account for language variability): *the compounding process, the derivation process*, and *the conversion process in* S. These are directly given in a component of the system, its *word-formation base*.

d. Each basic word-formation process in an idiolect system has two major *subcases* (which may be proper or improper parts of the process), one a *stem-form* or *morphological subcase* resulting in 'word-stems' of the system, the other a *word-form* or *syntactic subcase* resulting in word forms (forms of lexical words) of the system, forms that are non-analytic and, typically, non-synthetic, idiomatic. The word-stems are morphological units, the word forms are syntactic units; the three basic processes are therefore *morphosyntactic*, and the word-formation base containing them is a *morphosyntactic* component of the idiolect system, separate from and additional to its morphological and syntactic components. Six major subcases are obtained for the basic processes in an idiolect system S; these, too, are word-formation processes in S unless empty: *stem-form* or *morphological compounding / derivation / conversion processes*, and *word-form* or *syntactic compounding / derivation / conversion processes* in the system. When a stem-form process is used in an idiolect system, we will speak of *morphological word formation*, when a word-form process is used, of *syntactic word formation*.

e. All word-formation processes in an *idiolect* system, be they basic processes or subcases of basic processes, are construed as *functions* φ *of a single formal type*, in the logical sense of 'function'. Stem-form processes and word-form processes are distinguished not by their arguments but by their values (stem forms vs. word forms). A *subcase* of a basic process φ in a system is, informally, a function φ_1 that is a part of the basic process φ.

f. *Word formation in an idiolect system* is construed as a relation between three pairs ⟨P, b⟩, ⟨P_1, b_1⟩ and ⟨P_2, b_2⟩ as in (b) and a function φ as in (e); this relation is no function. The relation is supplemented by a function of *word-category assignment* (see below, Sec. 1.5 c).

g. *Word formation per se* is a function that takes arbitrary idiolect systems S in arbitrary languages as arguments and assigns to each S a relation as in (f).

h. *Types of word formation in* S, such as derivation in S, compounding in S, and conversion in S, are derivative on word formation in S and on specific word-formation processes, construed as functions as in (e).

i. All aspects of the construction of lexical words that are *not* covered by word formation—such as the construction of analytic forms—are covered by various *processes of inflection*, construed as functions of the same formal type as the functions involved in word formation.

1.2 b. Comments

The word lexicon of an idiolect system (to be distinguished from its lexeme lexicon) may be subdivided into the set of *actual lexical words* (words-in-use) of the system and the set—possibly non-finite—of its *potential lexical words* (words not in use). Independently, it may also be subdivided into the set of the system's *given* or *basic words* (the words not determined by word-formation processes) and the set of its *non-basic* words (determined by word-formation processes). All basic words are actual but some actual words may be non-basic: these are words in use that are 'transparently formed'. For example, the *door*-, *lock*-, and *door lock*-words in (2), below, are actual words in Standard English idiolect systems but only the first two are basic. All potential words (words not in use) are non-basic (are determined by word-formation processes). For example, a lexical word like *oxygen lock* meaning 'lock for a container filled with pure oxygen' should be a potential, hence, a non-basic word in many English idiolect systems, but not an actual word.

These distinctions are based on the ones made by Budde (2012). Traditionally, only one classification tends to be made in this context: 'actual word' vs. 'potential word', where 'actual' is understood roughly as above but 'potential' corresponds to our 'non-basic'. But this prevents us from recognizing a set of words-in-use that are determined by word-formation processes. Rainer (2012) makes a plea for considering so-called blocked words—English *stealer*, blocked by *thief*—as *virtual* words. We may include this by classifying the non-basic lexical words into virtual and non-virtual ones.

The Process Model construes word formation processes in an idiolect system as, roughly, *the non-inflectional processes used in the identification of the non-basic lexical words of the system*, including actual words that are 'transparently formed'. This means that complete or partial *semantic transparency* (for the speaker on reflection, not just for the linguist) not *productivity* is the leading criterion in characterizing a lexical word as determined, or not determined, through word formation; the criterion is to allow for lexicalization.

The present essay is concerned primarily with clarifying the nature of the functions φ in (1e), the functions with which word-formation processes in idiolect systems S are identified.

Basic processes and their interrelations as well as subcases of basic processes will be considered. It may come as a surprise that only three basic processes are assumed, given comprehensiveness as an aim of the Process Model. The compounding process, the derivation process, and the conversion process in a system S should be sufficient, though, due to a broader than usual conception of the conversion process: lifting two traditional restrictions on conversion—there must *not* be a change of form (excepting 'minor' changes), and there *must* be a part-of-speech change—we are able to distinguish many more subcases of the conversion process than are usually allowed, covering the processes involved in some troublesome kinds of word formation, such as back formation, short-word formation, or the formation of acronyms; at the same time, the notion of conversion process remains clearly defined, and defined in a way that covers the traditional cases.

On the conception of basic processes to be characterized, the compounding process, the derivation process ('derivation in a narrow sense') and the conversion process in an idiolect system are naturally ordered in this way: together with their non-empty subcases, they form a (non-continuous) three-part *process cline*, or scale. All word-formation processes involve a 'basic triple' and an 'added triple' in their arguments. The basic triple is what the process starts from; the added triple is what the basic triple is combined with by the process. With all three processes, the basic triple normally consists of a morphological or a syntactic unit, a categorization of the unit, and a lexical meaning of the unit. The three processes differ with respect to the added triples. In the case of the compounding process and the derivation process, the added triple is just like the basic triple, with the additional requirement that the lexical meaning in the added triple must be a 'non-empty concept' when we are dealing with the compounding process, and must be 'the empty concept' when dealing with the derivation process. The conversion process in turn requires an added triple that consists of three 'empty' components, in particular, has the empty concept as its third component; this triple is a purely formal entity. The compounding process, derivation process, and conversion process and their non-empty subcases, arranged in this order, constitute a cline of 'decreasing content' with respect to added triples. The derivation process takes the middle position, partly agreeing with the compounding process and partly with the conversion process. This would be blurred by introducing 'derivation in a broad sense' as a basic word-formation process, with the processes of derivation (in a narrow sense) and conversion as subcases. On the other hand, existence of the process cline is a strong argument against any further unification of the three basic word-formation processes.

From a systematic point of view, basic processes should be discussed before their subcases. This is not how I will proceed. For each basic process, the two major subcases—the stem-form subcase and the word-form subcase—will be presented first, and presented separately; the two subcases are word-formation processes in S unless empty. It will appear that the stem-form process (morphological) and the word-form process (syntactic) may indeed be construed as subcases of a single basic process (morphosyntactic). This is not self-understood: in earlier versions of the Process Model, stem-form processes and corresponding word-form processes were treated as merely analogous, not as subcases of more general processes. I therefore start not with the three basic processes but, for each of them, with the two major subcases. The basic processes will be discussed in Sec. 5, and formally characterized in Sec. 6 when the Process Model is outlined.

1.3 Historical remarks

The following remarks, which are far from exhaustive, may still be helpful in relating the Process Model of Word Formation, and the conception on which it is based, to linguistic tradition and to some of the more recent approaches.

For some time now, a certain move can be seen in the literature towards explicit recognition of word formation as morphological on the one hand and syntactic on the other; a process view of word formation may or may not be implied. One example is Drude, who in an extensive digression (2004: 184-192) comes to distinguish explicitly between *syntak-*

tische Wortbildung ('syntactic word formation') and *morphologische Wortbildung* ('morphological word formation'); again, Drude (2010: 293-300); see also Manova & Dressler (2005: e.g., 71-72) for 'morphological' vs. 'syntactic conversion', Manova (2011: Sec. 3.5.3) on 'syntactic conversion', and Eisenberg (1998: 280-285), (2006: 296-300) for *morphologische Konversion* vs. *syntaktische Konversion*, which, however, Eisenberg considers as not leading to a 'typical word-formation product'. Anderson (1992: Ch. 8) concludes "that 'special clitics' are actually the 'morphology' of phrases, parallel in fundamental ways to the morphology of words" (221), a parallelism worked out by Anderson in subsequent publications, especially Anderson (2005).[2] In his constructionist framework, Booij (2010: 101f) distinguishes 'syntactic compounding', or 'quasi-incorporation', from 'morphological compounding'; generally, "An important claim of Construction Morphology is that phrasal constructs may be similar in function to morphological constructs in that they function as lexical units and provide names for concepts." (Booij 2012: 344); see also Masini & Benigni (2012) on 'phrasal lexemes' in Russian as analysed in a Construction-Morphology framework, and the literature quoted there both for Russian and other languages. The general problem of the morphological vs. the syntactic in compounding is a recurrent major topic in the contributions to Lieber & Štekauer (2009).

Notions like 'inflection' or 'derivation' in 19th century historical linguistics are obviously process notions. A good discussion of such notions continues to be the one by Matthews (1974: Ch. VIII), where the nature of abstract linguistic processes as functions is recognized, as it is, implicitly, in informal accounts of 'morphological operations' like the one given by Booij (2007: Sec. 2.2). Authors working in a generative framework also tend to recognize that they are dealing with functions, irrespective of formal apparatus: as acknowledged by Anderson (1992: 186), and emphasized by Aronoff (2000: 205, on Aronoff 1976) for the 'Word Formation Rules' of lexicalist versions of generative morphology. Beard's Lexeme-Morpheme Base Morphology (Beard 1995) is explicitly function-oriented by allowing 'four and only four mutually independent types of operations' on components of 'any fully specified lexeme' (1995: 47), in a theory that is to cover both word formation and inflection.

A functional perspective is also implicit in some HPSG work. Müller (2002: 358) mentions a number of authors who have followed an Item-and-Process approach in an HPSG framework, using "lexical rules that relate stems to other stems or words" (358), and goes on to adopt such an approach himself (359).

It is, however, Hockett (1954) in his famous 'Two Models' paper who first distinguished the 'Item-and-Arrangement' model (IA) from the 'Item-and-Process' model (IP) and stated unequivocally that abstract processes, as in "the carryover of 'process' terminology from historical linguistics" (211), are best construed as functions in the logical (Hockett: mathematical) sense, which according to Hockett quite generally provides the formal underpinning for an Item-and-Process approach (1954: 227).

IA has remained dominant in typological linguistics. Construction Morphology (Booij 2010) also appears to be basically IA, notwithstanding a claim to the contrary recently made by two of its practitioners (viz., Masini & Benigni 2012: 445).[3]

---

[2] However, derivational *clitics* as recognized by Anderson in both (1992: 218) and (2005: 133, 169-170) appear not to include derivational *particles* (see below, fn. 37), due to the fact that clitics are introduced 'post-lexically' (Anderson 2005: 34).

[3] The following relationship appears to exist between Booij's Construction Morphology, in its word-formation part (IA), and the Process Model of Word Formation (IP). Given a formally explicit grammar that presupposes the Process Model (such a grammar is necessarily 'declarative'), a word-formation schema in

IP subsequently took an algorithmic turn in Generative Grammar, blurring Hockett's more general discovery. True, in the most basic sense an algorithm is simply a procedure for determining the values of a function, in particular where recursiveness is involved. It continues to be a disputed question if, or how far, processes that occur in word formation do involve non-trivial recursiveness. Still, even a non-recursive process may be considered for reconstruction as a function.

Hockett's own attempts in (1954) to construe word-related processes as functions are hardly satisfactory. Thirty years later, a serious attempt was made by Hoeksema (1985), in the framework of Categorial Grammar, to develop a strictly functional account of word formation (his 'lexical rules' are functions: Hoeksema 1985: Sec. 1.6.2); again, Hoeksema & Janda (1988). My own account will overlap with Hoeksema's at an important point: I assume 'fully specified units' that are similar to a construct used by Hoeksema, but also to constructs used by other authors (see Sec. 2.6, below).

Hockett (1954) mentions (but not characterizes) Word-and-Paradigm models as an alternative to both IA and IP. The word-formation framework outlined in the present paper combines a process conception of word formation with a WP conception of words, stems, and affixes, thus demonstrating, once again, that IP and WP are no alternatives of 'grammatical description' but may be used in conjunction.

Combining Word and Paradigm with Item and Process is an approach already followed by Anderson (1992), who uses WP for inflection and IP for word formation. I attempt to integrate the two models by, among other things, applying a single process view in dealing with both word formation and inflection. This is different again from Stump's extensive use of functions *within* his WP model (Stump 2001), where word formation is not yet treated in its own right. (Stump's proposal for dealing with word formation—as made in Stump 2005 and presupposed in Stump 2010—is, treating it *on the analogy of* a WP model for inflection.)

Manova (2011) combines Natural Morphology and Cognitive Linguistics and, according to Dressler (2011: vi), "This monograph felicitously mixes properties of item-and-arrangement grammars and of item-and-process grammars." And indeed, Manova assumes five basic 'morphological techniques' (2011: Sec. 2.4: 'addition', 'substitution', 'modification', 'conversion', 'subtraction') that correspond to traditional morphological operations both in word-formation and inflection and would have to be rendered by functions on a more formal account (not attempted by Manova). However, the 'morphological techniques', including conversion in Manova's sense, by and large correspond to functions like shortening or form change (or tuples or products of such functions) that on our account figure in the arguments of word-formation or inflection processes φ, rather than *being* such processes.

Manova (2011) also represents a recent example of prominently using the notion of a non-continuous cline, or scale, in morphology. (The use is informal, as is typical in linguistics; the formal explication given in Lieb 1992b: 186-188 is an exception.) Manova

---

the sense of Booij can be translated into an open sentential formula of the language in which the grammar is formulated, whereas word-formation rules of the grammar (below, Sec. 7.3) cannot be translated into word-formation schemata as assumed by Booij, for the simple reason that the schemata are insufficient as a basis for formally reconstructing the general notions of the compounding process (in S), the derivation process (in S), and the conversion process (in S). (Booij does speak of 'compounding', 'derivation', and 'conversion' but uses these terms informally as part of his extratheoretical language.) In this way, then, the Process Model of Word Formation is more general than Construction Morphology, while Construction Morphology is not incompatible with it.

introduces non-continuous clines in close connection with Prototype Theory and employs clines for grading the transition from word formation to inflection, but also for interrelating 'morphological techniques'. Only the last feature—interrelating techniques—is somewhat similar to the way the notion of cline has been used, independently, in the Process Model of Word Formation. As for a 'word-formation/inflection continuum', the following position is embodied in the Process Model: there is no such continuum, or scale; inflection processes are functions of the same formal type as word-formation processes but are distinguished from them by positive conditions that are *not* satisfied by word-formation processes (see Sec. 4.5d, below); therefore, inflection processes are not to be ordered behind word-formation processes by simply relaxing certain requirements that word-formation processes must meet.

Word formation has also been studied using versions of Integrational Linguistics, or inspired by them, notably by Lieb (1983: e.g., Secs 14.4, 15.2, on semantic aspects of word formation), Eisenberg (1998), (2006), Fuhrhop (1998), Eschenlohr (1999), and Drude (2010). Until very recently, Integrational Linguistics has been characterized by a combination of IA and WP, and this is typical, too, of all published Integrational work on word formation and morphology. This orientation is now being changed for word formation and morphology by the word-formation theories developed by Lieb (the Process Model of Word Formation, this essay) and Nolda (2012b) (to be discussed below, in Sec. 8); both theories represent an Item-and-Process approach. Integrational Linguistics is therefore beginning to combine all three approaches distinguished by Hockett some fifty years ago: IP, IA, and WP.

1.4 Word-formation statements: example and conventions

The following statement, given in a less formal version and a more formal one, is in agreement with (1); version (a) is to count as a *reading* of version (b):

(2) (Let S be a suitable English idiolect system.)

a. ⟨*door lock*^{P}, ·lock for door·⟩ is formed in S from ⟨*lock*$_{1}^{P}$, ·lock$_{1}$·⟩ and ⟨*door*^{P}, ·door·⟩ through stem-form compounding in S.

b. ⟨⟨*door lock*^{P}, ·lock for door·⟩, ⟨*lock*$_{1}^{P}$, ·lock$_{1}$·⟩, ⟨*door*^{P}, ·door·⟩, stem-form compounding in S⟩ ∈ word formation in S (wf(S)).

A number of *notational and terminological conventions* used in the present essay are introduced or presupposed in (2)[4]:

(i) An expression like 'stem-form compounding in S' is to be understood as short for 'the stem-form compounding process in S'. Such terms refer to *processes* in S that are involved in word formation; they do not refer to *word-formation types* in S. A word-formation type in S is a subrelation of word formation in S, which is a relation by (1e). Word-formation types are denoted by expressions like 'compounding in S', largely in agreement with linguistic tradition.

[4] In earlier versions of the theory, such as Lieb (2011/2012), I preferred a terminology based on the verb *build* to denote word-formation processes, while not excluding the *process* terminology. The *build*-terminology is now given up; it did not sufficiently prove itself in actual use.

(ii) *Concept names* (for lexical meanings) are formed by placing between raised dots an expression that is suggestive of the concept's intension, using number subscripts for differentiation. The concept may be left undetermined, as in the case of ·door·, or subsequently defined, like $\cdot\text{lock}_1\cdot$ (defined below, in (8)).

(iii) *Word paradigm names* are formed by means of a 'P'-superscript, with number subscripts for differentiation.

(iv) 'S', 'S_1', … each stand for arbitrary idiolect systems. The notion of idiolect system is understood as explained and defended in Lieb (1993a: Ch. 6); recourse to idiolect systems rather than to entire languages or varieties allows us to deal with linguistic variability, on the pattern of Lieb (1993a), and also allows us to integrate the theory of word formation into a general theory of language that applies to arbitrary languages. These two aspects, though basic, will not be discussed here any further. There is, however, an important consequence for the description of word formation in individual languages: it is variation-sensitive right from the start (below, Sec. 7).[5]

1.5 Word-formation statements: generalization

1.5 a. The format

Example (2) of a word-formation statement is representative; such statements have the following general form, where (a) is a reading of (b):

(3) a. $\langle P, b\rangle$ is formed in S from $\langle P_1, b_1\rangle$ and $\langle P_2, b_2\rangle$ through φ.
b. $\langle\langle P, b\rangle, \langle P_1, b_1\rangle, \langle P_2, b_2\rangle, \varphi\rangle \in \text{wf}(S)$.

Statements of this form are, strictly speaking, *specific*: they are no rule statements but statements on individual instances of word formation, and they are *primary*: they do not refer to *types* of word formation in S but to word formation in S as such.[6]

Whereas $\langle P, b\rangle$ must be a lexical word, $\langle P_1, b_1\rangle$ and $\langle P_2, b_2\rangle$ may be lexical words, (stem or affix) lexemes, or 'pseudo-words'. Word-groups (phrases) must be allowed to figure in word formation. However, such groups are not forms of word paradigms; no word paradigm is available when they turn up. Groups are accounted for by means of *pseudo-words*, certain constructs that are of the same formal type as lexical words but involve groups and group categories. (For an example of a pseudo-word, see (i) in Sec. 2.6 b, below.)

1.5 b. Functions involved in word-formation statements

By (1f), word formation per se is construed as a function. In a word-formation statement of the form (3b), this function is denoted by the term 'word formation in', or 'wf'. Formulations of the form (3a) are construed as readings of (3b)-formulations. (3) is in agreement with the way the word-formation function is characterized in (1e) and (1f).

---

[5] Relativization of terms to idiolect systems S may be left implicit in suitable contexts.

[6] Word-formation *rules* of a grammar—of a description of an individual language or language variety—are certain universally quantified sentences of the grammar: empirical statements on word-formation processes in the idiolect systems of the language or language variety. (See below, Sec. 7.3)

Functions φ in (3) are functions like *stem-form compounding in* S. *Stem-form compounding* in an absolute sense (the stem-form compounding process) is again a function: a function α that takes idiolect systems S as arguments and assigns to each S a function φ. For example, in the word-formation statements (2), α = stem-form compounding, and $\varphi = \alpha(S) = \alpha_S$ = stem-form compounding in S. Distinguishing between functions α and φ and interrelating them in this way is a fundamental step in solving a basic problem of long standing in linguistics: how to transcend, but still account for, the diversity of word-formation processes in the languages of the world when a general word-formation theory is to be formulated.

Functions α like stem-form compounding apply to individual idiolect systems S, not to pairs $\langle S, t\rangle$ where t is a time interval, nor to quadruples $\langle S, t, S_1, t_1\rangle$; similarly, time intervals do not figure in the arguments of functions $\alpha(S) = \varphi$, such as stem-form compounding in S. An extension of the theory would be needed to cover word formation as linguistic change in real time.

1.5 c. Accounting for lexical-word categories

Word-formation statements as characterized by (3) and exemplified in (2) do not yet mention any categories to which the newly formed lexical word $\langle P, b\rangle$ belongs, such as its part of speech. What is the reason for the apparent omission?

There are two aspects of word formation in a traditional sense: contributing to the identification of non-basic lexical words, and partially identifying the 'place' such a word has in the system of lexical-word categories (part-of-speech categories, valency categories, etc.). The *place* may be construed as the set of lexical-word categories to which the word belongs.

The two aspects are separated in the present theory. Word-formation processes φ help identify lexical words $\langle P, b\rangle$ without specifying lexical-word categories to which $\langle P, b\rangle$ belongs; as a result, we eventually arrive at the relation wf(S), or word formation in S. Given this relation, we account for the second aspect of word formation in a traditional sense, partial identification of place. The place of a non-basic lexical word in the word-category system may partly depend on the way it has been identified, but such determination of place should be treated as a consequence not a part of the identification process. The 'place' aspect of word formation is therefore accounted for by a separate function of word-category assignment based on the relation wf(S), or word formation in S. Since word formation in S also involves the formation of stem lexemes in S, we may assume a single function of *lexical-category assignment in* S, with *word-category assignment in* S and *stem-category assignment in* S as subfunctions. (See below, Secs 3.3 b and 6.5 d).

The present essay is mainly devoted to clarifying the nature of functions α and φ, both generally and with respect to specific functions. Before taking this up, some background will be introduced; in particular, the Word and Paradigm model that plays an essential role will be explained. The reader may, however, go directly to Sec. 3 for a first impression of how α and φ are conceived, returning to the explanations of background in Sec. 2 for a more detailed understanding.

2 Background: the WLP model

2.1 Words, lexemes, paradigms

2.1 a. Words and word paradigms

A *lexical word* of S is to be a pair ⟨P, b⟩ consisting of a word paradigm P and a concept b. A *word paradigm* P is conceived as a set of pairs ⟨f, J⟩, where f is a syntactic unit and J a set of syntactic categories K such that f is an element of each K in J. Adopting the view of noun paradigms proposed in Lieb (2005) for English, the following is an example, incompletely listed, of a noun paradigm in S (where S is any suitable English idiolect system):

(4) $lock_1^P$(S) =
{⟨$lock_1$, {$Unsp_{Case}$(-, S), Sg_N(-, S), $Unsp_{Def}$(-, S)}⟩,
⟨$locks_1$, {$Unsp_{Case}$(-, S), Pl_N(-, S), $Unsp_{Def}$(-, S)}⟩,
⟨the_1 $lock_2$, {$Unsp_{Case}$(-, S), Sg_N(-, S), Def(-, S)}⟩,
⟨the_1 $locks_2$, {$Unsp_{Case}$(-, S), Pl_N(-, S), Def(-, S)}⟩,
⋮
…}

'$lock_1^P$(S)' is to be read as: '*P*aradigm *1* for *lock* in S'. (Paradigm numbering is arbitrary.) In (4), subscripts '1', '2' etc. in the names of individual forms indicate that we are dealing with the first, second etc. member of a sequence—there may be only *one* member. The unlisted entries—indicated by dots—consist of the forms a_1 $lock_2$, $some_1$ $lock_2$, $some_1$ $locks_2$, any_1 $lock_2$, any_1 $locks_2$, no_1 $lock_2$, and no_1 $locks_2$, each one with an associated category set; $some_1$, any_1, and no_1 are 'unstressed'. It is assumed here that English nouns may have both synthetic forms, like $lock_1$, and analytic forms, like the_1 $lock_2$.

'$Unsp_{Case}$' is short for 'Unspecific for Case', and '$Unsp_{Def}$' is short for 'Unspecific for Definiteness'. An expression like '$Unsp_{Case}$(-, S)' reads: 'the set of all f such that f is Unspecific for Case in S'; the expression denotes a syntactic category of S consisting of forms of lexical words (nouns). The part '(-, S)' in category names may be omitted when it can be restored from context, same as '(S)' in an expression like '$lock_1^P$(S)'. [7]

---

[7] In English examples, assumptions on noun-stem and noun paradigms—their forms and categorizations—are taken over, without further justification, from Lieb (2005), but a classification of Noun forms (analogously, Noun-Stem forms) is added: into forms to which criteria for Case, Number, or Definiteness do not apply ($Neut_N$(-, S): Neutral Noun-form in S—most Adjective forms), and forms to which they do apply, if ambiguously ($Non\text{-}Neut_N$(-, S): Non-Neutral Noun-form in S); the two classes are allowed to overlap. If the criteria, say for Case, apply ambiguously, we obtain a category such as $Unsp_{Case}$(-, S) as a subset of $Non\text{-}Neut_N$(-, S). This exemplifies how I currently propose to deal with syncretism.—Categorizations of verb-stem forms and verb forms are based on a detailed analysis of the underlying systems, not yet published.

2.1 b. Comments

Ontologically, word paradigms in our sense are equivalent to paradigms as construed in the realizational WP model of Stump (2001: 43). The above conception was, however, developed earlier, in Lieb (1976), (1980a).[8]

The notion of word paradigm and lexeme paradigm is generalized by also allowing 'improper paradigms' in addition to the usual 'proper' ones: an *improper paradigm* P is a paradigm such that for some K and each ⟨f, J⟩ in P, J = {K}. All affix paradigms are improper ones. For an example of an affix paradigm, see (5c), below.

Word paradigms are *syntactic* due to the fact that their elements ⟨f, J⟩ each consist of (i) a syntactic unit f (obvious if f is an analytic form like the_1 $lock_2$, but also true of a synthetic form like $lock_1$, see below, Sec. 2.2), and (ii) a set J of categories K each of which is a set of syntactic units, as in (4).

The WLP model agrees with 'abstractive' word-based morphology as characterized in Blevins (2006) by taking phonological words as objects that are given independently of any morphological analysis to which they may or may not be subjected. 'Word-based morphology' agrees with tradition in assigning an essential role to 'principal forms' of a word paradigm, forms f of a word paradigm that may be used, together with a categorization J of f, to characterize the paradigm as a whole, due to implicational relationships that exist between these pairs ⟨f, J⟩ and all other elements of the paradigm. It is, however, by reference to morphological analyses that the implicational relationships are best stated (as they are in traditional lexicology when paradigms of words in a word lexicon are characterized), which takes us back to stems and affixes (see also Baerman & Corbett 2012: 55-56, on stem forms as implicitly assumed in Blevins 2006).

2.1 c. Lexemes and lexeme paradigms

The term 'lexeme' as used here does not apply to lexical words; thus, its use is more restricted than in part of the literature. On the other hand, it is also broader: the term allows not only *stem lexemes* (stem-form lexemes) but also *affix lexemes* (affix-form lexemes), which is an extension of traditional usage.

Lexemes are pairs ⟨P, b⟩ of a morphological paradigm P and a concept b. Stem lexemes in morphology are formally analogous to lexical words in syntax, and stem paradigms are analogous to word paradigms. As a rule, there is exactly one lexical word correlated with a given stem lexeme, *the lexical word for the stem lexeme*, and there is exactly one stem lexeme for a (non-idiomatic) lexical word, *the stem lexeme of the word*; stem lexemes for which there is no lexical word in the idiolect system will be called *trapped*.

Of the following *three lexeme paradigms* the first two are stem paradigms, the first is a noun-stem paradigm and the second a verb-stem paradigm. The third paradigm is an affix paradigm (superscripted 'LP' for 'lexeme paradigm'; 'S' and '(-, S)' to be added, see (4)):

(5) a. $lock_1^{LP}$ =
{⟨$lock_1$, {$Unsp_{Case}$-St, Sg_N-St, $Unsp_{Def}$-St}⟩,
⟨$lock_1$ s_2, {$Unsp_{Case}$-St, Pl_N-St, $Unsp_{Def}$-St}⟩}

---

[8] The Lieb and Stump versions both correspond to the notion 'arrangement of a paradigm' in Matthews (1965), in reverse order: Matthews has ⟨J, f⟩ instead of ⟨f, J⟩. Stump has modified his conception of paradigms in later work, since Stump (2002). Construing a paradigm simply as a *set* of forms, as still done in Anderson (1992: 134), can be shown to be inadequate; cf. Lieb (2005: 1621).

b. $lock_2{}^{LP}$ =
{⟨*lock*$_1$, {Inf-St, Pres-St}⟩,
⟨*lock*$_1$, {Unsp$_{Pers}$-St, Unsp$_{VN}$-St, Unsp$_{Mood}$-St, Pres-St}⟩,
⟨*lock*$_1$ *ing*$_2$, {Part-St, Pres-St}⟩,
⟨*lock*$_1$ *ed*$_2$, {Unsp$_{Pers}$-St, Unsp$_{VN}$-St, Unsp$_{Mood}$-St, Pret-St}⟩,
⟨*lock*$_1$ *ed*$_2$, {Part-St, Pret-St}⟩}

c. $s_1{}^{LP}$ =
{⟨*s*$_1$, {Af}⟩,
⟨*z*$_1$, {Af}⟩,
⟨*ez*$_1$, {Af}⟩}

Remarks on (5)

s$_2$ in (a) represents 'inherent inflection', not 'contextual inflection', by the well-known distinction advocated especially by Booij (e. g., Booij 2007: 104), and this is also true of *ing*$_2$ and *ed*$_2$ in (b). The position is taken here that *inherent inflection* without contextual inflection *results* (in morphology) *in stem forms*, which are analytic in case of affixal inflection. The respective forms in (a) and (b) are *contextually uninflected*.

Adding a form of the only Person affix in English to an appropriate verb-stem form creates not a stem form but an *inflected morphological word*. Therefore, *lock*$_1$ *s*$_2$ does not appear in the verbal stem-form paradigm in (b).

The affix paradigm in (c) is an improper morphological paradigm with three different forms; depending on one's view of inflection and the morphology-phonology interface, the three forms may or may not be reduced to one. All affix forms are assumed to have the categorization {Af}, which makes all affix paradigms improper ones. Functional distinctions that might be associated with affix *forms* are reconstructed as distinctions between complete affixes, i. e., between pairs ⟨P, b⟩ (construed as below, in Sec. 2.4 c). For example, there is just one Person affix in (Standard) English idiolect systems, with P as in (5c).

2.1 d. 'Stem alternation'

As appears from (5a) and (5b), the conception of stem lexemes developed in Integrational Linguistics since Lieb (1976) and adopted here in a modified version, differs from two current ways of dealing with 'stem allomorphy': *first*, differs from an approach that proceeds from single 'roots' but allows for roots to be changed by 'readjustment rules' to account for traditional stem allomorphy (e. g., Embick & Halle 2005, convincingly criticized in Aronoff 2012: 39-47); *second*, differs from an approach that does allow several stem forms for a single 'lexical item' but treats stem forms quite generally as 'morphomic' (Spencer 2012), in a sense going back to Aronoff (1994) as "stems that are pure forms and which are not the realization of any feature or property set" (Spencer 2012: 88; Spencer specifically claims that "*killed* is a morphomic stem", 2012: 99). The Integrational conception combines some features of the two approaches while rejecting others: several stem forms are allowed in a stem paradigm but each may be assigned a set, *or several sets*, of 'morphosyntactic (stem-form) properties'; in addition, each may have 'versions' that are not themselves stem forms but simply arise from applying form-change functions that are part of an inflection or word-formation process; these are indeed 'morphomic'.

As an example, consider German *täg*$_1$ as part of the stem form *täg*$_1$ *lich*$_2$ ·daily·. The stem form is obtained by stem-form derivation from *tag*$_1$ (no umlaut) ·day· and the suffix form *lich*$_1$. While *tag*$_1$ is a form of the stem paradigm *tag*^{LP}, this is not true of *täg*$_1$, arising from *tag*$_1$ by applying umlauting as part of the derivation process that adds *lich*$_1$ to *tag*$_1$; as a matter of fact, *täg*$_1$ is not a form of any stem paradigm but exists only as a version of *tag*$_1$, a version occurring in stem forms due to umlauting as part of a derivation process: *täg*$_1$ is a strictly morphomic version of the stem form *tag*$_1$. It is an empirical question how many different forms must be allowed for a given stem paradigm, and how many different versions for a given stem form.

2.1 e. Process-related forms and categories

The above account of *täg*$_1$ presupposes an Item-and-Process framework; it is not possible when Item and Arrangement is used, as it has been in Integrational Morphology ever since Lieb (1976), jointly with a Word-and-Paradigm approach. In such a framework we may proceed as follows.

We enrich the stem paradigm *tag*^{LP} by including a pair ⟨*täg*$_1$, {K, …}⟩, where K is a category identified either with (i) *lich/ig*, the set of stem forms that combine with the suffix forms *lich*$_1$ or *ig*$_1$ (exhibiting umlaut whenever possible), or else with (ii) Derivational Noun-stem form (Der-NSt), the set of Noun-stem forms that occur (possibly: occur only) with forms of derivational affixes. Whereas (i) was chosen in Lieb (1976: 31), (ii) is preferred in later IL work, as in Eisenberg (1998: 213), referring to Fuhrhop (1998). Let us call *derivation form* any f in an element ⟨f, J⟩ of a stem or word paradigm such that J contains a category Der-XSt or Der-Xform, where X = 'Noun', 'Verb', or 'Particle'; analogously, *compounding form*. A notion of compounding form was first introduced in Integrational Linguistics in Lieb (1976: 30); such notions were then applied by Eisenberg (1998) and, notably, Fuhrhop (1998), who assumes stem paradigms that include derivation stem forms (*Derivationsstammformen*), occurring only in a context of derivation, and compounding stem forms (*Kompositionsstammformen*), occurring only in a context of compounding.

Nolda (2012b), who uses an IP not an IA approach, introduces 'conversion forms' (foreshadowed by *Umwandlungsform* in Fuhrhop 1998) into stem paradigms, in addition to 'compounding forms' and 'derivation forms', categorizing all such forms by process-related categories: 'Noun-Stem compounding form in S', 'Verb-Stem conversion form in S', etc. (similarly, Barz 2005: 660-661, for German; Nolda, who refers to a later version of Barz's article, traces the assumption of process-related forms and categories as far back as Bloomfield 1933: 225-226).

Without conversion forms and their categories, Nolda's treatment of conversion processes in S would collapse, in the sense of no longer applying in cases where it is meant to apply. But should we really adopt, on an IP approach, process-related forms and categories for paradigms?

2.1 f. Excluding process-related forms and categories

Nolda's main argument in favour of *conversion* categories and forms in German paradigms is this: relevant forms are 'idiosyncratic' and therefore belong 'in the lexicon', where, it is claimed, they cannot be categorized in the usual way. But this can be turned around (and equally applied to 'compounding forms' and 'derivation forms'): the forms tend to be 'idiosyncratic' because they are introduced in the course of applying *word-formation processes*, where they may be introduced in an irregular fashion; therefore, the forms should not be represented in the lexicon in word or stem paradigms, or, if they appear, they should not appear with process-related categories.

Indeed, 'compounding forms', 'derivation forms', and 'conversion forms' arise in word-formation processes either as uncategorized parts of result forms, parts to be treated as purely morphomic that do not themselves figure in any paradigm: *täg*$_1$, above; or they arise as result forms that *are* paradigm forms but are categorized without reference to processes. The form *song*$_1$ as a Noun-Stem form is obtained by stem-form conversion from *sing*$_1$ as a Verb-Stem form but is to be categorized solely as a Singular Noun-Stem form, not as a conversion form, too. It would be utterly artificial to also assume *song*$_1$ as a Verb-Stem conversion form, introduced into the *sing*$_1$ Verb-Stem paradigm for a single purpose: to guarantee that stem-form conversion (now applied to *song*$_1$ as a Verb-Stem form) never introduces a segmental change. Indeed, it is the requirement of no segmental change, definitional on Nolda's account of conversion processes but rejected in the Process Model, that strongly suggests the retention of 'conversion' categories like 'Noun-Stem conversion form in S'.

In conclusion, process-related categories are not needed: either there are no relevant paradigm forms, or relevant paradigm forms are categorized without reference to processes. Phenomena of idiosyncracy to be observed in this area are a natural outcome of applying word-formation (or inflection) processes, and are not to be anchored in paradigms by introducing forms with process-related categories, artificially inflating the paradigms.

In contrast to my earlier position, then, I no longer assume categories like 'compounding form in S', 'derivation form in S', or 'conversion form in S' (relativized to Noun forms, Verb forms, Particle forms, or their stem forms). These may be needed on an IA approach to the formation of lexical words but are foreign to a framework that is consistently IP, making form-change functions available in all word-formation processes.

We next consider paradigm forms from an ontological point of view: what kind of entities are they?

2.2 Paradigm forms

Number subscripts such as the ones in (4) and (5) indicate that the *forms* f of a paradigm (its first-place members) are *sequences*, including *unit sequences*; more generally, this is to be true of any morphological or syntactic unit. A (non-empty) *sequence* is a one-place function whose arguments are the numbers 1, …, n, for some $n \geq 1$. Since a one-place function is a two-place relation, or set of ordered pairs, the paradigm forms f in (4) and (5) are sets of ordered pairs ⟨n, w⟩ where n is a positive integer and w is as follows. If f is a lexical-word form—or, generally, a *syntactic unit*—then w is a *phonological word*; if f is a lexeme form—or, generally, a *morphological unit*—then w is a *morph or a phonological word* (some member w of f must be a morph, though). For example:

(6) a. *the*$_1$ *lock*$_2$ = {⟨1, *the*⟩, ⟨2, *lock*⟩}
b. *lock*$_1$ = {⟨1, *lock*⟩} (= the unit sequence of *lock*, different from *lock* itself)

In (a) but not in (b) the number subscripts may be omitted: *the lock* = *the*$_1$ *lock*$_2$; but *lock*$_1$ ≠ *lock*.

Construing paradigm forms as sequences in this sense is important. Such a construal links with a standard notion of *concatenation*, symbolically: $^\frown$, as a purely formal operation on pairs of sequences. We also admit *the empty sequence* f^0 = the empty set Ø. This is the *identity element* for concatenation: for any f it is true that $f {}^\frown f^0 = f^0 {}^\frown f = f$. The empty sequence is neither a morphological nor a syntactic unit, it is a purely formal entity. Still, the empty sequence, being identical to the empty set, is a *part* (subset) of any unit: morphological and syntactic units are sets (of pairs), and the empty set is a subset of any set. Generally, we must distinguish between: a sequence, *the*$_1$ *lock*$_2$; its *parts* or subsets, such as *lock*$_2$ = {⟨2, *lock*⟩}; its *elements*, such as ⟨2, *lock*⟩; its *members*, such as *lock*; and its *positional variants*, such as *the*$_7$ *lock*$_{10}$ = {⟨7, *the*⟩, ⟨10, *lock*⟩} (any sequence is an improper positional variant of itself).

The forms both of lexical words and of lexemes are non-empty sequences f, they are entities of the same type. This is due to the fact that phonological words, the members of syntactic units, and morphs, the morphological members of morphological units, both are entities w, i. e. triples $\langle f^P, k^P, I^P \rangle$ consisting of a sound sequence f^P, a constituent structure k^P of the sound sequence, and an intonation structure I^P of the sound sequence, as exemplified in (7).

(7)

```
                              VcGr
                        ┌──────┼──────┐
                        C      V      C
                        1      2      3
lock = /'lɒk./ =  /     l      ɒ      k  /
                               H
```

= ⟨/lɒk/, {⟨{1}, C⟩, ⟨{2}, V⟩, ⟨{3}, C⟩, ⟨{1, 2, 3}, VcGr⟩}, {⟨1, {H}⟩}⟩

'VcGr' for 'Vocalic Group'; 'H' for 'High': in English, high pitch is an indicator of primary word stress, assuming a pitch accent theory for English word stress (at the phonological level, only pitches need be considered for English word stress); I^P is the unit sequence—because there is only *one* syllable—of the unit set of High. (See Lieb 2008 for the underlying phonological theory, which distinguishes just two levels of a phonological system: a phonological one, to which morphs and phonological words belong, and a narrowly phonetic one.)

lock happens to be both a morph and a phonological word, in contrast to *locks*, which is only a phonological word, and *s*, which is only a morph.[9]

2.3 Lexical meanings

2.3 a. Examples of concepts (1): the concept of a lock

Lexical meanings—of morphological or syntactic units, such as paradigm forms, and of paradigms—are to be concepts, in a specific sense. Very briefly, concepts are content-related properties of perceptions or conceptions. A non-empty concept—in contrast to the

[9] For the phonological facts of English, I follow, as much as possible, the entries for British English ('BBC English', formerly 'Received Pronunciation') in the *English Pronouncing Dictionary*, 16th edn (Roach et al. eds 2003). However, extrasyllabic consonants will be allowed, and a notion of secondary word stress is presupposed that is functional ('possibility of asemantic prominence'), rather than phonetically defined, which may lead to deviations from the *Dictionary*. There are notorious problems in the area of English word-stress patterns, especially with respect to compounding, where they have been under discussion for some forty years (see Plag & Kunter 2010: Sec. 2, for a short overview; Giegerich 2010 for incisive criticism of earlier claims; see also Bell & Plag 2012 and the literature discussed there).

empty concept (below, Sec. 2.3 c)—is the property of being a perception or conception that contains a certain set of attributes in its content. The attributes are properties of or relations between real-world entities. The set of attributes is the *intension* of the concept, the set of entities that has the attributes is the concept *extension*.[10]

A first example is provided by the concept that may be assumed for *the lock* in *the lock of the door*. Informally, this is the concept of being an appliance x, to be activated using a potentially private means (some sort of a key), such that the *purpose* of x (or its *function*: 'purpose' is preferred to avoid ambiguity) is as follows (it is definitional for appliances to *have* a purpose): when x is activated by some agent x_2, the action x_1 of activating x has the immediate result of creating a fixed but releasable connection between a movable object x_3 (such as a door) and some other object x_5 (e. g., the doorframe), thus barring access, at least partially, to an object x_4 to which access is possible.

A *purpose* may be construed as a *relation-in-intension* between an action x_1, actor x_2, and additional objects; equivalently, as an *attribute* of tuples $\langle x_1, x_2, \ldots\rangle$. The concept of a lock, ·lock$_1$·, may then be formally determined as follows:[11]

(8) ·lock$_1$· $=_{df}$ the property of being a perception or conception z such that {LOCK$_1$} $\subseteq$ the content of z, where:

LOCK$_1$ $=_{df}$ the property [one-place attribute] of being an x such that:

a. x is a (mechanical or electronic) appliance;

b. x is meant to be activated through a specific means that can be kept private;

c. the purpose (or function) of x = the relation-in-intension between any x_1, x_2, x_3, and x_4 [the four-place attribute of any $\langle x_1, x_2, x_3, x_4\rangle$] such that:

(i) x_1 is an action by x_2 of activating x;

(ii) x_3 is a movable object;

(iii) the immediate result of x_1 = creation, through x_1, of a fixed but releasable connection between x_3 and some x_5, with the effect of x_3 barring some access to x_4.

In (c.iii), the *immediate result* of x_1 may be construed as a state-of-affairs. We may also have accidental activation, say, through an inanimate agent, which may have an immediate result as in (c.iii); this does not correspond to the purpose of the lock.

The concept ·lock$_1$· is *one-place* because LOCK$_1$ is a one-place attribute, or property. The (one-place) *intension*, or i1, of ·lock$_1$· is {LOCK$_1$}, and the (one-place) *exten-*

---

[10] For the presupposed theory of lexical meanings, see Lieb (1983: Ch. 13), and Lieb (1992a). Any definition of a concept name becomes part of an empirical hypothesis when the concept is claimed to be a lexical meaning of a morphological or syntactic unit or paradigm. Whenever a concept name is defined in this essay to be used in lexical-meaning claims on English, the definitions were carefully checked against both major dictionaries and examples of actual use as found on the internet. Some definitions may appear to be overly precise and circumstantial but their details are needed to distinguish related concepts from one another and for properly interrelating concepts in word-formation processes by meaning-change functions. This is specifically true of the concept of a lock in (8) as underlying the concept of locking in (9), where distinctions such as the one between a lock and a bolt are relevant. (I would like to draw the reader's attention to an interrogation episode in a recent detective novel: Anne Perry, *Midnight at Marble Arch*, London: Headline Publishing Group, 2012, 40-42, where it is exactly the concepts of a lock and locking as defined in (8) and (9), with details as in (8), that appear to be relevant—a passage I hit upon only *after* defining the concepts.)

[11] Concept ·lock$_1$· will be the basic concept in the examples of stem-form compounding (Sec. 3.4) and stem-form conversion (Sec. 4.3).

sion, or e1, of $\cdot lock_1\cdot$ is the set of all x such that x has $LOCK_1 = \{x \mid x \text{ has } LOCK_1\}$ = the set of all locks.—A superscript notation is generally used for naming intensions and extensions: $^{i1}\cdot lock_1\cdot$, $^{e1}\cdot lock_1\cdot$. Names of attributes are in capitals.

2.3 b. Examples of concepts (2): two concepts of locking[12]

From $\cdot lock_1\cdot$ in (8) we obtain the verbal concept $\cdot lock_2\cdot$ in (9), which is three-place:

(9) $\cdot lock_2\cdot =_{df}$ the property of being a perception or conception z such that $\{LOCK_2\} \subseteq$ the content of z, where:

$LOCK_2 =_{df}$ the three-place attribute [relation-in-intension] of being an $\langle x_1, x_2, x_3\rangle$ such that: for some x and x_4,

a. $x \in {}^{e1}\cdot lock_1\cdot$;

b. $\langle x_1, x_2, x_3, x_4\rangle$ satisfies the purpose of x.

(‘satisfies’ in (b) means ‘has’: the purpose of x is a four-place attribute.) The intension and extension of $\cdot lock_2\cdot$ are, in superscript notation: $^{i3}\cdot lock_2\cdot = \{LOCK_2\}$; $^{e3}\cdot lock_2\cdot = \{\langle x_1, x_2, x_3\rangle \mid \langle x_1, x_2, x_3\rangle \text{ has } LOCK_2\}$.—Concept $\cdot lock_2\cdot$ is used in sentences like:

(i) *My friend locked the door*.

It is *not* used in:

(ii) *My friend locked the lock / the padlock.*
(iii) *My friend locked the chain*.
(iv) *My friend locked the house*.

Different if related concepts must be assumed for *locked*$_3$ in these sentences.

In particular, the meaning of *locked*$_3$ in (iv) is a concept $\cdot lock_3\cdot$, also obtained from $\cdot lock_1\cdot$, the concept of a lock (surprisingly, $\cdot lock_3\cdot$ cannot be obtained directly from $\cdot lock_2\cdot$, the concept of locking a door etc.). Roughly, $\cdot lock_3\cdot$ is the concept of satisfying the purpose of any ‘relevant’ lock (requiring that there should be such locks):

(10) $\cdot lock_3\cdot =_{df}$ the property of being a perception or conception z such that $\{LOCK_3\} \subseteq$ the content of z, where:

$LOCK_3 =_{df}$ the three-place attribute of being an $\langle x_1, x_2, x_3\rangle$ such that:

a. for some x and x_3',
   (i) $x \in {}^{e1}\cdot lock_1\cdot$,
   (ii) it is possible that for some x_1' and x_2', $\langle x_1', x_2', x_3', x_3\rangle$ satisfies the purpose of x;

b. for any x and x_3' as in (a), $\langle x_1, x_2, x_3', x_3\rangle$ satisfies the purpose of x.

x_3' in (10) (e. g., a door) corresponds to x_3 in (8), and x_3 in (10) (e. g., a house) corresponds to x_4 in (8). There may be many ‘doors’ x_3' into the house x_3, each one with its own lock

[12] The two concepts figure in later examples of word-formation processes as follows: $\cdot lock_2\cdot$ as a basic concept in stem-form derivation (Sec. 4.1) and word-form derivation (Sec. 4.2) and as a result concept in stem-form conversion (Sec. 4.3); $\cdot lock_3\cdot$ as a basic concept in word-form compounding (Sec. 3.5).

or locks x, but there is only *one* action x_1 by x_2 of activating all of them (this allows for successive activation).

2.3 c. The empty concept

For a general conception of lexical meanings as concepts, we also need *the empty concept*, defined as follows:

(11) b^0 (= *the empty concept*) $=_{df}$ the property of being a perception or conception z such that the content of z is empty.

There is *no* z with this property; b^0 is thus analogous to the empty set in set theory. b^0 is the only *zero-place* concept; the notions of intension and extension do not apply.

The empty concept, a well-defined non-linguistic entity, plays an important role in our conception of word-formation processes: it is by reference to the empty concept that the 'process cline' is established.

2.3 d. Concepts as meanings

The concept ·lock_1· in (8) is one meaning of the noun-stem paradigm $\textit{lock}_1^{\text{LP}}$ in (5a) and its forms. The two concepts ·lock_2· in (9) and ·lock_3· in (10) are different meanings of the verb-stem paradigm $\textit{lock}_2^{\text{LP}}$ in (5b) and its forms. Each concept is also a meaning of a corresponding word paradigm and its forms. We thus arrive at two different lexical words—two verbs: ⟨$\textit{lock}_2^{\text{P}}$, ·$\text{lock}_2$·⟩ and ⟨$\textit{lock}_2^{\text{P}}$, ·$\text{lock}_3$·⟩, with identical paradigms but different, if closely related, meanings. Assuming a single word in such cases, with a more general concept as a meaning, and re-introducing the more specific concepts as 'senses' or 'meaning shades', is inadequate for word formation understood as the partial identification of lexical words. For example, it appears from Plank (2010) that we would still have to treat the 'senses' as basic in conversion processes.

The affix paradigm s_1^{LP} in (5c) and its three forms have the empty concept b^0 as their only lexical meaning, which is true of affix paradigms and their forms in general: semantic differences between affixes are not accounted for by different lexical meanings.[13]

2.4 Paradigm bases: some essentials

2.4 a. General remarks

It is assumed that the *lexicon* of an idiolect system S has at least two parts, the *word lexicon* of S (see above, Sec 1.2 b) and the *lexeme lexicon* of S, consisting of the *stem lexicon* of S, the set of stem lexemes ⟨P, b⟩ of S, and the *affix lexicon* of S, the set of affixes (affix lexemes) of S.

The lexical words ⟨P, b⟩ in the non-basic word lexicon, and also analogous stem lexemes in the stem lexicon, are to be identified by means of word-formation processes and inflection processes. These must be supplemented by certain other components of the idio-

[13] Using the empty concept as a lexical meaning of affix forms and affix paradigms allows us to treat affixes as *lexical* without blurring the semantic difference between stem forms and stem paradigms on the one hand (typically, their lexical meanings are non-empty, and no function can be associated with them as a grammatical meaning), and affix forms and affix paradigms on the other (their lexical meanings are empty, and functions can be associated with them as grammatical meanings). See Sec. 2.4 c, below, for discussion.

lect system: *syntactic paradigm bases* for lexical words and *morphological paradigm bases* for stem lexemes; there are no paradigm bases for affixes ⟨P, b⟩. Paradigm bases are theoretical constructs too complex to be characterized here in any detail; the account given in Lieb (2005) (modified in some respects, yet to be mentioned) is being presupposed. There are some features of paradigm bases, though, that are of immediate importance to the present conception of word formation. The features may be explained by going back to some of the examples, again presupposing English idiolect systems S.

2.4 b. Stem lexemes and lexical words

Noun-stem lexemes and nouns are chosen for exemplification due to their simplicity in English; verb-stem lexemes as in (5b) and verbs might have been more convincing but are too complex for initial exemplification. Consider (note the superscripts):

(12) a. $\langle P_1, b\rangle = \langle \mathit{lock}_1^{LP}, \cdot\mathrm{lock}_1\cdot\rangle$, where lock_1^{LP} is as in (5a) and $\cdot\mathrm{lock}_1\cdot$ as in (8).
b. $\langle P_2, b\rangle = \langle \mathit{lock}_1^{P}, \cdot\mathrm{lock}_1\cdot\rangle$, where lock_1^{P} is as in (4).

$\langle P_1, b\rangle$ is a stem lexeme, and P_1 a stem paradigm. The underlying paradigm base for noun-stem paradigms allows us to identify the following pairs for the stem paradigm:

(i) $\langle f_1, J_1\rangle = \langle \mathit{lock}_1, \{\mathrm{Unsp}_{Case}\text{-St}, \mathrm{Sg}_N\text{-St}, \mathrm{Unsp}_{Def}\text{-St}\}\rangle$, and
(ii) $\langle f_2, J_2\rangle = \langle \mathit{lock}_1\ s_2, \{\mathrm{Unsp}_{Case}\text{-St}, \mathrm{Pl}_N\text{-St}, \mathrm{Unsp}_{Def}\text{-St}\}\rangle$

These are the only elements of the stem paradigm P_1.[14] $b = \cdot\mathrm{lock}_1\cdot$ is identified as a meaning of P_1 using the fact that b is a lexical meaning of f_1 given J_1, and of f_2 given J_2. Similarly, the pairs:

(iii) $\langle f_3, J_3\rangle = \langle \mathit{lock}_1, \{\mathrm{Unsp}_{Case}, \mathrm{Sg}_N, \mathrm{Unsp}_{Def}\}\rangle$ and
(iv) $\langle f_4, J_4\rangle = \langle \mathit{the}_1\ \mathit{lock}_2, \{\mathrm{Unsp}_{Case}, \mathrm{Sg}_N, \mathrm{Def}\}\rangle$

are identified by the underlying paradigm base for noun paradigms as some of the elements of P_2; and $b = \cdot\mathrm{lock}_1\cdot$ again is identified as a meaning of P_2.

Any paradigm base contains a component that is a classification system on the basic set of the paradigm base, such as the set of noun-stem forms, or the set of noun forms. This component, called *the functional system*, supplies the categories that may be elements of the set J in an element ⟨f, J⟩ of a paradigm P. Thus J_1 in (i) and J_2 in (ii) are determined using categories provided by the functional system of the noun-*stem* paradigm base, and J_3 in (iii) and J_4 in (iv) are determined using categories supplied by the functional system of the noun-*form* paradigm base.

The noun-stem lexeme $\langle \mathit{lock}_1^{LP}, \cdot\mathrm{lock}_1\cdot\rangle$ in (12a) is *the stem lexeme of* the lexical word $\langle \mathit{lock}_1^{P}, \cdot\mathrm{lock}_1\cdot\rangle$ in (12b). However, the function 'stem-lexeme-of' is not basic in the WLP model; forms of word paradigms are not construed—as in realizational WP models since Matthews (1972)—as forms that 'realize' certain sets of categories given a 'stem': the WLP model is not realizational.[15]

---

[14] For $\mathit{lock}_1\ s_2$ as a stem form, see above, *Remarks on (5).*

[15] Generally, the WLP model does not easily fit the classification of 'theories' proposed by Stump (2001: Sec. 1.1).

2.4 c. Affix lexemes

In contrast to Lieb (2005), paradigm bases are not assumed for affixes (affix lexemes ⟨P, b⟩), due to the fact that in addition to their shared lexical meaning b^0 all affix forms now also agree in their categorization, which is {Af}, as in Lieb (1983). The only remaining problem consists in determining the affix forms that go into the same affix paradigm.

In the case of stem paradigms and word paradigms, we rely on the lexical meanings b of forms to collect pairs ⟨f, J⟩ into a single paradigm. Similarly, *sameness of 'grammatical meaning'* is one criterion when pairs ⟨f, J⟩ are to be assigned to the same affix paradigm. Two kinds of grammatical meaning are distinguished in this context:

(v) category-change functions δ that assign sets of functional morphological categories to sets of morphological categories, sets of syntactic categories, or sets of 'pseudo-categories' (in the sense explained below, in Sec. 2.6 b)

(vi) meaning-change functions ε that assign concepts that are lexical meanings of forms to (pairs of) other such concepts

Identity functions are allowed as a limiting case in both (v) and (vi).

In speaking of 'grammatical meanings' I do not assume 'signs', represented independently in the idiolect system, that consist of an affix form or an affix paradigm plus a function δ or ε or both; rather, affix forms (therefore, affix paradigms) are directly related to functions δ and ε by appearing with them in (arguments of) word-formation or inflection processes.

Attempts to make affix forms into the *signifiants* of Sausurean signs by associating with them 'grammatical meanings' as their *signifiés* not only run into trouble because of a multitude of 'grammatical meanings' that may be associated with a single affix form, creating problems, quite unnecessarily, of meaning shade vs. meaning difference. Such attempts also run afoul of the fact that the form/meaning association is as a rule conditional on affix-external factors; the *signifiant* should be the affix form *together with* these factors, but this takes us well outside the domain of sensible notions of 'sign'. Our approach avoids such problems: a given affix form plus one 'grammatical meaning' of either type (function δ, function ε) plus a relevant tuple of conditioning factors are jointly represented in a single argument of a word-formation or inflection process in S, which does not imply that the affix form together with the 'grammatical meaning(s)' constitutes a Sausurean sign. This is in agreement with most of the morphological literature.[16]

Affix *lexemes* as here conceived are no Saussurean signs either. True, they are construed as paradigm/concept pairs, the concept being allowed as a lexical meaning of the paradigm; but such pairs would not be signs in any ordinary sense: 'empty' concepts and improper paradigms are not admitted for such signs. Construing affixes on the formal pattern of stem lexemes has the advantage that a single set of lexemes may be the origin of all morphological classifications that are not restricted to individual forms. Such classifications then also yield classes of affix *lexemes* that correspond to traditional classes of affix *forms*. The classification criteria are essentially provided by the way the forms of affix lexemes are related to functions δ and ε in word-formation and inflection processes.

---

[16] Vehement objections to a Saussurean-type sign conception of affix forms were prominently raised by Beard (1995: Ch. 2) in arguing against what he called 'the Lexical Morpheme Hypothesis' (1995: 6); affix forms, but not stem forms, are excluded as 'lexical' by Beard. Some approaches reject a sign status even for stem forms; see Sec. 2.1 d, above, for conceptions of stem forms as uniformly 'morphomic'.

Derivational affixes differ from *inflectional affixes* with respect to these functions. In particular, ε may have to be excluded as an identity function in the case of a derivational-affix form, but must be an identity function when we are dealing with an inflectional-affix form—inflection does not change lexical meaning.[17]

Derivational affixes may differ from each other with respect to the type of functions δ and ε that are correlated with their forms. It is by such differences, especially in relation to meaning-change functions ε, that the area of so-called affixoids (and analogous derivational particles) can be reconstructed: 'affix' is construed as an either-or notion but this does not preclude an affixoid scale or cline to be defined on the set of derivational affixes, or on the set of their forms (similarly, Boye and Harder 2012: 6, 30-32, on the distinction between 'grammatical expressions' and 'lexical expressions').

Some distinctions between paradigm forms that are essential to the present conception of word formation presuppose paradigm bases, hence, cannot be made for affix forms; in particular: simple vs. analytic form and idiomatic vs. non-idiomatic form. These distinctions will now be explained.

2.5 Paradigm bases and kinds of forms

It should be uncontroversial to claim that f_3 = *lock*$_1$ in (iii) is a *simple*, actually, a *synthetic* word form, and f_4 = *the*$_1$ *lock*$_2$ in (iv) an *analytic* word form (allowing analytic noun forms in English). The distinction may then be carried over to f_1 = *lock*$_1$ in (i)—a *simple* stem form—and f_2 = *lock*$_1$ *s*$_2$ in (ii)—an *analytic* stem form. Drawing these distinctions is justified as follows.

Any paradigm base is assumed to contain a function *main part* (mp) that takes each form f in the basic set of the paradigm base (the set of noun-stem forms, of noun forms etc.) and assigns to f a certain part (subset) of f, its *main part*, which may be identical to f. The *auxiliary part* of f (which may be empty) is f without its main part. A *simple form* is one whose auxiliary part is empty, an *analytic form* is a non-simple form, i. e. one with a non-empty auxiliary part. For example, *lock*$_1$ is both a simple noun-stem form and a simple noun form, whereas *lock*$_1$ *s*$_2$ is an analytic noun-stem form (auxiliary part: *s*$_2$), and *the*$_1$ *lock*$_2$ is an analytic noun form (auxiliary part: *the*$_1$).

Furthermore, any paradigm base contains a second function, *centre*, that takes any f in the basic set and assigns to it a certain non-empty part of the main part of f, the *centre* of f, which may be identical to the main part. The *periphery* of f (which may be empty) is the main part of f without the centre of f. A form with a many-word centre or a non-empty periphery will be called an *idiomatic form* (the term is defined non-semantically, but is largely co-extensive with its use on a semantic definition). For example, in *lock*$_1$ and *lock*$_1$ *s*$_2$ (stem forms), the main part = the centre = *lock*$_1$; the two forms are non-idiomatic. Similarly, the centre of *lock*$_1$ (noun form) = the main part = *lock*$_1$, and the centre of *the*$_1$ *lock*$_2$ = the main part = *lock*$_2$; once again, the two forms are non-idiomatic.

But consider *un*$_1$ *lock*$_2$ *ed*$_3$. Here the main part is *un*$_1$ *lock*$_2$, the centre *lock*$_2$, and the periphery *un*$_1$: the form is idiomatic. It is also analytic, because of its non-empty auxiliary

[17] Inflectional affixes and their forms are connected with *syntactic* categories that are sets of forms of lexical words, such as, in English, Pres(-, S), Pl_N(-, S), or Sg_V(-, S). Some of these categories are semantically interpreted: Pres(-, S), Pl_N(-, S), some are only indirectly relevant to sentence meaning: Sg_V(-, S) (for details, see Lieb 2005: 1625-1627). The interpretation of the *categories* does not, however, give rise to any 'meanings' of the affixes or their forms. See Viguier (2013) for a large-scale application of this approach.

part *ed*$_3$. Now take the form *un*$_1$ *lock*$_2$: this is idiomatic, because of its non-empty periphery *un*$_1$, and *simple*, because of an empty auxiliary part. Similarly, the word form *locks*$_1$ *up*$_2$ is a *simple* word form: *up*$_2$ is the periphery not the auxiliary part, which is empty.

Neither the main part functions nor the centre functions give rise to a division of stem forms or word forms into constituents: they are *no head functions*, either in morphology or in syntax. Analytic forms as well as idiomatic forms occur only as primitive constituents; in isolation, they are 'heads' of themselves, if heads at all. In particular, the distinction between simple forms and analytic forms is independent of constituent structure.

Word-formation processes in an idiolect system determine only *simple* forms of paradigms (which need not be synthetic); analytic forms are obtained from simple ones by functions of inflection.[18] Thus, *un*$_1$ *lock*$_2$ is obtained from *lock*$_1$ by stem-form derivation, and *un*$_1$ *lock*$_2$ *ed*$_3$ from *un*$_1$ *lock*$_2$ by stem-form inflection; similarly, *locked*$_1$ *up*$_2$ from *locked*$_1$ by word-form or syntactic derivation, and *has*$_1$ *locked*$_2$ *up*$_3$ from *locked*$_1$ *up*$_2$ by word-form or syntactic inflection.

2.6 Fully specified units

2.6 a. Definition

Strictly speaking, the *result* of applying a word-formation process is not an individual stem form or word form but a *fully specified unit*, a triple ⟨f, J, b⟩ where J is a category set and b is a meaning either of f or of some f_1 of which f is a 'citation form'; moreover, such triples also provide the *starting-point* for word-formation processes. The notion of a fully specified unit is as follows:

(13) a. ⟨f, J, b⟩ is a *fully specified syntactic unit* of S iff there is a P such that (i) or (ii):
- (i) ⟨P, b⟩ is a lexical word of S, and ⟨f, J⟩ ∈ P;
- (ii) ⟨P, b⟩ is a pseudo-word of S, and for some f_1,
  - α. ⟨f_1, J⟩ ∈ P,
  - β. f = the citation form in S of f_1 given J and b [f = cit_S(J, b)(f_1)].

b. ⟨f, J, b⟩ is a *fully specified morphological unit* of S iff there is a P such that (i) or (ii):
- (i) ⟨P, b⟩ is a lexeme of S, and ⟨f, J⟩ ∈ P,
- (ii) ⟨f, J, b⟩ is a fully specified morphological word of S.

c. ⟨f, J, b⟩ is a *fully specified unit* of S iff ⟨f, J, b⟩ is a fully specified syntactic or morphological unit of S.

(Condition (b.ii) is relevant only in a context of inflection and will not be discussed here.) The more specific notions of *fully specified word form / stem form / affix form / word group (phrase)* are easily defined, making obvious changes in (13).

2.6 b. Pseudo-words and pseudo-paradigms

The definitions in (13) presuppose an *extension* of standard Word and Paradigm models, and also of the Word, Lexeme and Paradigm model in Lieb (2005): in addition to word-form, stem-form, and affix-form paradigms, we now include paradigms of phrases, calling

---

[18] This is a deviation from Lieb (2005).

them (syntactic) *pseudo-paradigms*. Any pseudo-paradigm is a set {⟨f, {K}⟩} where f is a (syntactic) group and K is a (syntactic) pseudo-category to which f belongs. The following set P_{11} is a *syntactic pseudo-paradigm* of German idiolect systems S:

(i) P_{11} = {⟨´*fünf*$_1$ *táge*$_2$, {{Nom, Pl_N, $Unsp_{Gen}$, $Unsp_{Def}$}-group}⟩}

The pseudo-category—'group category'—is automatically obtained from a syntactic analysis of the phrase ´*fünf*$_1$ *táge*$_2$: {Nom, Pl_N, $Unsp_{Gen}$, $Unsp_{Def}$}-group = the set of all phrases f' such that, on some syntactic analysis of f', f' is 'marked' as Nom, Pl_N, $Unsp_{Gen}$, and $Unsp_{Def}$.[19] Given this pseudo-category and a suitable lexical interpretation for the phrase's primitive constituents, ·five days· is obtained as a lexical meaning of ´*fünf*$_1$ *táge*$_2$.

A *pseudo-word* consists of a syntactic pseudo-paradigm together with a lexical meaning of its forms; for example:

(ii) ⟨P_{11}, ·five days·⟩ is a pseudo-word of S, where S is a German idiolect system.

If, in addition to syntactic phrases, *morphological groups* are assumed, we may have pseudo-lexemes in addition to pseudo-words. A *pseudo-lexeme* then consists of a morphological pseudo-paradigm and a lexical meaning of its forms. For example, assume for English idiolect systems S that

(iii) P_{12} = {⟨*trúck*$_1$ *dríve*$_2$, {{Inf-St, Pres-St}-group}⟩}

is a *morphological pseudo-paradigm*, *trúck*$_1$ *dríve*$_2$ being a stem group that means ·drive trucks·. Then:

(iv) ⟨P_{12}, ·drive trucks·⟩ is a pseudo-lexeme of S.

However, I now prefer an analysis by which *trúck*$_1$ *drive*$_2$, initially stressed, is a stem form obtained by stem-form compounding, available for use in stem-form derivation (*trúck*$_1$ *drive*$_2$ *er*$_3$); *trúck*$_1$ *drive*$_2$, categorized as {Inf-St, Pres-St}, is the only form of a trapped stem lexeme meaning ·drive trucks·, a stem lexeme for which there is no corresponding lexical word. Either way, the notorious 'bracketing paradox' presented by 'synthetic compounds'—an apparent contradiction between build-up of meaning and build-up of form—is solved.[20]

---

[19] The correctness of these categories is presupposed here. Marking is via a so-called 'marking function'; cf. Lieb (1993b: 438).

[20] Following Lieber (2009: 367), the literature dealing with so-called synthetic compounds in English offers two ways of assigning internal structure to relevant stem forms such as *trúck*$_1$ *drive*$_2$ *er*$_3$: one that would correspond to applying the compounding process after the derivation process (the majority view), the other, to applying the derivation process after the compounding process (minority view). Since we are confronted with a semantic ambiguity—'driver of a certain truck' vs. 'person who professionally drives trucks'—both analyses may be correct: the majority view in relation to the first meaning, the minority view for the second meaning. Problems arise when the majority view is extended to also account for the second meaning; this is avoided on our proposal. (Indeed, there is another possibility that may be considered for dealing with the second meaning: instead of the compounding process we may apply the conversion process, starting from a Verb Phrase such as *dríve*$_1$ *trúcks*$_2$; the derivation process is then applied to the conversion-process result.)

2.6 c. Citation forms

In (13a), lexical words are treated differently from pseudo-words: forms of lexical words may enter directly into fully specified units. In the case of a phrase f_1 that is a form of a pseudo-word, this is true of the *citation form* of f_1: the form in which it is quoted as a linguistic unit. While the citation form f of f_1 may be identical to f_1, it may also differ from f_1, in particular, differ with respect to intonation. For example (omitting the relativization to idiolect systems S, categorizations J and meanings b):

(v) The citation form of *´fünf*$_1$ *táge*$_2$ = *fünf*$_1$ *táge*$_2$. (*´fünf*$_1$ is destressed.)

It is (v) that is used in the stem-form compounding process leading up to the result form

(vi) *fünf*$_1$ *táge*$_2$ *wòch*$_3$ *e*$_4$,

meaning ·five-day week·. The intonation change in (v)—partial destressing—should not be construed as due to applying a word-formation process, in contrast to the stress reduction on *wóch*$_1$ *e*$_2$ that occurs in the transition from *fünf*$_1$ *táge*$_2$ and *wóch*$_1$ *e*$_2$ to *fünf*$_1$ *táge*$_2$ *wòch*$_3$ *e*$_4$.[21]

2.6 d. Reliance on fully specified units

On our conception of word formation, applying a word-formation process involves a transition from a fully specified unit to a fully specified unit, not from simply a unit to a unit (this proves untenable).

In its reliance on fully specified units the conception is similar to the approach—within a Categorial-Grammar framework and restricted to stem-form compounding—that is adopted by Hoeksema (1985: 12), who considers similar triples to be elements of the 'lexicon', as does Aronoff (2007: 803 fn. 1) in using the term 'lexeme' (however, 'lexemes' in the sense of Aronoff 1994: 9-11 are more abstract than our fully specified units). There are also similarities with word conceptions used implicitly or explicitly in declarative frameworks like HPSG, or in 'parallel architecture' (e. g., Jackendoff 2002a), from where it is taken over by Booij (2010: 5-6) into his Construction Morphology; also compare *atomare Konstruktion* (atomic construction) in Jacobs (2008: 6), and 'indivisible [lexical] nuclei' in Beard (1995: 46); for Nolda (2012b), see below, Sec. 8.3. Still, there are major differences; in particular, the notion of fully specified unit presupposes a Word and Paradigm model extended to account for pseudo-words, with a different view of the lexicon.

The conception of word-formation processes, generally and in an idiolect system, will now be characterized by means of examples from English for the three basic processes, the compounding process, the derivation process, and the conversion process. For each basic process in English idiolect systems S, examples will be given for the two major subcases separately: the morphological or stem-form process in S and the syntactic or word-form process in S. Syntactic groups (phrases) will not figure in the key examples, which reduces the fully specified units that occur to the simpler case where pseudo-words and their citation forms are not involved.

[21] Citation forms are used in this context also by Nolda (2012b: Sec. 5.2.3), who first noticed the intonational change for German; it also holds in English.

B. The conception

3 Stem-form and word-form compounding in S: examples

3.1 *dóor*$_1$ *lock*$_2$: Introduction. The result and its basis

3.1 a. Introduction

Word-formation processes in the general sense and their subcases are to be identified with functions α, and word-formation processes in an idiolect system S with functions $\alpha(S) = \alpha_S = \varphi$. The basic idea for these functions will now be explained, using α = stem-form compounding = st-comp, and φ = stem-form compounding in (suitable English idiolect systems) S = st-comp$_S$.

It is the function st-comp$_S$, the morphological subcase of the compounding process in S (comp$_S$), that is the word-formation process involved in the formation of the lexical word ⟨*dóor lock*^{P}, ·lock for door·⟩. This function is used for the construction of simple (i.e., non-analytic) stem forms, as informally indicated in (i) (accent symbols for word stress):

(i) The stem-form compounding in S
of: *lóck* as a specific Noun-stem form, meaning ·lock$_1$·,
and
dóor as a specific Noun-stem form, meaning ·door·,
by means of:
a concatenation function and other functions,
is:
dóor lock as a specific Noun-stem form, meaning ·lock for door·.

The *of*-part and the *by means of*-part of (i) jointly render an argument of φ = st-comp$_S$, the *of*-part representing the *basis* of the argument and the *by means of*-part its *construction mode*; the *is*-part represents the *result*—the value—of st-comp$_S$ for this argument. The construction mode consists of a number of functions that apply to components of the basis to yield the result. All word-formation processes φ are to be functions of essentially this type.

3.1 b. Result triple, basic triple, and added triple

The result of applying φ = st-comp$_S$ must be a fully specified stem form ⟨f, J, b⟩:

(ii) ⟨f, J, b⟩ = ⟨*dóor*$_1$ *lock*$_2$, {Unsp$_{Case}$-St, Sg$_N$-St, Unsp$_{Def}$-St}, ·lock for door·⟩

The accent sign means primary word stress, lack of an accent sign means no stress; arguably, *lock*$_2$ has secondary stress.[22]

The basis for applying st-comp$_S$ must contain *lóck*$_1$ and *dóor*$_1$. Since the result is to be a triple ⟨f, J, b⟩, we also start from such triples rather than from the two individual forms: we proceed from a fully specified stem form with f_1 = *lóck*$_1$ to the fully specified stem form in (ii), having added a fully specified stem form ⟨f_2, J_2, b_2⟩ with f_2 = *dóor*$_1$:

(iii) ⟨f_1, J_1, b_1⟩ = ⟨*lóck*$_1$, {Unsp$_{Case}$-St, Sg$_N$-St, Unsp$_{Def}$-St}, ·lock$_1$·⟩,

where ·lock$_1$· is as in (8), Sec. 2.3; and

(iv) ⟨f_2, J_2, b_2⟩ = ⟨*dóor*$_1$, {Unsp$_{Case}$-St, Sg$_N$-St, Unsp$_{Def}$-St}, ·door·⟩,

where ·door· is a one-place concept that applies to devices for closing and opening (otherwise, the concept will be left unspecified). The two triples (iii) and (iv), in this order, form *the basis* of the application of st-comp$_S$ that yields triple (ii) as a result; ⟨f_1, J_1, b_1⟩ is *the basic triple*, with the *basic unit* f_1, the *basic category set* or *basic categorization* J_1, and the *basic concept* or *basic meaning* b_1; ⟨f_2, J_2, b_2⟩ is the *added triple*, with the *addition* f_2, the *added category set* J_2, and the *added concept* b_2.

3.1 c. Comments

The basic triple precedes the added triple in the basis because the *centre* (cf. Sec. 2.5) of the result form *dóor*$_1$ *lock*$_2$, i. e., *lock*$_2$, 'goes back to' f_1 = *lóck*$_1$, and the *periphery* of the result form, i. e., *dóor*$_1$, 'goes back to' f_2 = *dóor*$_1$ (with which it happens to be identical). Generally, the basic triple and the added triple are distinguished in this way, and in an argument of the function st-comp$_S$, the basic triple precedes the added triple. The basic triple is also the one containing the category set J_1 to which the 'category-change function'—Sec. 3.4, below—applies.

In the *dóor*$_1$ *lock*$_2$-example the result triple is from an English idiolect system S, and both the basic and the added triples are from the same system as the result triple. Being from the same system should not be made into a general requirement, though. In German systems S, we may have stem-form compounding in S with added triples from a system S_1 different from S; S_1 may even be from another language, as in *búsiness*$_1$ *àn*$_2$ *zùg*$_3$ ·business suit· (attested example), where the periphery *búsiness*$_1$ presupposes an added triple with *búsiness*$_1$, from an English system S_1. (Complete or partial borrowing from English is currently endemic in many languages, especially so in German.)

---

[22] For the phonological assumptions on English, see fn. 9 above; for the categorizations, fn. 7, and generally Sec. 2.1 a.

3.1 d. From basis to result: the construction mode

In any construction mode there is a first pair of functions that may change the basic unit f_1; a second pair may change the addition f_2; function number five, 'arrangement', combines the sequences that are obtained from f_1 and f_2 by applying the two function pairs. (Relating the aspects of *change* and *arrangement* in this way is not self-understood; it has simply turned out to be optimal.) Two more functions account for category change and semantic change, respectively. The various functions will now be specified for our example, and motivated by some additional data.

3.2 *dóor*$_1$ *lock*$_2$: shortening, form change, and arrangement

3.2 a. Shortening

In an instance of a stem-form process in S, the basic unit may have to be shortened before being combined with the addition. *True shortening* of the basic unit occurs in typical cases of *stem-form cutting in* S, as in the transition from *mòbile*$_1$ *phóne*$_2$ to *móbile*$_1$, or of stem-form back formation in (German systems) S, as in the transition from *nót*$_1$ *land*$_2$ *ung*$_3$ ·emergency landing· to *nót*$_1$ *land*$_2$ ·perform an emergency landing·. In the first case, we apply the shortening function short_S(*phóne*$_2$), or the shortening in S by *phóne*$_2$, to *mòbile*$_1$ *phóne*$_2$, taking away its part *phóne*$_2$ to obtain the new sequence *mòbile*$_1$ (still with secondary not primary stress); in the second case, the shortening function short_S(*ung*$_3$) is applied to *nót*$_1$ *land*$_2$ *ung*$_3$ to yield the sequence *nót*$_1$ *land*$_2$.[23]

The result of applying a shortening function must always be a sequence. This happens to be the case in the two examples but is no longer true when we proceed from *téle*$_1$ *phone*$_2$ to *phone*$_2$. Here, *phone*$_2$ must be adjusted, by a purely formal operation of *adjustment*, to obtain the sequence *phone*$_1$. (The adjustment of a sequence is the sequence itself.)

Generally, when a *shortening function* is applied to any (empty or non-empty) sequence of morphs or phonological words of S, it does not change any members of the sequence but yields the difference of the sequence and one of its (empty or non-empty) parts, adjusted to make it a sequence. If the part is empty, we have *zero-shortening in* S, with no overt effect.

Indeed, *the basic shortening function* β_{11} in the *dóor*$_1$ *lock*$_2$-example, a shortening function for f_1 = *lóck*$_1$, is zero-shortening in S:

(i) $\beta_{11}(\textit{lóck}_1) = \text{zero-short}_S(\textit{lóck}_1) = 0_S(\textit{lóck}_1) = \text{short}_S(\varnothing)\ (\textit{lóck}_1)$
$= \text{id}(\textit{lóck}_1) = \textit{lóck}_1$.

(id = the set-theoretic *identity function*.)

True shortening must be admitted also for the *addition* in applying a stem-form process; when stem-form compounding in German systems S is applied to *haús*$_1$ ·house· and *schúl*$_1$ *e*$_2$ ·school· to obtain *schúl*$_1$ *haùs*$_2$ ·school house·, there is *e*$_2$-shortening of the addition *schúl*$_1$ *e*$_2$. In the *dóor*$_1$ *lock*$_2$-example, we therefore assume zero-shortening also as the *added shortening* β_{21}, for the addition f_2 = *dóor*$_1$.

[23] See also below, Secs 4.5 a and 4.5 b.—Shortening functions must specify a relevant part of the unit to which they apply, and are therefore quite specific. This does not preclude general ways of determining them.

3.2 b. The importance of shortening functions

Shortening functions may be far from marginal. This has recently been emphasized by Masini & Benigni (2012) for word formation in Russian, where, among others, the Russian version of the *mòbile*$_1$ *phóne*$_2$-example is analysed in a constructionist framework (the analysis differs from, but is compatible with, the one given below, in Sec. 4.5 a).

The English and Russian examples are similar to instances of 'subtraction', on the level of syntactic rather than morphological units. In Manova (2011: 45), 'subtraction' is recognized as a 'basic morphological technique' and subsequently (Ch. 4) discussed in detail for three Slavic languages. Shortening functions may be taken to reconstruct subtraction to the extent that subtraction is meant to remove complete morphs or complete phonological words from a sequence of morphs or words, rather than remove proper parts of morphs or phonological words; this would constitute *form change* on our account.

It will eventually turn out that we are dealing with *conversion* in S (below, Sec. 5.4) in the two examples, with shortening of the basic unit (basic shortening). In Sec. 3.2 a, 'added shortening' was justified for *stem-form compounding* in S. Basic shortening can also be used to show that so-called 'analogical word formation' (a case of 'paradigmatically related words' as assumed in Construction Morphology: Booij 2012: 354) is, as a matter of fact, based on the *compounding process* in S or the *derivation process* in S.

For example, there are the following two fully specified stem forms in German systems S, ⟨*haús*$_1$ *fraù*$_2$, {Sg$_N$-St, …}, ·house wife·⟩ and ⟨*haús*$_1$ *mànn*$_2$, {Sg$_N$-St, …}, ·male partner having taken over the house duties of a house wife·⟩. It may be claimed that the second triple is either formed directly from a *mánn*$_1$-triple and a *haús*$_1$-triple, or formed from the *haús*$_1$ *fraù*$_2$-triple 'by analogical substitution' (two solutions discussed in Nolda 2012b on the basis of Becker 1993). However, the example may be rendered as a case of stem-form compounding in S where the basic triple is the *haús*$_1$ *fraù*$_2$-triple, the added triple is a *mánn*$_1$-triple, and the shortening function is short$_S$(*fraù*$_2$) applied to *haús*$_1$ *fraù*$_2$; stress reduction is applied to *mánn*, and concatenation yields the result form. On this analysis, neither does the *haús*$_1$ *mànn*$_2$-triple result from applying stem-form compounding to a *mánn*$_1$-triple and a *haús*$_1$-triple, which is counterintuitive, nor do we have to apply a special word-formation process of 'analogical substitution' to obtain the *haús*$_1$ *mànn*$_2$-triple directly from the *haús*$_1$ *fraù*$_2$-triple, an unnecessary complication; the alternative analysis is to be preferred.

Our analysis of the example involves 'substitution' in the sense of first applying a shortening function and then applying concatenation as a special case of an arrangement function. In Manova (2011: 45), 'substitution' is still listed as a 'basic morphological technique' and is exemplified by the Bulgarian version of the English pair *márx*$_1$ *ism*$_2$ / *márx*$_1$ *ist*$_2$. This 'technique' is *not* basic, though. It can be rendered as a combination of a shortening function and an arrangement function (roughly, Manova's 'addition', another 'basic morphological technique'). In the *márx*$_1$ *ism*$_2$ / *márx*$_1$ *ist*$_2$-example, we are dealing with derivation in S not with compounding in S but again, basic shortening is followed by concatenation: informally, *márx*$_1$ *ist*$_2$ is obtained from *márx*$_1$ *ism*$_2$ by *ist*$_1$-derivation in S, first applying short$_S$(*ism*$_2$) to *márx*$_1$ *ism*$_2$ and then concatenating the result with *ist*$_1$.

Actually, there are two ways of understanding 'substitution': either as the combination of a *shortening* function and an arrangement function (typically, a concatenation function), as in the example just discussed, or as the combination of a *form-change* function (below, Sec. 3.2 c) and an arrangement function. The second interpretation would apply if an unanalysed stem form *márxism*$_1$ were to be assumed in the example, plus the replacement of the *ism*-part of *márxism* (taken to be a single morph) by an *ist*-part, the replacement being effected by a form-change function; we would then be confronted with the conversion process in S, not the derivation process. It may not always be easy—compare some examples in Stump (2010: Sec. 2)—to decide which interpretation should be assumed.

As demonstrated by the *mòbile*$_1$ *phóne*$_2$- and *téle*$_1$ *phone*$_2$-examples, we may still have to apply proper form change to the sequence resulting from shortening. This is also true of the *dóor*$_1$ *lock*$_2$-example, where *lóck*$_1$, the result of zero-shortening, must be changed to unstressed *lock*$_1$.

3.2 c. Form change and primitive change

In order to account for unstressed *lock*$_2$, we introduce a function β_{12} = *destressing in* S, or destr$_S$, that operates on *lóck*$_1$ and assigns to it the sequence *lock*$_1$, i. e., removes word stress:

(ii) destr$_S$(*lóck*$_1$) = *lock*$_1$

Formally, the effect of destr$_S$ on *lóck*$_1$ is *the replacement of* {H} *by* {L}, or {H}/{L}. This function takes *lóck*$_1$ and changes the intonation structure of its only member *lóck* (as in (7), Sec. 2.2) by replacing the indicator of primary word stress, {H}, by the indicator of lack of stress, {L}. (If secondary stress on *lock*$_2$ is assumed, i. e., if f is to be *dóor*$_1$ *lòck*$_2$, we would have a function by which {H} in *lóck* is changed to {H, L}, the indicator of secondary stress.) {H}/{L} is a *primitive change function*, and destr$_S$ is a *form-change function*, *the basic form-change*. In this example, destr$_S$ has an effect only on intonation structures since destr$_S$(*lóck*$_1$) = {H}/{L}(*lóck*$_1$); in other cases, sound sequences and constituent structures may also be affected by destressing.

There may of course be form changes that affect only the sound sequences, or both the sound sequences and the constituent structures, as in the transition from *déep*$_1$ to *dép*$_1$ when stem-form derivation with *th*$_1$ is applied to yield the result form *dép*$_1$ *th*$_2$. Here, the basic form-change β_{12} is *vowel-reduction in* S, which involves two primitive change functions, a sound-change function, /ii/-replacement in S, that changes /ii/ to /e/ in the sound sequence of *déep*, and a structure-change function that subsequently adjusts the constituent structure to the new sound sequence /dep/. There is no intonation change; this may be expressed by allowing the *identity function*, id, as an *intonation-change function*. Since there need not be either a sound change or a structure change, identity should be admitted also as a sound-change function and a structure-change function.

On the other hand, primitive change functions, one of each type and none of them identity, may co-occur. In the transition from the stem form *sudán*$_1$, assuming a long vowel on the second syllable and an extrasyllabic /n/, to the stem form *sùdan*$_1$ *ése*$_2$, with the short central vowel on the second syllable of *sùdan* and again with an extrasyllabic /n/, we apply a basic form-change function to *sudán*$_1$, to be called *stress-fronting with vowel reduction in* S, that simultaneously affects the intonation structure, the sound sequence, and also the syllable structure assuming that long vowels are represented through doubling.

It should be emphasized that form-change functions operate at the phonological not the phonetic level (distinguished as in Lieb 2008); changes irrelevant to word formation—such as final consonant devoicing (obstruent tensing) in German or in Dutch—are not covered. But consider a form change that is a phonological effect of applying a word-formation process, such as the form change in a basic unit that is due to combining *ése*$_1$

with the unit: this will indeed be represented through a form-change function that is connected with the process, rather than be treated independently of the process.

In proceeding from f_1 = *lóck*$_1$ and f_2 = *dóor*$_1$ to f = *dóor*$_1$ *lock*$_2$, the members (there is only one) of f_2 remain unchanged; this is also true of f_2 in the transition from f_1 = *déep*$_1$ and f_2 = *th*$_1$ to f = *dép*$_1$ *th*$_2$. But such constancy of the addition f_2 must not be generally assumed: not only do we need a form-change function β_{12} for the basic unit f_1 but also a form-change function β_{22} for the addition f_2: *the added form-change*. For example, the German affix form f_2 = *heit*$_1$ is changed to *keit*$_1$ under certain morphophonological conditions (to be satisfied by the basic unit), and to *igkeit*$_1$ under others (for the much-discussed details, see Barz 2005: 732-733). More importantly, form change for f_2 must be allowed when *blending processes in* S or *reduplication processes in* S are applied, as subcases of compounding processes in S.

In the *dóor*$_1$ *lock*$_2$-example, the basic form-change β_{12} and the added form-change β_{22} are as follows:

(iii) β_{12} = destr_S = destressing in S

(iv) β_{22} = id = (set-theoretic) identity

(The addition f_2 = *dóor*$_1$ is left unchanged by the form change β_{22}.)

Applying the two form changes to the shortening results, we obtain the pair of sequences to which an arrangement function γ applies:

(v) $\langle(\beta_{12}(\beta_{11}(f_1)),\ \beta_{22}(\beta_{21}(f_2))\rangle$
$= \langle \text{destr}_S(0_S(\textit{lóck}_1)),\ \text{id}(0_S(\textit{dóor}_1))\rangle$
$= \langle \textit{lock}_1, \textit{dóor}_1\rangle$

3.2 d. The arrangement function

The pair $\langle \textit{lock}_1, \textit{dóor}_1\rangle$ is a pair of (unit) sequences. The form f = *dóor*$_1$ *lock*$_2$ is another sequence, obtained from the pair by *inverse concatenation*, $^{\cup}$.

This function applies as in:

(vi) $\gamma(\beta_{12}(\beta_{11}(f_1)),\ \beta_{22}(\beta_{21}(f_2)))$
$= {}^{\cup}(\text{destr}_S(0_S(\textit{lóck}_1)),\ \text{id}(0_S(\textit{dóor}_1)))$
$= \textit{lock}_1 {}^{\cup} \textit{dóor}_1$
$= \textit{dóor}_1\ \textit{lock}_2$

In this example, then:

(vii) γ = the arrangement function = inverse concatenation = $^{\cup}$

It is here that the construal of all morphological and syntactic units as sequences—including unit sequences—is of vital importance: sequences are sets, and functions γ are purely formal, set-theoretic operations on pairs $\langle f', f''\rangle$ of sequences.

Concatenation functions are a special case of set-theoretic operations on sequences but not the only one; 'non-concatenative morphology' is just as easily accounted for as concatenative versions are, in both its replacive aspect (through form-change functions) and its combinatory aspect (through arrangement functions other than concatenation or inverse concatenation).

Arrangement functions may change the order of members of f' or f" but may not shorten or lengthen the sequences, nor may they have access to the members themselves: all changes within the members of the basic form f_1 or the addition f_2 and any removal of pairs ⟨n, w⟩ that are necessary to arrive at the result form f must have been effected by the shortening and form-change functions before the arrangement function γ applies.

One way or the other, the arrangement function and the modifying functions, shortening and form change, are sufficient to deal with the transition from the basic form and the addition, say, f_1 = *lóck*$_1$ and f_2 = *dóor*$_1$, to the result form, e.g., f = *dóor*$_1$ *lock*$_2$. We must still consider how the result categorization J and the result meaning b of f are obtained.

3.3 *dóor*$_1$ *lock*$_2$: category change

3.3 a. The category-change function

In the *dóor*$_1$ *lock*$_2$-example, the basic category set J_1, added set J_2, and result set J are identical: {$Unsp_{Case}$-St, Sg_N-St, $Unsp_{Def}$-St}. But from a general point of view, this is accidental; in particular, we must allow for a change in going from the basic set J_1 to the result set J. This is typical of stem-form derivation, as in the transition from ⟨*déep*$_1$, J_1, ·deep·⟩ to ⟨*dép*$_1$ *th*$_2$, J, ·depth·⟩, where J_1 = {$Neut_N$-St} (Neutral with respect to Noun-stem distinctions) and J = {$Unsp_{Case}$-St, Sg_N-St, $Unsp_{Def}$-St}. Generally, we need a *category-change function* δ that takes the basic category set J_1 and changes it into the result set J.

Intuitively, δ in a construction mode is to change J_1, the basic category set, into J, the result category set, by taking a proper or improper non-empty subset J' of J_1 and replacing it by a non-empty set J"; that is, we first form the set-theoretic difference J_1\J' and then the union of J_1\J' and J":

(i) $\delta(J_1) = (J_1 \backslash J') \cup J'' = J$.

The two sets J' and J" must be given as part and parcel of δ. This is achieved by construing δ, a function, as the value of another function *to* (definable in set-theoretic terms), on the pattern of:

(ii) δ = J'-to-J".

Combining (i) and (ii), we obtain

(iii) $\delta(J_1) = J'\text{-to-}J''(J_1) = (J_1 \backslash J') \cup J'' = J$.

In the *déep*$_1$ / *dép*$_1$ *th*$_2$-example, J' = J_1 = {$Neut_N$-St}, and J" = {$Unsp_{Case}$-St, Sg_N-St, $Unsp_{Def}$-St}, that is,

(iv) δ = {$Neut_N$-St}-to-{$Unsp_{Case}$-St, Sg_N-St, $Unsp_{Def}$-St}

Applying δ as in (iii), we obtain:

(v) $\delta(\{\text{Neut}_N\text{-St}\})$
$= \{\text{Unsp}_{\text{Case}}\text{-St}, \text{Sg}_N\text{-St}, \text{Unsp}_{\text{Def}}\text{-St}\}$

Here, we have *complete replacement* of the basic category set J_1. The other extreme is no change at all, as in the *dóor*$_1$ *lock*$_2$-example. This is covered again by allowing δ to be the identity function.

Remark. As appears from the example, the added set J_2 does not figure in the arguments of the category-change function δ; rather, it helps to choose the function δ itself. The added category set J_2 is excluded from the arguments of the category-change function because what is subject to change is the basic categorization, not the added category set, which plays only an auxiliary role.

3.3 b. The word-category assignment

Category sets like J_1, J_2, and J consist of categories that are sets of forms. Such category sets, plus a function δ to provide for transitions between them, are not yet sufficient to account for all categorial aspects of word formation.

Consider (2) in Sec. 1.4, repeated here as:

(i) ⟨⟨*dóor lock*^{P}, ·lock for door·⟩, ⟨*lóck*$_1^P$, ·lock$_1$·⟩, ⟨*dóor*^{P}, ·door·⟩, stem-form compounding in S⟩ ∈ word formation in S.

Given (i) and the information that ⟨*lóck*$_1^P$, ·lock$_1$·⟩ is an element of the lexical-word category COUNT NOUN in S, we know that ⟨*dóor lock*^{P}, ·lock for door·⟩ also is. Arguably, this is implied by Sg_N-St being an element of both J_1 and J as these are specified in (iii) and (iv) of Sec. 3.1 b, and ·lock$_1$· being a concept that applies to things. Still, the fact that ⟨*dóor lock*^{P}, ·lock for door·⟩ is a count noun due to ⟨*lóck*$_1^P$, ·lock$_1$·⟩ being one is a fact about a complete lexical word ⟨P, b⟩ not about one of its forms or about a form of its stem paradigm. This fact should therefore be accounted for after the lexical word has become available.[24]

I therefore assume a function ψ whose domain is the word formation in S such that, if ⟨⟨P, b⟩, ⟨P_1, b_1⟩, ⟨P_2, b_2⟩, φ⟩ ∈ word formation in S, then ψ(⟨P, b⟩, ⟨P_1, b_1⟩, ⟨P_2, b_2⟩, φ)—i. e., *the word-category assignment in* S *to* ⟨P, b⟩ *given* ⟨P_1, b_1⟩, ⟨P_2, b_2⟩, *and* φ—is a set of lexical-word categories of S to which ⟨P, b⟩ belongs. An analogous function of *stem-category assignment in* S is also needed. Both assignment functions may be introduced as subcases of a single function, *lexical-category assignment in* S, if ⟨P, b⟩ is allowed to be any lexical item of S, i.e., any lexical word or stem lexeme of S.

Remark. The lexical-category assignment of S may be introduced as a separate morphosyntactic component of S (below, Sec. 6.5 d). The following types of lexical-word categories (analogously, stem-lexeme categories) may be considered as potentially determined through word formation: (categories of) part of speech, gender, government, and inflection class. It may well turn out that in word or lexeme formation in S all of these categories are determined for ⟨P, b⟩ by the word-formation process φ alone, due to proper-

[24] This is contrary to Nolda (2012b) where lexical-word and stem-lexeme categories are already accounted for by a component of a quadruple that is otherwise analogous to a fully specified unit ⟨f, J, b⟩ in our sense; however, Nolda does not attempt to directly identify lexical words in word formation (see below, Secs 8.1 and 8.3).

ties of the basic, added, or result triples. In this case, the lexical-category assignment need not be a separate component of the idiolect system.

Returning to the *dóor lock*-example, we now consider its semantic aspect.

3.4 *dóor₁ lock₂*: meaning change. Summary

3.4 a. Identifying the result meaning [25]

The *result meaning* b must be determined from the *basic meaning* b_1 and the *added concept* b_2; in our example, b = ·lock for door· is to be determined from b_1 = ·$lock_1$· and b_2 = ·door·.

The intension of the result concept b = ·lock for door· consists of a property that is informally characterized by the following formula:

(i) the property of being an x such that:
x is a lock, and there is an x' such that x is meant for x' and x' is a door.

What is involved in (i) is two *properties*, being-a-lock and being-a-door, plus one *relation*, is-meant-for. The two properties are equivalent to, respectively, being in the extension of the concept ·$lock_1$·, defined as in (8), and being in the extension of the concept ·door·. Being related by the relation is-meant-for is equivalent to being in the extension of a two-place concept ·meant-for·:

(ii) $\langle x, x' \rangle \in {}^{e2}$·meant-for· iff: x is a [lock-type] appliance, and it is intended that it is possible that there are x_1, x_2, and x_4 such that $\langle x_1, x_2, x', x_4 \rangle$ satisfies the purpose of x.

A lock-type appliance is an appliance whose purpose is formally like the purpose of a lock, see (8). This must be required because of ‘purpose’ in (ii). Formula (i) is equivalent to:

(iii) the property of being an x such that:
$x \in {}^{e1}$·$lock_1$·, and there is an x' such that: $\langle x, x' \rangle \in {}^{e2}$·meant-for· and $x' \in {}^{e1}$·door·.

Property (iii) is a property of real-world objects x, such as locks. This property may be obtained by a function that applies to the pair of concepts ⟨·$lock_1$·, ·door·⟩ while at the same time newly introducing the relational concept ·meant-for·. Call this function ‘·meant-for·-SPECIFYING’, or ‘SPEC(·meant-for·)’, for short. We then have:

(iv) SPEC(·meant-for·)(·$lock_1$·, ·door·) = (iii); to be read as: ‘the ·meant-for·-specifying of ·$lock_1$· by ·door· is (iii)’

Property (i) = property (iii) is identified by (iv) as follows: it is the value adopted by the function SPEC(·meant-for·) for the argument that is the concept pair consisting of: the basic meaning ·$lock_1$· and the added concept ·door·.

[25] See Sec. 2.3, above, for the presupposed background in semantics.

Property (iii), a *property of objects* x, is not yet the concept ·lock for door·. Assuming the presupposed theory of concepts (above, Sec. 2.3 a), the concept ·lock for door· is to be a *content-based property of perceptions or conceptions* z that consists in each perception or conception z having a *content* that is a set of properties of objects x. Informally, ·lock for door· is to be the property of being a perception or conception z in whose content property (i), identified as in (iv), occurs. More precisely, ·lock of door· is:

(v) the property of being a perception or conception z such that {SPEC(·meant-for·)(·$lock_1$·, ·door·)} $\subseteq$ the content of z

Property (v), a property of perceptions or conceptions, may be obtained by a function called 'spec(·meant-for·)', or 'the ·meant-for·-specification', a function that is like SPEC(·meant-for·) but has properties of perceptions or conceptions z as its values instead of properties of objects x:

(vi) spec(·meant-for·)(·$lock_1$·, ·door·) = (v); to be read as: 'the ·meant-for·-specification of ·$lock_1$· by ·door· is (v)'

This finally allows us to actually *define the term* '·lock for door·':

(vii) ·lock for door· $=_{df}$ spec(·meant-for·)(·$lock_1$·, ·door·)

Presupposing the theory of concepts, it follows from (vii) that ·lock for door· is indeed a concept (more precisely, a 'potential' concept as long as nobody 'has' it), whose *intension* is the unit set of property (iii) and whose *extension* is the set of all objects x that have property (iii), i.e., the set of all locks that are meant for doors.

The status of ·lock for door· as a concept is independent of being a meaning in any language. It is an additional *empirical assumption*, to be made in a grammar of English, that ·lock for door· defined as in (vii) is indeed a *meaning* of the word paradigm *dóor lock*^{P} in English idiolect systems S, an assumption that may be wrong (it should be right).

3.4 b. The meaning-change function. Specifying and specification

The way the result meaning ·lock for door· has been identified may seem unduly circumstantial. However, the identification presents a non-trivial example for the way result meanings are identified when the compounding process in idiolect systems S is applied in one important case: applied to obtain 'determinative' compounds of a basic type that may well be represented in any language that has compounds. This type may be called *external-relation compound* due to a feature of the underlying subcase of the compounding process: the subcase involves meaning-change functions that introduce an additional relational concept 'from the outside'; using this concept, the basic meaning b_1, itself *non*-relational, is further specified by means of the added concept b_2. (If b_1 is a relational concept, b_1 itself provides the relation to be used by the meaning-change function, and we obtain an *internal-relation compound*, variously called 'Rektionskompositum' in German, or 'VN-compound'—one type—in English.)

There are two interrelated functions in the *dóor*$_1$ *lock*$_2$-example that may be suggested as the meaning-change function ε: either the ·meant-for·-*specifying* (SPEC(·meant-for·)), or the ·meant-for·-*specification* (spec(·meant-for·)). It is only the second function, based on the first, that may be chosen: both functions have pairs of concepts as their arguments but only the second directly assigns to each pair another concept. The meaning-change function in the *dóor*$_1$ *lock*$_2$-example is therefore assumed to be:

(viii) ε = spec(·meant-for·) [‘the ·meant-for·-specification’]

It appears from (viii) that the meaning-change function ε itself is the value of a function, viz. the value assigned to the concept ·meant-for· by the function *spec*, or *specification*. Indeed, ·meant-for· is not the only concept that qualifies as an argument of spec: we may have not only ·meant-for·-specification but also ·shaped-like·-specification, ·placed-on·-specification, etc., functions that apply to ⟨·lock$_1$·, ·door·⟩ or to some other concept pair. All these functions share one basic feature: they are *specification functions*, values of the single function spec. Therefore, a definition for ‘spec’ or ‘specification’ is in order, to be based on a definition of ‘SPEC’ or ‘specifying’ (certain generalizations of the definitions may have to be considered):

(14) Suppose that b_1 and b_2 are one-place concepts and b' is a two-place concept.
a. SPEC(b')(b_1, b_2) [‘the b'-specifying of b_1 by b_2’] $=_{df}$ the property of being an x such that:
(i) $x \in {}^{e1}b_{1;}$
(ii) there is an x' such that: $\langle x, x' \rangle \in {}^{e2}b'$ and $x' \in {}^{e1}b_2$.
b. *spec*(b')(b_1, b_2) [‘the b'-specification of b_1 by b_2’] $=_{df}$ the property of being a perception or conception z such that {SPEC(b')(b_1, b_2)} $\subseteq$ the content of z.

This applies as exemplified in (iii) to (vi); in particular, ·lock for door· is the ·meant-for·-specification of ·lock$_1$· by ·door·.

Remark on spec. There is a well-known *problem* raised by the basic type of determinative compounds that I called ‘external-relation compounds’: how to identify the relations used in forming such compounds. Clearly, there are typical relations; is-meant-for is one. However, the existence of ad-hoc compounds appears to ruin any attempt to specify such relations by enumerating them. Definition (14) offers a way out. For a given idiolect system S, the relations to be allowed are exactly the ones introduced by meaning-change functions ε as follows: there is a two-place relational concept b' that is an argument of the spec function, and ε is the value of spec for b'; and for any argument $\langle b_1, b_2 \rangle$ of ε, b_1 and b_2 are ‘formally compatible’ with b' and are lexical meanings in S. No restrictions need be imposed on b', which allows us to include ad-hoc compounds however far-fetched the underlying relation; in particular, given the versatile nature of concepts b', we may safely assume that application of a meaning-change function ε directly results in a concept that is a lexical meaning of the result unit, rather than resulting in a concept that is only ‘implied by’ such a meaning (an approach followed in Nolda 2012b, which now turns out to be unnecessary for determinative compounding processes in S; nor have I come upon any other case where ‘underspecified meanings’ would be required, or even helpful, in dealing with word-formation processes). Types of relations may still be distinguished, and relations may be established that are typically used.

3.4 c. Summary of the *dóor*$_1$ *lock*$_2$-example

The functional character of stem-form compounding in any idiolect system S now appears clearly. We are dealing with a function $\varphi = \alpha$ in S whose arguments consist of a basis and a construction mode as just exemplified. The added triple has a non-empty lexical meaning, which in the example is guaranteed by not having used an affix form. Added triples with a non-empty lexical meaning are generally required for the compounding process in any S. The word-formation process in the example is used for the formation of a determinative compound. The meaning-change function is construed as a function that is the specification of the basic meaning by the added concept; a two-place relational concept is newly introduced by the meaning-change function itself. This approach can apparently be generalized to the formation of all determinative compounds of a certain basic type.

The word-formation process used in the example, stem-form compounding in S, results in a fully specified stem-form of S; for this reason, it is a morphological word-formation process in S. Analogously, word-form compounding in S, to which we now turn, is syntactic because its results are fully specified word forms, which are fully specified syntactic units.

3.5 Word-form compounding (syntactic compounding) in S: *lòck*$_1$ *óut*$_2$

Consider a sentence like:

(15) I lock you out from my house.

(The *from*-group is taken to be a complement, a ‘prepositional object’.) This can be uttered as, among other things, a threat whose content may be paraphrased as (*you*: one person):

(16) Speaker is going to lock all entrances to his or her house so that the addressee has to be on the outside.

There are good syntactic and semantic reasons—such as possible modifiers for *lock*$_2$ *out*$_4$ jointly, rather than for *lock*$_2$ or *out*$_4$ separately (apparent counterexamples notwithstanding), and properties of sentence-meaning composition—to postulate a single verb form *lòck*$_1$ *oút*$_2$, with the centre *lòck*$_1$ and periphery *oút*$_2$.[26] The verb form *lòck*$_1$ *oút*$_2$ is the first component in a fully specified *syntactic* unit, once again the *result* of applying a function φ:

(i) ⟨f, J, b⟩
  = ⟨*lòck*$_1$ *oút*$_2$,
  {$\text{Unsp}_{\text{Pers}}$, Unsp_{VN}, $\text{Unsp}_{\text{Mood}}$, Pos_{V}, Minus-*do*, Pres, Act, Non-Cont},
  ·lock somebody out from some place·⟩[27]

The *basic triple* may be assumed as follows:

[26] A recent overview of research on particle verbs in English and German, with a thirty-page bibliography, still incomplete, is Étoré (2007), relevant to both the present Section and to Sec. 4.2., below; see also Booij (2010: Ch. 5), with special reference to Dutch, and the literature quoted there. The ‘phrasal’ vs. the ‘lexical’ nature of particle verbs continues to be a disputed topic, but there is no space here for a detailed discussion of the literature.

[27] In (i), ‘Unsp’ abbreviates ‘Unspecific’; Pos_{V} = the set of positive verb forms (as opposed to forms such as *has nt locked out*), Minus-*do* = the set of verb forms that do not exhibit *do*-periphrasis, but still allow it. See also fn. 7, above.

(ii) $\langle f_1, J_1, b_1\rangle = \langle$*lóck*$_1$, J, ·lock$_3$·$\rangle$

Here, J_1 = J in (i). The concept ·lock$_3$· is as defined in (10), Sec. 2.3. As an *added triple*, we have:

(iii) $\langle f_2, J_2, b_2\rangle = \langle$*oút*$_1$, {Pf}, ·on the outside of·$\rangle$

What is used here, is the only element $\langle$*oút*$_1$, {Pf}$\rangle$ of the improper paradigm of the adverb $\langle$*oút*^{P}, ·on the outside of·$\rangle$, or $\langle$*oút*^{P}, ·out$_1$·$\rangle$, for short.

The *basic shortening* β_{11} is zero-shortening, and the *basic form-change* β_{12} is the function *stress-reduction in* S:

(iv) $\beta_{11} = 0_S$; β_{12} = stress-red$_S$

red$_S$ is the function that assigns to a sequence f' the sequence f" obtained from f' by replacing the indicator of primary word stress, {H}, by the indicator of secondary word stress, {H, L}, in the intonation structure of any member of f' whose intonation structure contains {H} (there must be at least one such member).

The *added shortening* β_{21} is, once again, zero-shortening in S; the *added form-change* β_{22} is identity, and so is the *category-change function* δ. The *meaning-change function* ε may be identified with a result function, to be called 'res':

(v) ε = res[ult],

defined as follows:

(17) Let b_1 be a three-place action concept and b_2 a two-place concept.
res$(b_1, b_2) =_{df}$ the property of being a perception or conception z such that $\{RES(b_1, b_2)\} \subseteq$ the content of z, where

RES$(b_1, b_2) =_{df}$ the attribute of being an $\langle x_1, x_2, x_3, x_4\rangle$ such that:
a. $\langle x_1, x_2, x_4\rangle \in {}^{e3}b_1$,
b. $\langle x_3, x_4\rangle \in {}^{e2}b_2$,
c. the state-of-affairs that $\langle x_3, x_4\rangle \in {}^{e2}b_2$ is a result of x_1.

This applies to our example as follows: RES(·lock$_3$·, ·out$_1$·) = the attribute of being an $\langle x_1, x_2, x_3, x_4\rangle$ such that: $\langle x_1, x_2, x_4\rangle \in {}^{e3}$·lock$_3$· and $\langle x_3, x_4\rangle \in {}^{e2}$·out$_1$· and the state-of-affairs that $\langle x_3, x_4\rangle \in {}^{e2}$·out$_1$· is a result of x_1. Informally: x_1 is an action by x_2 of locking all entrances to x_4 with the result—intended or not—that x_3 is on the outside of x_4. We may now define concept b in (i) as follows:

(vi) ·lock somebody out from some place· [·lock out$_1$·] $=_{df}$ res(·lock$_3$·, ·out$_1$·)

This is a four-place concept, by (17).

In summary, we may again assume that we are confronted with a function φ that takes a basis $\langle f_1, J_1, b_1, f_2, J_2, b_2\rangle$ and a construction mode $\langle \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \gamma, \delta, \varepsilon\rangle$, and assigns to them a fully specified word form (form of a lexical word) $\langle$f, J, b$\rangle$, a function that is analogous to stem-form compounding in S and does not allow the added concept b_2 to be empty (cf. non-empty ·out$_1$·): we are dealing with *word-form compounding* in S because the *result triples*—not necessarily the basic or added triples—are fully specified word forms. This is *syntactic compounding* in S because word forms (forms of lexical words) are syntactic units.[28]

The *dóor lock*- and *lòck óut*-examples are cases of standard non-copulative compounding (determinative compounding). It is not immediately clear if or how the functions involved in copulative compounding fit the same mould.

3.6 Dealing with copulative compounding in S

3.6 a. The problem

This is not the place for a more thorough analysis of the various types of copulative compounding discussed in the literature (including 'co-compounding': Arcodia, Grandi & Wälchli 2010). Only one problem raised specifically by the copulative compounding process will be considered here, a problem whose solution is crucial to the conception of word-formation processes that is being proposed.

Consider the following fully specified word forms of English idiolect systems S:

(i) ⟨*paìnter*$_1$ *póet*$_2$, {Unsp$_{\text{Case}}$, Sg$_{\text{N}}$, Unsp$_{\text{Def}}$}, ·painter-poet·⟩,
(ii) ⟨*paìnter*$_1$ *pòet*$_2$ *scúlptor*$_3$, {Unsp$_{\text{Case}}$, Sg$_{\text{N}}$, Unsp$_{\text{Def}}$}, ·painter-poet-sculptor·⟩[29]

Word form (i) may be obtained by applying the function of word-form compounding (syntactic compounding) in S to a basis consisting of a *paínter*-triple and a *póet*-triple, the first being the basic triple and the second the added triple rather than the other way around.[30]

How do we obtain word form (ii)? Obviously, there are three, not two, fully specified word forms to start from, one with *paínter*$_1$, one with *póet*$_1$, and the third with *scúlptor*$_1$. This might require a basis of three not two triples, throwing into doubt the very distinction between a basis and an added triple.

Generally, we are confronted here with the problem of how to deal with cases where a function of copulative (word-form or stem-form) compounding in S appears to apply to

---

[28] There are other kinds of syntactic compounding, such as, in German and Dutch systems S, syntactic compounding that results in 'noun + verb combinations' like German *auto fahren* ·drive (a car)·, discussed by Booij (2010: Ch. 4) as 'quasi-noun incorporation'. So-called phrasal names like German *heiliger vater* ·Holy Father· are extensively discussed for Dutch and German by Booij (2010: Ch. 7) (English *holy father* ·Holy Father· may be subsumed here). Booij considers them as 'syntactic compounds'; I would rather suggest syntactic conversion as the underlying process. For further discussion, see the special issue of *Word Structure*, 2.2 (2009); cf. Schlücking & Hüning (2009).

[29] It is assumed here that we are confronted with sequences of, respectively, two and three phonological words: *paìnter*$_1$ *póet*$_2$; *paìnter*$_1$ *pòet*$_2$ *scúlptor*$_3$, not with unit sequences of a single phonological word: **paìnterpóet*$_1$, **paìnterpòetscúlptor*$_1$. Both the two-member sequence and the three-member sequence are forms of lexical words.

[30] *paìnter*$_1$ in (i) will then be the centre and *póet*$_2$ the periphery of *paìnter*$_1$ *póet*$_2$. This is a case of centre and head not coinciding: on a traditional notion of head (cf. Scalise & Fabregas 2010), the head in (i) would be *póet*$_2$ not *paìnter*$_1$. The stress assignment—secondary stress on the centre but primary stress on the periphery, which is on the right—agrees with the results obtained by Bell & Plag (2012) for (non-copulative?) compounds: "In general, the more informative N_2 is, the more likely is the compound to be right-stressed." (2012: 516). In copulative compounding, the centre of a compound form may be associated with greater relevance (below, (19)) but the periphery with greater informativeness (Sec. 3.6 d, below), and this is marked in English—as it is in German—by primary stress on the periphery.

a basis of more than two triples (the problem does not arise with non-copulative compounding).

3.6 b. The solution

There is a fourth fully specified word form involved in (ii): the *paìnter*$_1$ *póet*$_2$-triple in (i); the *paìnter*$_1$ *pòet*$_2$ *scúlptor*$_3$-triple in (ii) may be obtained by applying word-form compounding in S to a bipartite basis consisting of the *paìnter*$_1$ *póet*$_2$-triple as a basic triple and a *scúlptor*$_1$-triple as the added triple. We arrive at (ii) in two steps.

First, word-form compounding in S is applied to obtain the *paìnter*$_1$ *póet*$_2$ word form in (i):

(iii) wf-comp$_S$(*paínter*$_1$, J_1, ·painter·, *póet*$_1$, J_2, ·poet·, zero-short$_S$, stress-red$_S$, zero-short$_S$, identity, $^\frown$, identity, ε) = (i),

where $J_1 = J_2 = \{Unsp_{Case}, Sg_N, Unsp_{Def}\}$; the meaning change ε remains to be specified.

We next apply word-form compounding in S to (i) and a *scúlptor*$_1$-triple:

(iv) wf-comp$_S$(*paìnter*$_1$ *póet*$_1$, J_1, ·painter-poet·, *scúlptor*$_1$, J_2, ·sculptor·, zero-short$_S$, stress-red$_S$, zero-short$_S$, identity, $^\frown$, identity, ε) = (ii),

where J_1, J_2, and ε are as before. In both (iii) and (iv) the same function, word-form (syntactic) compounding in S, is applied, each time to a bipartite basis; the second application uses the result of the first as a starting-point. Obviously, such recursive application could be repeated any number of times, yielding four-member word forms, five-member word forms etc.

Generally, the following *solution* to the problem of a bipartite basis is proposed (analogously, in part of the literature):

(18) Traditional copulative compounding of more than two word forms or more than two stem forms is reconstructed as recursive application of word-form compounding in S or stem-form compounding in S, each construed as a function that applies to a bipartite basis consisting of a basic triple and an added triple; the recursive application uses a single meaning-change function ε of a specific type.

As an alternative, *n* functions wf-comp^n(S) or st-comp^n(S) could be considered, for $n \geq 2$: functions that apply to a basis consisting of, respectively, 2, 3, 4, etc. triples. The alternative may be rejected, provided solution (18) is tenable.

This depends on suitable meaning-change functions being available; *copulative word-form compounding in* S (cop-wf-comp$_S$) and *copulative stem-form compounding in* S (cop-st-comp$_S$) may then be defined as, respectively, wf-comp$_S$ and st-comp$_S$ that uses an appropriate meaning-change function. I will now argue that types of suitable functions ε do exist. (The argument and its result are summarized in Sec. 3.6 e, below.)

3.6 c. Meaning-change functions for copulative compounding in S: *and*-functions

It is standard practice to assume that copulative compounding typically involves a 'semantic and', in some sense, by which result meanings are obtained. Attempts to make this idea more precise are rare and, I believe, not yet entirely successful.

The following definition of '*and'* (term in italics) turns out to provide a suitable basis for the meaning change functions ε that figure in typical cases of copulative compounding; as an example, the change function ε in (iii) and (iv) will be identified on the basis of the definition:

(19) Let b_1 and b_2 be one-place concepts, and b' a two-place concept.[31]
and(b')(b_1, b_2) $=_{df}$ the property of being a perception or conception z such that {AND(b')(b_1, b_2)} is a subset of the content of z, where

AND(b')(b_1, b_2) $=_{df}$ the property of being an x such that, for some x_1 and x_2:

a. $x_1 \in {}^{e1}b_1$:
x_1 is in the extension of b_1;
b. $x_2 \in {}^{e1}b_2$:
x_2 is in the extension of b_2;
c. $\langle x, \{x_1, x_2\}\rangle \in {}^{e2}b'$:
the pair consisting of x and the set $\{x_1, x_2\}$ is in the extension of b';
d. $\langle [x_1 \in {}^{e1}b_1], [x_2 \in {}^{e1}b_2], x\rangle \in {}^{e3}$·more relevant·:
the state-of-affairs that x_1 is in the extension of b_1 is more relevant with respect to x than the state-of-affairs that x_2 is in the extension of b_2.

By an ‘*and*-function’ I understand any function ε such that, for some concept b', ε = *and*(b'). The meaning-change functions ε used in (word-form or stem-form) compounding in S are, typically, *and*-functions: this determines their type.

In the *paìnter*$_1$ *póet*$_2$-example (iii) and the *paìnter*$_1$ *pòet*$_2$ *scúlptor*$_3$-example (iv), ε is as follows:

(v) ε = *and*(·mult-id·)

·mult-id·, or ·multiple identity·, is a two-place concept whose extension is the set of pairs ⟨x, y⟩ such that y is a non-empty set and x is identical to every element of y (which implies that y = {x}).

In (iii), the function ε in (v) applies in the following way. Replacing b' in (19) by ·mult-id·, b_1 by ·painter·, and b_2 by ·poet·, we obtain:

(20) *and*(·mult-id·)(·painter·, ·poet·) = the property of being a perception or conception z such that {AND(·mult-id·)(·painter·, ·poet·)} is a subset of the content of z, where

AND(·mult-id·)(·painter·, ·poet·) = the property of being an x such that, for some x_1 and x_2:

a. $x_1 \in {}^{e1}$·painter·:
x_1 is a painter;
b. $x_2 \in {}^{e1}$·poet·:
x_2 is a poet;
c. $\langle x, \{x_1, x_2\}\rangle \in {}^{e2}$·mult-id·:
x is identical to x_1 and to x_2;
d. $\langle [x_1 \in {}^{e1}$·painter·$], [x_2 \in {}^{e1}$·poet·$], x\rangle \in {}^{e3}$·more relevant·:

[31] This may have to be generalized.—b' is a concept for, roughly, forming a complete entity out of parts, which justifies ‘and’: ·multiple identity·, ·mereological sum·, ·group·, ·pair·, and similar concepts. But this need not be specified in the definition beyond the inclusion of condition (c).

x_1 (= x) being a painter is more relevant with respect to x than x_2 (= x) being a poet.

(a) to (d) mean, informally, that x is a painter and a poet, and being a painter is more relevant with respect to x than being a poet.

The result meaning ·painter-poet· in the *paìnter*$_1$ *póet*$_2$-example (iii) may now be identified as follows:

(vi) ·painter-poet· $=_{df}$ *and*(·mult-id·)(·painter·, ·poet·)

·painter-poet· is a concept; its (one-place) intension consists of the property identified in the AND-part of (20); its extension is the set of all x that have this property.

In the same way we identify the result meaning ·painter-poet-sculptor· in the *paìnter*$_1$ *pòet*$_2$ *scúlptor*$_3$-example (iv), using ·painter-poet· as defined in (vi) as a basic meaning:

(vii) ·painter-poet-sculptor· $=_{df}$ *and*(·mult-id·)(·painter-poet·, ·sculptor·).

3.6 d. Comments

The notion of *and*-function based on (19) characterizes the type of semantic change functions ε in all standard types of copulative compounding in a traditional sense, not only the functions ε in the copulative N+N type from which examples, such as the *paìnter*$_1$ *póet*$_2$- and *paìnter*$_1$ *pòet*$_2$ *scúlptor*$_3$-examples, are usually drawn.

For instance, we also have an adjectival type in German, as in *blaù*$_1$ *weíß*$_2$ ·blue and white· and *blaù*$_1$ *weìß*$_2$ *rót*$_3$ ·blue, white and red·, applied to flags and other objects consisting of parts that each have a distinct colour (in English, phrases using the conjunction *and* appear to be preferred in these cases). We are dealing here either with word-form or with stem-form compounding in S (a decision is not relevant to the problem under discussion). The meaning-change function ε is *and*(·sum·), where *and* is as in (19) and ·sum· is the two-place concept of a mereological sum; $\langle x, \{x_1, x_2\}\rangle \in {}^{e2}$·sum· means that x is a (the) sum of x_1 and x_2 (is the sum of two parts that make up x).

It may be questioned that a concept like ·multiple identity·, used in (20), is actually needed for identifying the meaning-change function in the *paìnter*$_1$ *póet*$_2$-example. This is no longer true of ·sum· in the *blaù*$_1$ *weíß*$_2$-example. The concept ·multiple identity· now turns out to be a special case of b' in (19), the definition of ‘*and*’.

The *blaù*$_1$ *weíß*$_2$-example also motivates the relevance condition (d) in (19) and (20): a flag that is striped blue and white is different from a flag that is striped white and blue; there is an ordering of stripes such that the blue stripe comes first in a blue-and-white flag and the white stripe comes first in a white-and-blue flag. Being the blue stripe is, in this sense, more relevant with respect to the blue-and-white flag than being the white stripe, and conversely for the white-and-blue flag.[32]

---

[32] The concept ·blue-and-white· involves a relevance condition as in (19d); so does the concept ·white-and-blue·; and one condition is the opposite of the other: these are facts of lexical meaning with respect to *blaù*$_1$ *weíß*$_2$ and *weìß*$_1$ *blaú*$_2$, hence, facts of the meaning of suitable sentences with these forms as constituents. When a suitable sentence with *blaù*$_1$ *wéiß*$_2$ is uttered, the greater relevance of being the blue part of x vs. being the white part (whatever x is in the utterance situation) becomes a part of the utterance meaning that is determined by the meaning of the sentence. However, what the greater relevance of being the blue part of x consists in, such as being the left-most stripe of a vertically striped flag, is not determined by the meaning of

Even in the *paìnter*$_1$ *póet*$_2$-example there is evidence, easily overlooked, for a relevance condition as in (20d). Otherwise, discussions of the following kind would not make sense: ‘X was a paìnter-póet.—No, X was a pòet-paínter: a poet who was also a painter, but more of a poet than a painter.’ (The semantic difference is not due to contrastive accentuation, as it would be if ‘paìnter scúlptor’ were opposed to ‘paìnter póet’.)

Generally, *and*-functions, like *and*(·sum·) or *and*(·mult-id·), represent a kind of ‘ordered and’, not an ‘unordered’ one, which excludes a commutative view of copulative compounding, contrary to what tends to be implied in the literature.

3.6 e. Summary

Copulative compounding raises the problem of how to deal with cases where more than two fully specified word forms or stem forms appear to constitute a basis. A solution was formulated in (18) that would allow us to retain a bipartite basis also in such cases. The solution was conditional on appropriate meaning-change functions being available. I argued that *and*-functions defined on the basis of (19) are appropriate for standard types of copulative compounding (the generality of the relevance condition (19d) remains to be established, though).

I have checked all types of compounding reported in the literature that are clearly compounding and are clearly copulative (see also Štekauer, Valera & Körtvélyessy 2012: 88-92); some types *claimed* to be copulative compounding should be excluded as such, either for not being compounding: *móther*$_1$ *chíld*$_2$ is not a form of a compound but occurs only as part of a compound form, or for not being copulative: *óak*$_1$ *tree*$_2$ is ‘determinative’. I have come up with a single problem case for the assumption that *and*-functions may be sufficient: in view of certain kinds of co-compounding (such as the ‘brother-sister = siblings’ type, see Arcodia, Grandi and Wälchli 2010: Secs 2 and 3), it appears necessary to also allow *or*-functions based on a definition of ‘*or*(b')(b_1, b_2)’ that is by and large analogous to (19), the definition of ‘*and*(b')(b_1, b_2)’.

The solution proposed in (18) to the problem of a non-bipartite basis in copulative compounding in S is now accepted. We next turn to word-formation processes in S other than compounding processes.

---

the sentence but is a part of the utterance meaning that is provided by situational context; it is, in this sense, ‘pragmatic’. This could possibly be made more precise using notions from ‘Relevance Theory’ (where ‘Relevance’ is unrelated to my term ‘more relevant’), cf. Carston (2006); we may also consider applying the sentence semantics of Lieb (1983), which is not truth-conditional and is utterance-oriented.

4 Other processes

4.1 Stem-form derivation in S (st-der$_S$): *un*$_1$ *lóck*$_2$

4.1 a. Example

In the sentence

(21) My friend unlocked the door.

a verb is used whose stem lexeme has, among other forms, *un*$_1$ *lóck*$_2$, fully specified as:

(i) $\langle$f, J, b$\rangle$ = $\langle$*un*$_1$ *lóck*$_2$, {Inf-St, Pres-St}, ·undo all locking·$\rangle$[33]

The triple in (i) is to be the result of applying stem-form derivation in (English idiolect systems) S; this may be construed along the same lines as stem-form compounding.

First, we assume the following *basic triple*:

(ii) $\langle f_1, J_1, b_1\rangle$ = $\langle$*lóck*$_1$, {Inf-St, Pres-St}, ·lock$_2$·$\rangle$,

where ·lock$_2$· is the concept of locking defined in (9), Sec. 2.3, which was obtained from ·lock$_1$·, the concept of a lock defined in (8).

The basic triple in (ii) is a fully specified stem-form, just like the basic triple in our example of stem-form compounding. The difference from stem-form compounding arises with the *added triple*:

(iii) $\langle f_2, J_2, b_2\rangle$ = $\langle$*un*$_1$, {Af}, $b^0\rangle$

un$_1$ is a form of a *derivational* affix. We do not allow categories such as 'Derivational-Affix form in S' (in contrast to 'Derivational Affix in S'). It is, however, sufficient to state in (iii) that the added triple is a fully specified *affix* form to exclude stem-form compounding, where the added triple must be a fully specified (morphological or syntactic) unit with a non-empty lexical meaning, hence, cannot be an affix form, and to exclude stem-form conversion, where the added triple is not even a fully specified unit.

We may also identify a *construction mode*. The *arrangement function* γ is inverse concatenation (this exemplifies prefixation, since we are dealing with an affix form). The *shortening functions* β_{11} and β_{21} are zero-shortening in S; the *form-change functions* β_{12} and β_{22} are identity, and so is the *category-change function* δ. The *meaning-change function* ε is harder to determine.

As utterances of (21) would show, there must be a previous act of locking whose result continues to be relevant. Roughly, unlocking a door consists in deactivating ('releasing') all relevant locks that have been, and still are, locked. We may therefore call *annulling* the function ε that takes us from the concepts ·lock$_2$· in (ii) and b^0 in (iii) to the con-

[33] It is well known that there are 'several *un*' in English, of different origin and with difficult word-stress behaviour; *un* in *unlock* is unstressed according to the *English Pronouncing Dictionary* (Roach et al. eds 2003). Assuming *ùn*$_1$ instead of *un*$_1$ would leave the following analysis essentially unaffected.

cept ·undo all locking· in (i). We must presuppose a concept of ‘appliance activation’, like ·lock$_2$·, so that reference can be made to the *purpose* of an appliance.

Unlocking a door x_3 then consists in an action x_1 by an actor x_2 such that: through some action x_1' by some actor x_2', the door x_3 was successfully locked, and continues to be so at x_1; and for every lock x of the door x_3 involved in the purpose of any relevant locking action x_1' by any actor x_2', it is true that the action x_1 by actor x_2 consists in deactivating x. (If this sounds complicated, it simply reflects the complexity of concepts that involve appliance use.)

The meaning-change function ε may now be defined as follows:

(22) Suppose that b_1 is a three-place concept of ‘appliance activation’, and $b_2 = b^0$.
annul(b_1, b_2) $=_{df}$ the property of being a perception or conception z such that {ANNUL(b_1, b_2)} $\subseteq$ the content of z, where

ANNUL(b_1, b_2) $=_{df}$ the attribute of being an $\langle x_1, x_2, x_3\rangle$ such that:
a. for some x_1' and x_2',
(i) $\langle x_1', x_2', x_3\rangle \in {}^{e3}b_1$,
(ii) x_1' is successful at the beginning of x_1;
b. for all x_1', x_2', x_4, and x, if
(i) x_1' and x_2' are as in (a.i) and (a.ii), and
(ii) $\langle x_1', x_2', x_3, x_4\rangle$ satisfies the purpose of x,
then
(iii) x_1 is an action by x_2 of deactivating x.

We may now *define*: ·undo all locking· $=_{df}$ annul(·lock$_2$·, b^0). A different concept of unlocking, based on ·lock$_3$· in (10) and obtained through a slightly different function of annulling, is used in *My friend unlocked the house.*, which may be paraphrased as ‘My friend undid all locking of some entrance to the house.’

In summary, the *un*$_1$ *lóck*$_2$-example demonstrates how stem-form derivation in S may be construed as a function of essentially the same type as stem-form compounding in S, the differences residing in the added triple.[34]

4.2 Word-form derivation (syntactic derivation) in S (wf-der$_S$): *lòck*$_1$ *úp*$_2$

4.2 a. The result and its basis

As a sample sentence, take:

(23) I lock the door up.

This may be paraphrased as, roughly:

(24) Speaker is going to securely lock the door.

The word form *lòck*$_1$ *úp*$_2$ used in (23) is to be obtained by applying a word-form process, with the following *result triple*:

[34] Arguably, the added triple may be from some system other than S, provided the basic triple is from S.

(i) ⟨f, J, b⟩
= ⟨*lòck*$_1$ *úp*$_2$,
{Unsp$_{Pers}$, Unsp$_{VN}$, Unsp$_{Mood}$, Pos$_V$, Minus-*do*, Pres, Act, Non-Cont},
·lock securely·⟩[35]

There is no problem with identifying the *basic triple*:

(ii) ⟨f$_1$, J$_1$, b$_1$⟩ = ⟨*lóck*$_1$, J, ·lock$_2$·⟩

What about the *added triple*?

It may be suggested that we treat *lòck*$_1$ *úp*$_2$ on the pattern of *lòck*$_1$ *oút*$_2$: the *lòck*$_1$ *oút*$_2$-triple in (i) of Sec. 3.5 was taken to be the result of applying word-form or syntactic *compounding* (assuming, in Sec. 3.5 (ii), the basic meaning ·lock$_3$· not ·lock$_2$·); so why not construe the *lòck*$_1$ *úp*$_2$-triple in (i) also as a word-form compounding result, and assume the added triple accordingly?

The problem is the semantic difference between *oút*$_1$ as used in (15) *I lock you out from my house*, and *úp*$_1$ as used in (23) *I lock the door up*: we may indeed assume a non-empty concept, ·on the outside of·, as a lexical meaning of *oút*$_1$ given {Pf}, as required by word-form compounding with a particle form; ·securely·, though, can hardly be construed as a lexical meaning of *úp*$_1$ given {Pf}, and none of the non-empty lexical meanings that *úp*$_1$ does have given {Pf} is relevant in the context of (23); only the empty concept b^0 remains. We thus arrive at the following added triple:

(iii) ⟨f$_2$, J$_2$, b$_2$⟩ = ⟨*úp*$_1$, {Pf}, b^0⟩

Remark on the added category set. *úp*$_1$ in (iii) is a form of a *derivational* particle but, again, this need not be made explicit: it is sufficient in (iii) to state that the added triple is a fully specified particle form *with an empty meaning*. This excludes syntactic compounding, where the added triple may well be a fully specified particle form but must then have a *non-empty* meaning (such as ·out$_1$· in Sec. 3.5), and also excludes syntactic conversion, where the added triple is not a fully specified unit.

4.2 b. The construction mode

From (ii) and (iii), we continue as in the *lòck*$_1$ *oút*$_2$-example: the basic and added shortenings β_{11} and β_{21} are zero-shortening in S; the basic form-change β_{12} is *stress-reduction in* S (stress-red$_S$), the added form-change β_{22} is identity; the arrangement function γ is concatenation; and the category-change δ, identity.

The semantic function ε is different, though, from ε in the *lòck*$_1$ *oút*$_2$-example: it must introduce the 'security'-feature of the result meaning. Intuitively, any secure locking is a locking x_1 by an actor x_2 of an object x_3 (e. g., a door) such that the result of x_1 is secure. Now locking means that ⟨x_1, x_2, x_3⟩ together with some x_4 satisfies the purpose of some lock x. This again means (see (8)) that x_1 is an action by x_2 of activating x, and the immediate result of x_1 is: creation, through x_1, of a fixed but releasable connection between x_3 and some x_5 with the effect of barring some access to x_4 by x_3. *This result, then, the creation of the connection, is secure.*

---

[35] For the set J, see fns 7 and 27.

The meaning-change function ε is defined accordingly, as a special case of the following definition schema:

(25) Suppose that b_1 is an n-place action concept, n > 2, and $b_2 = b^0$.
secure-res$^n(b_1, b_2) =_{df}$ the property of being a perception or conception z such that $\{\text{SECURE-RES}^n(b_1, b_2)\} \subseteq$ the content of z, where

SECURE-RES$^n(b_1, b_2) =_{df}$ the attribute of being an $\langle x_1, x_2, \ldots, x_n\rangle$ such that:
a. $\langle x_1, x_2, \ldots, x_n\rangle \in {}^{en}b_1$;
b. the result of x_1 is secure.

We now *define*: ·lock securely· $=_{df}$ secure-res^3(·lock$_2$·, b^0).[36]

4.2 c. Summary

The triple ⟨f, J, b⟩ in (i) may be taken to be the result of applying a word-form process φ = α in S to a basis that consists of the basic triple $\langle f_1, J_1, b_1\rangle$ in (ii), the added triple $\langle f_2, J_2, b_2\rangle$ in (iii), and the construction mode $\langle \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \gamma, \delta, \varepsilon\rangle$ in Sec. 4.2 b, where ε = secure-res^3.

The function φ is analogous, in all relevant respects, to stem-form derivation in S, as appears from a comparison of φ with the example in Sec. 4.1 for stem-form derivation. In particular, the particle form *úp*$_1$ in (iii), with its empty lexical meaning b^0, corresponds to the affix form *un*$_1$ in (iii) of Sec. 4.1, again with its empty meaning, and the semantic function secure-res^3 in (25) may be construed as a 'grammatical meaning' correlated with *úp*$_1$, just as the function annul in (22) is a 'grammatical meaning' correlated with *un*$_1$. Due to these correspondences, the function α in S may be identified with *word-form derivation in* S (wf-der$_S$), or *syntactic derivation in* S. In word-form derivation, derivational particles such as ⟨*úp*$_1^P$, b^0⟩, where *úp*$_1^P$ = {⟨*úp*$_1$, {Pf}⟩}, are analogous to derivational affixes in stem-form derivation, such as ⟨*un*$_1^{LP}$, b^0⟩, where *un*$_1^{LP}$ = {⟨*un*$_1$, {Af}⟩}.[37]

---

[36] It may be argued that secure-result functions are unduly specific. A single meaning-change function for all relevant uses of the particle may indeed be suggested, a function that assigns concepts somewhat like 'going through with the action to its natural end' (compare Jackendoff 2002b: 76, on sample sentences with *up* as an 'aspectual particle': "This means roughly 'V NP completely'"). My argument for syntactic derivation would remain unaffected. Still, proposals for a single function meant to cover *all* relevant *up*-uses (such as Etoré 2007: 119) turn out to be quite vague. It should be more adequate to assume several related functions, or families of functions. In particular, secure-result functions apply to all concepts of creating a fixed connection, as in *fasten/fix/link/lock/nail/seal/tie up*.—See also Rich (2003) for the complex semantics of particle verbs in German.

[37] The *derivational particles* of a given idiolect system are the 'core particles' of the system whose forms may be used for added triples in syntactic derivation. Core particles, in the sense of *Kernpartikel* in Budde (2000), are, roughly, the particles ⟨P, b⟩ with $b = b^0$ that cannot figure in the occurrence of any basic grammatical relation (nucleus or head, complement, modifier, coordination, and syntactic topic: see Lieb 1993b, 2011; Nolda 2007: Sec. 7.3.7). Derivational particles are recognized explicitly by Drude (2004: Sec. 6.4): *modale Derivationspartikel* in Guaraní. In Drude (2010), a case is made for derivational verbs. (For further discussion, see Sec. 5.3 b, below.) I restrict the use of the term 'particle verb' to verbs that involve a particle in the narrow sense, in contrast to a much broader usage found in the literature, e. g., in Müller (2002: Ch. 6). The broader usage is critically discussed already in Lüdeling (2001).

4.3 Stem-form conversion in S (st-conv$_S$): from lock to locking

4.3 a. Result triple and basic triple

It is an acid test for any theory of word formation how it manages to deal with conversion, in a sense where conversion is to be a word-formation process or type rather than simply recategorization.[38] Once again, a ‘lock’-example will be used in arguing that the present conception passes the test: there is a function $\varphi = \alpha$ in S, of the kind we are discussing, that can be applied to yield (i) as a *result*:

(i) $\langle$f, J, b$\rangle$ = $\langle$*lóck*$_1$, {Inf-St, Pres-St}, ·lock$_2$·$\rangle$,

where ·lock$_2$· is the concept of locking defined in (9), Sec. 2.3.

As a *basic triple*, we choose:

(ii) $\langle$f$_1$, J$_1$, b$_1$$\rangle$ = $\langle$*lóck*$_1$, {Unsp$_{Case}$-St, Sg$_N$-St, Unsp$_{Def}$-St}, ·lock$_1$·$\rangle$,

where ·lock$_1$· is the concept of a lock defined in (8), Sec. 2.3. It is intuitively clear from the definitions (8) and (9) that there is a systematic way, i. e., a function yet to be specified, for getting from ·lock$_1$· to ·lock$_2$·.

4.3 b. Added triple and arrangement function

The decisive question is this: what exactly is the added triple, and what is the arrangement function? The *added triple* will be assumed as in:

(iii) $\langle$f$_2$, J$_2$, b$_2$$\rangle$ = $\langle$f^0, Ø, b^0$\rangle$,

where f^0 = the empty sequence = the empty set, and Ø, too, is the empty set. The arrangement function is:

(iv) $\gamma = {}^\frown$ = concatenation

Assumptions (iii) and (iv) appear to open up a notorious minefield: are we proposing that f$_1$ = *lóck*$_1$ should be concatenated with an empty morphological unit, possibly, an empty affix form? The arguments against ‘empty affix forms’, or empty units in general, are indeed overwhelming.

But no such unit is introduced in (iii): the empty sequence f^0 (= the empty set) has *no linguistic status*; in particular, since it is not a linguistic unit, the question of how it should be categorized does not arise.

Accordingly, J$_2$ in (iii) is the empty set, and asking if the empty concept b^0 in (iii) is a lexical meaning of f^0 is meaningless. As a consequence, *the added triple* in (iii) is *not* a

[38] I here presuppose the recent comprehensive review of conversion research, with special reference to English, to be found in Balteiro (2007a), (2007b); see also Manova (2011: Ch. 3). I uniformly use the term 'conversion', avoiding 'zero-affixation' but still accounting for relevant phenomena discussed under that heading.—An advanced treatment of conversion is Nolda (2012b: Part C), for German, as yet unpublished. For a critical evaluation of Nolda’s general framework, which is related to the Process Model, see below, Sec. 8. Despite a different conception of conversion, his analyses could be replicated in the Process Model. (I would not subscribe, though, to his semantic analyses of instances of noun-to-verb conversion.)

fully specified form; it is an ‘empty shell’ of such a form, which has no linguistic status other than serving as an added triple.

The importance of f_2 being f^0 appears once the arrangement function γ is identified with concatenation: on purely logical grounds, the empty sequence is the *identity element* for concatenation (of sequences); i. e. for any sequence f, $f \frown f^0 = f^0 \frown f = f$ (as already pointed out in Sec. 2.2). Therefore, f_1 in (ii) concatenated with f_2 in (iii) = *lóck*$_1 \frown f^0$ = *lóck*$_1$.[39]

4.3 c. The shortening and form-change functions

The addition f_2 is the empty sequence: this has immediate consequences for the added shortening β_{21}, which must be zero-shortening in S, and the added form change β_{22}, which must be identity. Moreover, in core cases of conversion, the result form f is identical to the basic form f_1, as indeed it is in our example; i. e., the basic shortening β_{11} must also be zero-shortening in S, and the basic form-change function β_{12} identity. But this should not be made into a general requirement on stem-form conversion: we should not exclude as conversion the transition from, say, *sing*$_1$ to *song*$_1$, *fall*$_1$ to *fell*$_2$; or from *pérfect*$_1$ to *perféct*$_1$—a point to be taken up below, in Sec. 5.4 b.

Allowing non-zero shortening and proper form change for the basic unit when a conversion process in S is applied increases the importance of such processes: we may now construe processes of short-word formation, clipping, or even back formation as subcases of the conversion process in S, provided the category change and meaning change are allowed to be identity functions; for details, see Sec. 4.5, below.

4.3 d. Category change and meaning change

In our lock-to-locking example, the *category-change function* δ is not identity but is a function {$\text{Unsp}_{\text{Case}}$-St, Sg_{N}-St, Unsp_{Def}-St}-to-{Inf-St, Pres-St}. However, differently from what is typically required or implied for conversion in the literature, identity should be allowed as the category-change function δ to increase the range of conversion processes.

Next, consider the meaning-change function ε. In our example, we have an action x_1 by actor x_2 of locking (·lock$_2$·) an object x_3 (a door). The action x_1 consists in satisfying the purpose of some lock x (·lock$_1$·), i.e. consists in activating x, with the result of creating, through the activation of the lock, a fixed but releasable connection between x_3 (the door) and some x_5, etc. (see the definitions in (8) and (9), Sec. 2.3). The function ε that takes us from the concept of a lock (·lock$_1$·) to the concept of locking (·lock$_2$·) can be defined as *use-of*:

(26) Suppose that b_1 is an appliance concept, and $b_2 = b^0$.
use-of(b_1, b_2) $=_{\text{df}}$ the property of being a perception or conception z such that {USE-OF(b_1, b_2)} ⊆ the content of z, where

USE-OF(b_1, b_2) $=_{\text{df}}$ the attribute of being an $\langle x_1, x_2, x_3\rangle$ such that for some x and x_4,
a. $x \in {}^{1e}b_1$
b. $\langle x_1, x_2, x_3, x_4\rangle$ satisfies the purpose of x.

---

[39] Strictly speaking, proper subfunctions of concatenation must also be allowed, to account for concatenation with rearrangement, as in English f = *úp*$_1$ *kèep*$_2$ from f_1 = *kèep*$_1$ *úp*$_2$: concatenation of f_1 and f^0 (= f_2) with rearrangement of f_1.

We now *define*: ·lock$_2$· $=_{df}$ use-of(·lock$_1$·, b^0).—It is primarily the semantic function that determines the 'direction' in which a conversion function is to apply. And indeed, defining ·lock$_2$· in terms of ·lock$_1$·, rather than conversely, happens to be historically correct, too.

In summary, the function φ = α in S that is characterized by this example is rightly called *stem-form conversion in* S (st-conv$_S$).[40]

4.4 Word-form conversion (syntactic conversion) in S: from locking-out to locking-out from work-place

4.4 a. Example

It appears desirable to have a word-formation process α in S = φ such that

(i) ⟨*lòck*$_1$ *oút*$_2$, {Inf, Pres, Act, Non-Cont}, ·lock out from work-place·⟩,

is the *result* of applying φ to the argument

(ii) ⟨*lòck*$_1$ *oút*$_2$, {Inf, Pres, Act, Non-Cont}, ·lock out$_1$·, f^0, Ø, b^0, zero-short, id, zero-short, id, ⁀, id, ε⟩,

where ε remains to be specified.

The *result concept* ·lock out from work-place·, to be obtained by applying ε, must be a three-place concept b whose intension consists of the following attribute:

(iii) the RESTRICTION^{3,4} to ·work-place· of ·lock out$_1$· and b^0 = the attribute of being an ⟨x_1, x_2, x_3⟩ such that:
α. x_1 is an action of x_2;
β. for some x_4, ⟨x_4, x_3, x_2⟩ ∈ ^{e3}·work-place· (i.e., x_4 is a work-place of x_3 supplied by x_2);
γ. for all x_4, if ⟨x_4, x_3, x_2⟩ ∈ ^{e3}·work-place·, then the effect of x_1 on x_3 and x_4 is the state-of-affairs that, for some x_1' and x_2', ⟨x_1', x_2', x_3, x_4⟩ ∈ ^{e4}·lock out$_1$· (i. e. given any work-place x_4 of employee x_3 supplied by employer x_2, it is an effect of x_2's action x_1 that somebody—not necessarily x_2—locks x_3 out from x_4).

While this may not yet do justice to all aspects involved, legal and non-legal, it should be roughly adequate, assuming that some physical locking out is necessarily included in an employer's lock-out action. The essential point is this: in order to arrive at our result concept b by means of ε, the *meaning-change function* ε must have access to the concept of

[40] Construing stem-form conversion in S as in Sec. 4.3 is not the only kind of analysis suggested by the literature. As an alternative, we might consider an analysis of stem-form conversion in S—or any other conversion function—that does away with an added triple, leading to one-place word-formation processes in S (Nolda's approach in Nolda 2012b, see below, Sec. 8.6). There is, however, a major empirical reason for identifying the conversion process in S and its subcases with functions φ as above: existence of the *process cline* in which the conversion process in S and its subcases are readily included on our approach (see Secs 1.2 b and 6.3 b) but less readily, if at all, on the alternative analysis (see Secs 8.6 and 5.4 e).—Taking up a suggestion made to her by Lieb, Eschenlohr (1999) proposes a conception of stem-form conversion in a framework that is Integrational but still of the IA type; the conception, which does without empty entities, too, is then applied to a study of conversion in German; its results continue to be of interest.

work-place, not yet available in the basis, see (ii). The problem may be solved by taking ε itself as the value of a function whose arguments are concepts. This function will be called *restriction*^{3,4}; it assigns ε to the concept of work-place:

(iv) $\varepsilon = \text{restr}^{3,4}$(·work-place·)

restr^{3,4}, easily defined on the basis of RESTRICTION^{3,4} as applied in (iii), is a function that takes *three*-place concepts b' (superscript 3) as arguments and assigns to each a function ε that takes concept pairs $\langle b_1, b_2\rangle$ as arguments, where b_1 is a *four*-place concept (superscript 4); the values of ε are *three*-place concepts b.

‘ε’ in (ii) may then be replaced by ‘restr^{3,4}(·work-place·)’, and ‘·lock out from work-place·’ in (i) by ‘restr^{3,4}(·work-place·)(·lock out$_1$·, b^0)’, adopting as a *definition*: ·lock out from work-place· $=_{df}$ restr^{3,4}(·work-place·)(·lock out$_1$·, b^0).

The only function that φ can be identified with is a conversion function, more specifically, word-form conversion in S since the results of φ are to be fully specified word forms.

4.4 b. Objection to a conversion analysis

There is a major objection that can be raised, from a traditional point of view, to assuming a conversion function in (a): ⟨*lòck oút*^{P}, ·lock out from work-place·⟩ belongs to the same part of speech—verb—as ⟨*lòck oút*^{P}, ·lock out$_1$·⟩. True, there is a valency difference: one verb is two-valued, the other three-valued (used with a prepositional group as a complement), but this would not suffice for having conversion in a strict traditional sense.

There is a forty-page review in Balteiro (2007a: Ch. 3) of the positions taken in the literature on this question—necessity of a part-of-speech change; Balteiro herself eventually adopts the traditional requirement (2007a: 114). I draw the opposite conclusion from her discussion: the requirement should be rejected (as recently it has been by Manova 2011: Sec. 3.3.2); otherwise, clear cases—like the one above—that may be treated as word formation are left in a word-formation limbo.

4.4 c. Other examples

The *lòck*$_1$ *oút*$_2$-example discussed in this Section is not yet representative of the full range of word-form or syntactic conversion. Another classical case is provided by the nominalization of adjective forms in German (most recently discussed, and competently analysed, in Nolda 2012b: Sec. 8.2), in contradistinction to the nominalization of infinitives and infinitival groups, which are syntactic units but whose nominalization results in stem forms in German (as shown by Nolda 2012b: Sec. 8.1, contrary to Manova 2011: 113, who discusses ‘syntactic conversion’ more generally for Bulgarian, Russian, and Serbo-Croatian, 2011: 111-121). Also, a large part of idiom formation, typically assigned to phraseology or to lexicology in a broad sense, may be reclaimed for word formation proper once due attention is paid to word-form conversion. For example, the adverb ⟨P, b⟩ = ⟨*òff and ón*^{P}, ·intermittently·⟩ may be obtained from the citation form of the particle group *óff*$_1$ *and*$_2$ *ón*$_3$ meaning ·stopped and then continued· (or rather, obtained from a corresponding pseudo-word). Similarly, phrasal names like *heiliger vater* ·Holy Father· (·Pope·) are obtained by syntactic conversion rather than syntactic compounding. (In Polish, the phrase is *święty ojciec*, A + N; order is reversed in the phrasal name: $f_2 = f^0$ and γ = concatenation with rearrangement.) In addition, there are unexpected word-formation examples due to syntactic conversion, such as, in English, the transition from ⟨*háve*^{P}, ·possess·⟩ to the tense auxiliary ⟨*hàve*^{P}, b^0⟩.

Generally, the range of subcases of the conversion process in S has now been broadened to include processes whose place is traditionally hard to define.

4.5 The range of conversion functions

4.5 a. Cutting and clipping in S

Given the conception of conversion functions and of basic shortening and form-change functions introduced in Sec. 3.2, we may now have instances of stem-form conversion in S like the following one:

(i) st-conv$_S$(*mòbile*$_1$ *phóne*$_2$, {Unsp$_{Case}$, Sg$_N$, Unsp$_{Def}$}, ·mobile phone·, f^0, Ø, b^0, short$_S$(*phóne*$_2$), stress$_S$, zero-short$_S$, id, $^\frown$, J_1-to-J, sem-id)
= ⟨*móbile*$_1$, J, ·mobile phone·⟩,

for J_1 = {Unsp$_{Case}$, Sg$_N$, Unsp$_{Def}$}, J = {Unsp$_{Case}$-St, Sg$_N$-St, Unsp$_{Def}$-St}, and the meaning-change function sem-id, *semantic identity*; this is the function ε whose arguments are any concept pairs ⟨b_1, b_2⟩ such that ε(b_1, b_2) = b_1. Because of a non-zero *basic shortening*, the subcase of st-conv$_S$ exemplified in (i) will be called *stem-form cutting in* S, or *st-cut*$_S$.

There is another subcase of st-conv$_S$ that is similar to st-cut$_S$ but must be distinguished from it; this is *stem-form clipping in* S, or *st-clip*$_S$, where *the basic form-change* β_{12} is a (stress-sensitive) *truncation function* (essentially, 'clipping' in a traditional sense; cf. Berg 2011), as in:

(ii) st-conv$_S$(*jàpanése*$_1$, J_1, ·citizen of Japan·, f^0, Ø, b^0, zero-short$_S$, trunc$_S$(*anése*, 1), zero-short$_S$, id, $^\frown$, J_1-to-J, derogatory)
= ⟨*jáp*$_1$, J, ·speaker-disliked citizen of Japan·⟩,

for J_1 = {Unsp$_{Case}$, Unsp$_{Num}$, Unsp$_{Def}$}, J = {Unsp$_{Case}$-St, Sg$_N$-St, Unsp$_{Def}$-St}, and trunc$_S$(*anése*, 1) (*jàpanése*$_1$) = *jáp*$_1$ (read: 'the truncation in S of *anése* in member 1 of *jàpanése*$_1$ is *jáp*$_1$'). The meaning of *jáp*$_1$ is derogatory, while the meaning of *jàpanése*$_1$ is not. Being derogatory should be included in the intension of a 'deictic' concept (for such concepts, see Richter 1988: Ch. 4, Lieb 1993b: Sec. 5.6) obtained through use of a function derogatory (der) such that, for any argument ⟨b_1, b_2⟩ of the function, der(b_1, b_2) = speaker-disliked(b_1) (this would have to be spelled out).

Truncation functions in S are form-change functions, not shortening functions: in one or several members w of a sequence f, they remove one or several parts; they do not completely remove any member from the sequence, as a non-zero shortening function would. Since several members may be affected simultaneously by a truncation function, such functions allow us to construe *acronym building in* S as a special case of stem-form clipping in S.—*Syntactic* cutting and clipping can be treated in a largely analogous way.

4.5 b. Back formation in S

In a classical case like German ⟨*nót*$_1$ *land*$_2$, {Unsp$_{Pers}$-St, …, Pres-St}, ·perform an emergency landing·⟩, from ⟨*nót*$_1$ *land*$_2$ *ung*$_3$, {Unsp$_{Case}$-St, Sg$_N$-St, Unsp$_{Def}$-St, Fem-St}, ·emergency landing·⟩, we may now assume stem-form conversion in S, with a basic shortening β_{11} that eliminates ⟨3, *ung*⟩ from the sequence f_1 = *nót*$_1$ *land*$_2$ *ung*$_3$. A more detailed discussion of back formation is not possible here but it appears feasible to construe all types of backformation processes in S as subcases of stem-form conversion in S, or, possibly, word-form conversion in S. The meaning change in back-formation use creates the typical effect of 'going back to something that was presupposed'.

4.5 c. Strict conversion in S

Differentiating between subcases of stem-form or word-form conversion in any S is both possible and necessary. In particular, stem-form conversion in S that requires β_{11} to be zero-shortening, β_{12} to be identity, and δ or ε to involve a part-of-speech change—the classical case of conversion—could be singled out as *strict conversion in* S. Narrowing conversion functions down to strict conversion turns out to be problematic by blurring the overall picture of word-formation processes, a point to be taken up below, in Sec. 5.4.

Applying strict conversion in S implies a change of lexical meaning. This need not be true of stem-form cutting in S, as exemplified by (i), nor need it be true of stem-form clipping in S. How do we separate such cases from inflection?

4.5 d. The problem of meaning constancy

It is a defining property of inflection processes that they leave lexical meanings unaffected, which normally does not hold of word-formation processes in S. Could we not construe all word-formation processes in S in a way that makes a change of lexical meaning obligatory, neatly separating them from inflection processes? Unfortunately, identity of basic meaning and result meaning is not easily excluded, if excluded at all, for word-formation processes such as clipping and cutting in S. How, then, are word formation processes in S to be distinguished from inflection processes if constancy of lexical meaning is allowed to occur also with word-formation processes?

Let us assume that stem-form and word-form inflection in S are functions analogous to word-formation processes in S. Then the following three-part condition may be formulated as a necessary condition for *inflection processes* φ in S.

Let ⟨f, J, b⟩ and ⟨f_1, J_1, b_1⟩ be the result triple and the basic triple, respectively, of an application of φ. Then:

(iii) $b = b_1$; and $J \neq J_1$; and J and J_1 both are sets of morphological categories of S, or both are sets of syntactic categories of S.

The partial condition of $J \neq J_1$ requires *some* grammatical effect of inflection; thus, the relationship between, say, German *mannes*$_1$ and *manns*$_1$, both of them Gen Sg_N, is not reconstructed through inflection. The third part of (iii) requires that morphological inflection in S must have a morphological basis in S, and syntactic inflection in S a syntactic basis (apparent problem cases such as English *forget-me-nots* or *lookers-on* can be analysed in agreement with this requirement).

Now consider a *word-formation process* φ in S. Again, let ⟨f, J, b⟩ and ⟨f_1, J_1, b_1⟩ be the result triple and the basic triple, respectively, of an application of φ. We then assume that:

(iv) $b \neq b_1$; or $b = b_1$ and J and J_1 are non-inflective in S,

where 'non-inflective' is understood as follows:

(v) J and J_1 are *non-inflective* in S iff (α), (β) or (γ):
α. $J = J_1$;
β. J but not J_1 is a set of morphological categories of S;
γ. J but not J_1 is a set of syntactic categories of S.

Because of (iii) to (v), it now follows that no inflection process in S is a word formation process in S, and conversely, despite the fact that $b = b_1$ is not excluded for the applications of word-formation processes in S: there may be constancy of lexical meaning in an application if the category sets J and J_1 are non-inflective.

In particular, there may be constancy of lexical meaning in cases of stem-form or word-form *conversion in* S. For instance, it is true of the *mòbile*$_1$ *phóne*$_2$-example in (i) of Sec. 4.5 a that there is no meaning change. However, J is a set of morphological categories and J_1 a set of syntactic categories, contrary to (v.β): J and J_1 are non-inflective. Therefore, (i) as an instance of stem-form conversion in S is not a case of inflection, by (iii), despite the lack of a meaning change. Moreover, conversion functions in S may also apply now in borrowing: both (β) and (γ) in (v) are trivially satisfied if J_1 is not a set of categories of S but of some other system S_1.

We may also have constancy of lexical meaning where it has been suggested for *compounding functions* in S, such as ‘compounds with synonymous constituents’ in Chinese (Scalise & Bisetto 2009: 52): $b = b_1$, but also $J = J_1$, which excludes such cases as instances of inflection, by (iii).

I do not know of any clear cases where $b = b_1$ must be assumed with respect to *derivation processes* in some S. If there are such cases, J and J_1 must be non-inflective by (v).

It was announced in Sec. 1.2 b that the two major subcases, the morphological and the syntactic, of the three basic word-formation processes in S were to be characterized before the basic processes themselves. We now consider the defining properties of the three basic processes.

5 The basic processes in S: defining properties

5.1 Role of the added triple. The process cline

We are characterizing a conception by which both the basic word-formation processes in any idiolect system S and their two major subcases are identified with functions φ of a single formal type. If the two major subcases of a basic process are independently given, obtaining the basic process itself is trivial: we simply apply set theoretic union to the subcases. This was the approach I followed in earlier versions of the Process Model. If, however, the basic processes are taken to be primary, which leads to a simpler theory, there must be a defining property for each basic process in S that is shared by its two major subcases and also serves to distinguish this process from the other two basic processes. It was difficulties with finding such properties that originally made me start from the morphological and syntactic subcases as independently given.

In principle, a defining property may be obtained by referring to any component or components of the arguments or values of a basic process φ. It eventually turned out that defining properties suitable for the three basic word-formation processes should be based on properties of the added triples (*not* the basic triples) in the arguments of a given process; apparently, this is the only way. Roughly, the *compounding process* in S is characterized by having as added triples fully specified (morphological or syntactic) units of some S_1 (not necessarily S) whose meaning components are non-empty (are concepts different from the empty concept b^0); the *derivation process* in S is characterized by having as added triples fully specified units of some S_1 whose meaning component is the empty concept; and the *conversion process* in S is characterized by having as added triples the triples where *all* components are empty. We thus arrive at a three-part *process cline* of subcases of the compounding / derivation / conversion process in S that is based on the decreasing 'content' of the added triples. (See also Sec. 6.4 b, below.)

It may be argued that concepts of iconicity and prototypicality as applied in Natural Morphology (for instance, in Manova 2011) also underlie the process cline: the added triples associated with the compounding process in S, it may be claimed, are both most 'iconic' in relation to a general form/meaning relationship and 'prototypical' by not involving empty entities; in this respect the triples precede the added triples that are associated with the derivation process in S, which in turn precede the conversion triples. But this is only an analogy: in establishing the process cline no use is made either of a concept of sign or of Prototype Theory (as formally reconstructed already in Lieb 1980b).

Assuming added triples in this way is not without its problems. The requirements for the added triples will be discussed separately for the three processes.

5.2 The compounding process in S: the requirement of a non-empty added concept

5.2 a. Introduction

The defining property for the compounding process in any S requires the added concept b_2 (but not the basic meaning b_1) to be non-empty. This requirement must be compatible with the following condition: core instances of lexical-word compounding in a traditional sense must continue to be instances of lexical-word compounding, and more marginal instances must be covered either through the compounding process in S or through one of the other two basic processes. The first part of this condition is obviously satisfied, but there are at least three marginal cases traditionally claimed to be compounding where the compounding process in S might not meet the requirement of a non-empty added concept b_2; in each case, the problem arises with the process of stem-form or morphological compounding in S.

5.2 b. Problems raised by so-called empty morphs

This is the notorious 'cranberry' case, with an 'empty morph' *cran*.

Suppose that we wish to assume compounding in S here (derivation may be excluded).We must then have a fully specified morphological unit ⟨*crán*$_1$ *berry*$_2$, {Sg_N-St, …}, ·cranberry shrub·⟩ (certain botanical subtleties are disregarded in assuming ·cranberry shrub·; the concept ·berry of the cranberry shrub· is not chosen because it obviously presupposes the concept ·cranberry shrub·). This fully specified morphological unit must now be obtained as the result of applying stem-form compounding in S to a basic triple with *bérry*$_1$ and an added triple ⟨*crán*$_1$, J_2, b_2⟩. A reasonable choice (possibly, the only one) for the added concept b_2 *in this context* appears to be the empty concept b^0 (some non-empty concepts that may indeed be associated with *crán*$_1$ are irrelevant here). It is unclear what the category set J_2 should be, other than being a set of stem-form categories; we leave J_2 undetermined.

While empty-meaning triples are admitted as *basic* triples when the compounding process in S is applied, we must exclude them as *added* triples. At least the following three solutions may be proposed for the problem presented by the 'cranberry' case:

(i) a non-empty concept is constructed for *crán*$_1$, to be used as the added concept b_2, and the compounding process in S is applied; or
(ii) *crán*$_1$ *berry*$_2$—possibly, *cránberry*$_1$—is a stem form that does not result from any word-formation process; or
(iii) *cránberry*$_1$ is a stem form that results from the stem-form *conversion* process (the derivation process has been excluded) applied to a triple with *bérry*$_1$, a solution rendered possible by our generalization of traditional notions of conversion.

I do not know of any tenable proposal for a non-empty meaning of *crán*$_1$ in the present context.[41] The failure to come up with a non-empty meaning here is not an accident: intuitively, it is simply the change of form—the very occurrence of the morph-like *cran* as a 'form', independently of any 'meaning' it may or may not have—that creates a name for the cranberry shrub directly from the *bérry*-word.

The second solution—there simply is no word-formation here—does not account for the fact that the *berry*-part of *cránberry* can be interpreted.

I therefore adopt the third solution.[42] The stem-form conversion process in S is applied to ⟨*bérry*$_1$, {Sg_N-St,…}, ·berry·⟩ as follows: a form-change function '*crán*-extension'

[41] This includes the proposal made in Nolda (2012b: Sec. 4.3.3) to use in such cases the concept of being a 'basic' entity x identical to itself: assuming that the compounding process is applied, any semantic contribution made by *crán*$_1$ should *restrict* the contribution made by *bérry*$_1$, which excludes admitting all 'basic' entities x; in fact, Nolda subsequently outlines an alternative by which the *cránberry*-word is no compound at all but is a 'pseudo-compound'; for such a category, I see no need.

[42] This solution is not available to Nolda in his (2012b) model.

simultaneously changes the sound-sequence, syllable sequence, and intonation structure of *bérry*$_1$ to obtain *cránberry*$_1$, the category-change function is identity, and a meaning-change function is applied to ⟨·berry·, b^0⟩ to yield ·cranberry shrub· (b^0 is from the 'empty' added triple here that is used in applying the conversion process). As a result, we have the fully specified stem-form ⟨*cránberry*$_1$, {Sg$_N$-St, …}, ·cranberry shrub·⟩.

This analysis is in agreement with the way a fully specified stem-form like ⟨*bláck*$_1$ *berry*$_2$, {Sg$_N$-St, …}, ·blackberry bramble·⟩ is obtained: we first apply stem-form compounding to obtain ⟨*bláck*$_1$ *berry*$_2$, {Sg$_N$-St, …}, ·black berry·⟩ (where ·black berry· is the concept of being a berry that is black), with segmentally unreduced *berry*, and go on to apply stem-form conversion—it is just the first step that is missing in the *cránberry*$_1$-example.

In summary, the 'cranberry' case is compatible with the empty-concept requirement that is imposed on the added triple when the compounding process in S is applied; the case can be treated not as an instance of using the compounding process but as covered by the conversion process in S; this is possible because on our conception, both segmental and suprasegmental form change is compatible with applying the conversion process in idiolect systems S.

5.2 c. Problems raised by so-called neoclassical compounds

Lexical words like ⟨*ànthropólogy*^{P}, ·study of mankind·⟩ or ⟨*ànthropocéntric*^{P}, ·regarding man as central fact of universe·⟩ must be due to word-formation processes, more specifically, stem-form processes. Among the vexed problems presented by such words, only one question need be considered here: do we have to assume that these processes are, or may be, subcases of stem-form compounding requiring added triples with an *empty* lexical meaning?

On one traditional analysis of so-called neoclassical compounds, we would assume stem-form *derivation* with a 'combining form' *ánthropo*$_1$ and two suffix forms *lógy*$_1$ and *céntric*$_1$ (with change of stress due to form-change functions), so there would be no compounding at all. Still, a derivational analysis may not be adequate, or not adequate in all cases.

Suppose, then, that we are dealing with *stem-form compounding* in S in the *first example*. As an added triple, we choose:

(i) ⟨*ánthropo*$_1$, {Neut$_N$-St}, ·concerning mankind·⟩,

and as a basic triple:

(ii) ⟨*lógy*$_1$, {Sg$_N$-St, …}, ·study·⟩,

both associated with stem lexemes (adjectival for (i), substantival for (ii)) that are 'trapped', i.e., are not stem lexemes of lexical words. We now apply stem-form compounding in S (with accent-change functions) to obtain:

(iii) ⟨*ànthropó*$_1$ *logy*$_2$, {Sg$_N$-St. …}, ·study whose subject matter is mankind·⟩

The *second example* may be analysed as an instance of stem-form derivation in S applied after stem-form compounding in S. We first apply stem-form compounding in S, choosing a *céntre*$_1$-triple as the basic triple and (i) as the added triple, to obtain ⟨*ànthropo*$_1$

céntre$_2$, {Sg$_N$-St, …}, ·man as central fact of universe·⟩; we then take this as the basic triple in applying stem-form derivation in S, with an *ic*$_1$-triple as the added triple, obtaining:

(iv) ⟨*ànthropo*$_1$ *céntr*$_2$ *ic*$_3$, {Neut$_N$-St}, ·regarding man as central fact of universe·⟩

As an alternative, we may consider stem-form compounding in S with an adjectival *céntr*$_1$ *ic*$_2$-triple as the basic triple, itself obtained by stem-form derivation in S.

In summary, either there is no subcase of the compounding process in such cases, or there is stem-form compounding with a non-empty added concept, as required.[43]

5.2 d. Problems raised by syntactic units with an empty lexical meaning: German *nìcht*$_1$

Certain particles (lexical words) should be assumed with the empty concept as their lexical meaning, their semantic effects being non-lexical. This is true, in particular, of the sentence negation and of qualifying particles. For example, in English there is *not*^{W} = ⟨*nòt*^{P}, b^0⟩, with an improper paradigm whose forms are *nòt*$_1$ and (arguably) *n't*$_1$; the forms are *syntactic* units whose lexical meaning is the empty concept b^0.[44] Empty lexical meanings are also assumed for the forms of qualifying particles, such as *ònly*$_1$ used in sentences like: 'Only a boy can have done this.'.

In German, but arguably not in English, sentence-negation forms and forms of some qualifying particles may appear in stem-form compounding as additions: as the first components of added triples. For instance, the sentence negation in Standard German idiolect systems, ⟨*nìcht*^{W} = *nìcht*^{P}, b^0⟩, has the single form *nìcht*$_1$. The form may be combined with substantival and adjectival stem forms by *stem-form* or *morphological compounding* in S, replacing secondary by primary word stress, as in: *nícht*$_1$ *haus*$_2$, a stem form with the meaning ·non-house·, or *nícht*$_1$ *klug*$_2$, meaning ·non-clever·. So do we have to admit fully specified syntactic units with an empty lexical meaning as added triples in stem-form compounding? Contrary to appearances this is not the case, as I will now argue, using the *nícht*$_1$ *haus*$_2$-example.

We start from the fully specified syntactic unit ⟨*nìcht*$_1$, {Pf}, b^0⟩, with an empty lexical meaning. To this we apply stem-form *conversion* (not compounding) in S, with a change of secondary word stress to primary stress, obtaining the fully specified morpho-

[43] Beyond this result, the problems raised by neoclassical compounds are here left undiscussed. My position towards 'combining form' as a *separate morphological category* in addition to Affix form and Stem form is, however, negative, as it is towards the analogous category of *Konfix* or *Konfixform*, much discussed in recent German linguistics but studied exhaustively by Eins (2008) with largely negative results. In contradistinction to Eisenberg (2012: Sec 6.4), who continues to defend and use *Konfix* as a separate morphological category, I fail to see any cogent reason for such a category. In some cases of 'confixes' or 'combining forms', we may be dealing with affixes or their forms; in other cases, with 'trapped' stem lexemes or their forms: only artificially are these lexemes excluded from, and opposed to, the class of trapped (or 'bound') stem lexemes in general, as required by Eisenberg's position.

[44] Apparently, the intonational prominence that *not*$_1$ may have in a sentence (in non-citational use) may be either asemantic—that is, have no effect on sentence meaning—or contrastive, i.e., go with a contrastive sentence meaning; non-contrastive sentence stress on *not*$_1$ used as a sentence negation appears to be excluded. This state of affairs is expressed by assuming inherent *secondary word stress* for the form: *nòt*$_1$. Similarly, for the sentence negation in German, and for forms of qualifying particles in English and German; cf. Lieb (1999).

logical unit (stem-form) ⟨*nícht*$_1$, {P-St}, ·nicht·⟩, where P-St is the set of Particle-Stem forms (of the given idiolect system), and ·nicht· is the (non-empty) two-place concept whose intension consists of the relation between any x and set y such that x is *not* an element of y. For the transition from b^0 to this concept, a semantic function of '*nìcht*$_1$-filling' is used. (A similar conversion step may be considered in English for obtaining a triple with *nón*$_1$ from a triple with *nòt*$_1$, allowing for a segmental in addition to the suprasegmental change to go with the change of meaning.)

Only then do we apply stem-form *compounding*: using ⟨*haús*$_1$, {Sg$_N$-St, …}, ·house$_1$·⟩ as the basic triple and ⟨*nícht*$_1$, {P-St}, ·nicht·⟩ as the added triple, we obtain the result triple ⟨*nícht*$_1$ *haus*$_2$, {Sg$_N$-St, …}, ·non-house·⟩, where the intension of ·non-house· consists of the property of being an x such that ⟨x, ^{e1}·house·⟩ ∈ ^{e2}·nicht·; as part of using the compounding process, a meaning-change function of 'semantic qualifying' is applied to ⟨·house·, ·nicht⟩. (Assuming the triple with non-empty ·nicht· is not *ad hoc*; the triple also serves as a *basic* triple in stem-form derivation in S: in the derivation of the Verb-Stem form *ver*$_1$ *nícht*$_2$, meaning ·destroy·, i.e., 'make into nothing'; or the Adjective-Stem form *nícht*$_1$ *ig*$_2$, meaning ·invalid·.)

For the first step—use of stem-form conversion in S—we must again allow a 'trapped' stem lexeme: ⟨*nícht*^{LP}, ·nicht·⟩, whose improper paradigm is {⟨*nícht*$_1$, {P-St}⟩}. It is the non-empty concept ·nicht·, not the empty concept b^0, that is used as the added concept in the second step, application of stem-form compounding in S.

5.2 e. Generalizations

The *nícht*$_1$-example generalizes in German to all cases of 'stem-form negation': we are dealing with *stem-form compounding* in S; the *added* triple is a fully specified stem-form with the *non-empty* lexical meaning ·nicht·, and is itself the result of *stem-form conversion* in S where the *basic* triple is a fully specified word-form with the *empty* lexical meaning b^0. Such a two-step procedure, use of the compounding process after the conversion process, can also be shown to be adequate in German for stem-form compounding when qualifying particles rather than negation particles are involved, leading to stem forms such as *núr*$_1$ *haus*$_2$, meaning ·house and house only·, or *aúch*$_1$ *haus*$_2$, meaning ·house but not only house·.

True, alternative analyses of these examples, including the *nícht*$_1$-example, may be suggested depending on one's framework; what had to be shown here is simply this: the examples do not necessarily vitiate the general requirement of a non-empty added concept for the compounding process in S.

There is only one additional problem case I am aware of, stem-form compounding that involves quoting, as in (one type): 'the F-word'. Such compounding, which cannot be discussed here, also turns out to be compatible with the requirement of a non-empty lexical meaning for added triples.

5.3 The derivation process in S: the requirement of an empty added concept

5.3 a. Stem-form or morphological derivation in S

On our conception, stem-form derivation in S is to correspond to traditional derivation 'in the narrow sense', where use of affixes (including infixes) is required; this excludes, in particular, stem-form conversion in S but does not exclude 'internal' form change or 'modification' of the basic unit when the change is due to affixation; simple 'modification', which may or may not be allowed as sufficient for derivation in the narrow sense, is excluded on our conception. Stem-form or morphological derivation in S may then be construed as derivation by means of added triples $\langle f_2, J_2, b_2\rangle$ that are fully specified affix forms. Since the empty concept b^0 is assumed to be the lexical meaning of any affix form, the added concept b_2 in stem-form or morphological derivation in S is the empty concept b^0.

5.3 b. Word-form or syntactic derivation in S

The situation is more complex when we turn to syntactic derivation in S. What is required here, is exemplified by the *lòck*$_1$ *úp*$_2$-example in Sec. 4.4: (i) an *added* triple that is a fully specified syntactic unit with an empty lexical meaning; (ii) a *basic* triple that is a fully specified syntactic unit; and (iii) *non-identity* of the basic meaning and the result meaning. Condition (iii) is essential. For suppose that (i) and (ii) are satisfied, the basic category set J_1 and the result category set J are different (both are syntactic), and requirement (iii) is *not* met, i.e., the basic meaning and the result meaning are identical (the situation arising in English from the use of auxiliary verbs): we are then confronted not with a *word-formation* process in S but with a process of *syntactic inflection* in S (see Sec. 4.5 d, above).

The added triples in *syntactic derivation* in English systems S are fully specified particle forms with an empty lexical meaning; it is an empirical question which particle forms qualify. Whenever a fully specified particle form with a *non*-empty lexical meaning is used as an added triple in a syntactic word-formation process, we are dealing with *syntactic compounding* in S, as exemplified by the *lòck*$_1$ *oút*$_2$-example in Sec. 3.5.

It is not immediately clear whether the restriction to particle forms may be generalized. In particular, Drude (2010: Sec. 8) argues for *derivational verbs* in Awetí (a Tupí language), which would be the source of added triples on our approach. But the verbs in question have a *non*-empty lexical meaning. We are therefore confronted with the following alternative: either we are not dealing with word-formation here, or we have to assume *syntactic compounding* in S. And indeed, the reasons given by Drude *against* syntactic compounding are unconvincing: they do not exclude determinative syntactic compounding in S for the relevant two-verb constructions. It has to be admitted, though, that the entire area of 'serial verb constructions' as discussed in Aikhenvald & Dixon (2006), or 'multi-verb constructions' in Aikhenvald & Muysken (2010), remains to be analysed from a Process-Model point of view to determine the status of the constructions with respect to word formation (non-word vs. compounding vs. derivation vs. conversion) and inflection.

In *stem-form* derivation in S, the changes that the basic unit f_1 may undergo must all be due to use of f_2 in the added triple. This also holds of *word-form* derivation in S: f_2, whatever its word-form status, is analogous to an affix form in stem-form derivation.

5.4 The conversion process in S: going beyond strict conversion

5.4 a. Strict conversion in S: the requirement of an 'empty' added triple

As stated above (Sec. 4.5 c), the process involved in conversion in a traditional, strict sense can be construed as *strict conversion in* S: a function φ such that in each argument of φ, the basic triple $\langle f_1, J_1, b_1\rangle$ is a fully specified unit of S; the added triple $\langle f_2, J_2, b_2\rangle$ is the 'empty' triple $\langle f^0, \varnothing, b^0\rangle$, where f^0 is the empty sequence (= the empty set); the shortening, form-change, arrangement, and category-change functions are such that: if $\langle f, J, b\rangle$

is the result triple for the argument, then f is identical to the basic unit f_1, and a part-of-speech change is implied by the transition from J_1 to J or by the transition from b_1 to b (b is different from b_1). This characterization holds of both strict *morphological* or *stem-form* conversion in S (*wálk*$_1$ as a Noun-Stem form, from *wálk*$_1$ as a Verb-Stem form), and strict *syntactic* or *word-form* conversion in S (Lat. *bonum*$_1$ as a Substantive form meaning ·advantage·, from *bonum*$_1$ as an Adjective form meaning ·favourable·). Since there must be a change of lexical meaning, strict conversion in S is not an inflection function.

In summary, the requirement of an empty added triple is compatible with conversion in a traditional, strict sense. This remains true when traditional restrictions on conversion are lifted, as some of them frequently are. The following condition must always be satisfied, though, in order to keep the conversion process in S apart from inflection processes (see Sec. 4.5 d, above):

(i) The result meaning b is different from the basic meaning b_1, or the two are identical and the result categorization J and the basic category set J_1 are non-inflective in S.

5.4 b. Lifting the restrictions (1): the construction mode

When the compounding process in S or the derivation process in S are applied, both non-zero shortening and true form change should be allowed. It would be an artificial restriction, unmotivated in view of the conversion process in S as a basic process, to exclude change functions of the two types from the conversion process in S, i.e. proper shortening and form change—both segmental and suprasegmental—should be permitted for the basic unit f_1 (the addition f_2 is empty, hence, remains unaffected).

Furthermore, requiring a part-of-speech change through the category-change function, the meaning-change function, or both, is unduly restrictive (Sec. 4.4 b, above): identity of the result meaning and the basic meaning is permitted by (i), and is even compatible with the basic category set being identical to the result categorization; in such cases, a part-of-speech change is excluded. In summary, not only may f_1 change but the requirement of a part-of-speech is lifted, too:

(ii) The conversion process in S allows for shortening and form change of the basic unit; the requirement of a part-of-speech change—implying a meaning change—is replaced by condition (i).

Stem-form *derivation* in S as previously conceived excludes any changes of the basic unit that are not due to affixation. Such changes are now accounted for by stem-fom *conversion* in S. Moreover, cutting, clipping and back-formation processes in S can now be construed as subcases of the conversion process in S (Sec. 4.5, above).

5.4 c. Lifting the restrictions (2): basic triples

Strict conversion in S is to apply to basic triples that are fully specified (morphological or syntactic) units of S. Two ways of liberalizing this requirement may be considered.

First, the basic triples may be allowed to be fully specified units not of S but of some other system S_1 that may well belong to another language. This results in a subcase of the conversion process in S, a subcase that may be taken to be the basis of lexical borrowing (Sec. 4.5 d).

Second, we may permit basic triples ⟨f_1, J_1, b_1⟩ where f_1 is a 'sound-imitation' in S, i.e. , is a sequence of phonological triples w of S that 'imitates sound-events', and where J_1 is the empty set and b_1 the empty concept. The corresponding subcase of the conversion process in S may serve as the word-formation process involved in onomatopoeia, arguably the only true 'word creation' *ex nihilo*.

Whether we do or do not subscribe to these two generalizations in addition to the previous ones, they are worth considering, which shows the power of the conception being proposed here for the conversion process.

5.4 d. Discussion

The following three topics are briefly discussed: empty entities, inactive functions, and the generalization of strict conversion.

Empty entities. Traditional objections to 'empty morphs', and empty linguistic entities in general, do not apply to the 'empty' added triple: its components are non-linguistic entities, two are purely formal, one is a specific concept; nowhere in the idiolect system are they separately represented. These entities serve to relate the conversion process in S to the derivation and compounding processes in S, and the derivation process to the compounding process.

Inactive functions. When the conversion process in S is applied, some functions in the construction mode remain 'inactive', i.e., leave the basic triple and the added triple unchanged: zero-shortening is the shortening function and identity the form-change function for the empty sequence f_2 in the added triple; and if the arrangement function is simply concatenation, it is 'inactive' because concatenating the basic unit f_1 with the empty sequence f_2 leaves f_1 unchanged. While these functions may simply be disregarded in practical work, they are essential to the theory, also marking the conversion process in S as being at the lower end of the process cline.

Going beyond strict conversion. There is a twofold motivation for lifting the restrictions on the *construction mode*. *First*, there are no such restrictions for the compounding and the derivation processes in S, and there is no cogent reason for treating the conversion process in S differently. In addition, we have to cover changes—outside compounding—of the basic unit that are not due to the use of an affix form or the use of an analogous syntactic unit, and are therefore excluded from the derivation process in S. *Second*, lifting these restrictions allows us to treat processes that would otherwise remain isolated, like cutting or clipping in S, as subcases of the conversion process in S.

A truly comprehensive theory of word-formation processes appears to be within reach if in addition to the restrictions on the construction mode, the restrictions on the basic triple are also lifted: even the processes involved in lexical borrowing or in onomatopoeia may then be included in the conversion process in S.

5.4 e. Alternatives

There are at least two alternatives to the present conception of the conversion process in S.

Alternative 1 (following a traditional approach). The conversion process in S continues to be identified with strict conversion in S; processes like cutting—let alone processes of borrowing—are treated as unrelated to the conversion, derivation, and compounding processes in S; and the derivation process in S is modified by also allowing changes of the basic unit that are *not* due to affixation or use of affix-like word forms; strict conversion in S and the (modified) derivation process in S may or may not be treated as subcases of the derivation process in S in a broad sense.

The alternative misses the very possibility of a unified theory of word-formation processes and is therefore rejected.

Alternative 2 (essentially the approach of Nolda 2012b, allowing for differences in the conception of processes). The restrictions on shortening and form change (and, possibly, part-of-speech change) are lifted except for one requirement: there must not be any *segmental* change of the basic unit when the conversion process in S is applied; segmental changes are to be accounted for by that part of the *derivation process in* S *in a broad sense* which is different from the conversion process in S; the *derivation process in* S *in a narrow sense* is reconceived as that part.

From a theoretical point of view, allowing suprasegmental changes for the conversion process in S but excluding segmental ones is arbitrary; indeed, the reasons given in Nolda (2012b: Sec. 3.1.3) for allowing *suprasegmental* changes appear to be just as valid for segmental ones. Moreover, on Nolda's approach, special 'conversion forms' must be introduced into stem and word paradigms in order to guarantee exclusion of segmental changes: when such a change *might* occur in the transition from the basic unit to the result unit, a 'conversion form' that is segmentally identical to the result unit is used by Nolda as the basic unit. Unfortunately, this trivializes the requirement of not having a segmental change. Furthermore, segmental and suprasegmental changes may occur in bundles and should not be separated then. (See also Sec. 2.1 f, above.)

Forbidding segmental changes for the conversion process in S has the general consequence that processes like cutting in S or clipping in S can no longer be treated as subcases of the conversion process in S, but must be treated as subcases of the derivation process in S in a narrow sense. Already in the case of traditional 'conversion' and traditional 'derivation in a narrow sense', there are well-known problems with *justifying* the relevant word-formation processes as subcases of a single process of 'derivation in a broad sense', problems further aggravated by any attempt to treat processes of 'abbreviation', like cutting or clipping, as subcases of 'derivation in a narrow sense'. On the other hand, such processes are naturally treated as subcases of the *conversion* process in S: obtained in a natural way by systematically lifting strict-conversion requirements.

For these reasons, Alternative 2 is rejected in favour of the present conception of the conversion process in S.

Sections 3 to 5 exemplify, motivate, and justify the process aspect of the Process Model of Word Formation. A major part of the model will now be characterized in a more explicit way (Sec. 6). It is then shown, briefly and incompletely, how the Process Model is used in actual language description (Sec. 7). Finally, the Process Model will be confronted with the theory of word formation proposed in Nolda (2012b) (Sec. 8).

C. The Process Model of Word Formation

6 An outline of the Model

6.1 General aspects

6.1 a. Introduction

I am envisaging a theory of language in axiomatic form that has, among others, the following, closely linked parts: the WLP model as characterized in Sec. 2; the Process Model of Word Formation (PMF), dealing with word formation; and the Process Model of Inflection (PMI), dealing with inflection; the last two jointly constitute the Process Model of Word Formation and Inflection (PMFI). Each one of the three theories may be referred to simply as 'the Process Model' (PM) if the context allows us to identify the theory that is meant.

It is only a part of the Process Model of Word Formation that is going to be characterized here: essentially the part dealing with word-formation processes. All of Secs 3 to 5 may be read as an exemplification of the theory sketch that is to follow.

From the very beginning let me emphasize that the distinction between axioms, definitions, and theorems is a *formal* one: axioms and definitions are underived, theorems are derived; definitions are so-called nominal definitions, laying down how a term is to be understood. In particular, being an axiom does *not* mean being a sentence whose truth, or even whose meaning, is self-understood. *No* sentence of an empirical theory, including its definitions, is exempt from revision, depending on how the theory as a whole stands up to the requirements made by the domain it is meant to cover.

The theory sketch is semi-formal, and is presented in isolation from the theory of language of which the Process Model is to be a part. The numbering of axioms, theorems and definitions (and the status of the axioms as underived) is therefore preliminary. It is essentially a part of the model's *conceptual core* that is given here: the definitions for terms introducing basic concepts, plus axioms that are needed, mostly, for formally justifying the definitions, plus a few theorems (without proof). Quite a few empirical assumptions made in previous Sections on specific word-formation processes or on word-formation processes in general are not yet represented.

6.1 b. The format of word-formation processes

Word-formation processes α in an absolute sense are construed as functions whose arguments are arbitrary idiolect systems S and whose values are certain functions φ: $\alpha(S) = \alpha_S = \alpha$ in S is a function φ. The formal nature of the functions $\alpha_S = \varphi$ is as follows, for word-formation processes in S of any kind:

Format Formula

$\alpha_S(f_1, J_1, b_1, f_2, J_2, b_2, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \gamma, \delta, \varepsilon)$
$= \langle\gamma(\beta_{12}(\beta_{11}(f_1)), \beta_{22}(\beta_{21}(f_2))), \delta(J_1), \varepsilon(b_1, b_2)\rangle$
$= \langle f, J, b\rangle$

The *second line* in the formula indicates how the functions β_{11} to ε operate on f_1 to b_2, yielding the triple in the *third line* as the value that the function α_S assigns to the 13-tuple $\langle f_1, \ldots, \varepsilon\rangle$ in the *first line* (for δ operating on J_1 but not J_2, see Sec. 3.3 a, *Remark*). This should be made more precise.

We first interpret the variables in the formula, except for 'α':

First List of Variables

'S', 'S_1', etc.: for any idiolect system of any language
'f', 'f_1', etc.: for any (empty or non-empty) sequence, or part of a sequence, of triples w, triples with the ontological status of morphs and phonological words of idiolect systems S (see above, Sec. 2.2); *example*: syntactic units
'J', 'J_1', etc.: for any set of sets K, where K is a non-empty set of non-empty sequences f; *example*: sets J of syntactic categories
'b', 'b_1', etc.: for any property of perceptions or conceptions; *example*: concepts
'β_{nm}' (n and m: 1, 2): for any function whose arguments and values are (parts of) sequences f; *example*: shortening and form-change functions
'γ', 'γ_1', etc.: for any function whose arguments are pairs $\langle f_1, f_2\rangle$ of sequences and whose values are sequences f; *example*: arrangement functions
'δ', 'δ_1', etc.: for any function whose arguments and values are sets J; *example*: category functions
'ε', 'ε_1', etc.: for any function whose arguments are pairs $\langle b_1, b_2\rangle$ and whose values are properties b; *example*: meaning-change functions

Remarks on the First List of Variables. Variables may also be primed: f', etc.—It appears from the list, especially from the interpretation of 'f', 'f_1', etc., that the theory to be presented is still restricted to 'spoken language'; its extension to 'written language', which is obviously possible, or even to 'signed language' requires adaptations not to be discussed here.

Next, consider the 13-tuple in the first line of the Format Formula:
$\langle f_1, J_1, b_1, f_2, J_2, b_2, \beta_{11}, \beta_{12}, , \beta_{21}, \beta_{22}, \gamma, \delta, \varepsilon\rangle$

We must make sure that the seven functions in this tuple always apply correctly with respect to the first six components f_1 to b_2; otherwise, the second line of the Format Formula does not make sense. The domain of each function must be assumed accordingly; the 13-tuple is to be 'properly constructed':

Argument Construction

$\langle f_1, J_1, b_1, f_2, J_2, b_2, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \gamma, \delta, \varepsilon\rangle$ is *properly constructed* iff:

a. f_1 is in the domain of β_{11};
b. $\beta_{11}(f_1)$ is in the domain of β_{12};
c. f_2 is in the domain of β_{21};
d. $\beta_{21}(f_2)$ is in the domain of β_{22};

e. $\langle \beta_{12}(\beta_{11}(f_1)), \beta_{22}(\beta_{21}(f_2)) \rangle$ is in the domain of γ;
f. J_1 is in the domain of δ;
g. $\langle b_1, b_2 \rangle$ is in the domain of ε.

Properly constructed 13-tuples determine triples $\langle f, J, b \rangle$:

Triple Determination

Let $\langle f_1, \ldots, \varepsilon \rangle$ be properly constructed.
$\langle f_1, \ldots, \varepsilon \rangle$ *determines* $\langle f, J, b \rangle$ iff:
$\langle f, J, b \rangle = \langle \gamma(\beta_{12}(\beta_{11}(f_1)), \beta_{22}(\beta_{21}(f_2))), \delta(J_1), \varepsilon(b_1, b_2) \rangle$.

Given this definition, we introduce the variables that are still missing:

Second List of Variables

'φ', 'φ_1', etc.: for any function (possibly empty) whose arguments are properly constructed 13-tuples $\langle f_1, \ldots, \varepsilon \rangle$ such that the value assigned by the function to $\langle f_1, \ldots, \varepsilon \rangle$ is the triple $\langle f, J, b \rangle$ determined by $\langle f_1, \ldots, \varepsilon \rangle$; *example*: word-formation processes in idiolect systems S

'α', 'α_1', etc.: for any function whose arguments are idiolect systems S and whose values are functions φ (possibly empty); *example*: word-formation processes per se (comp, st-comp, etc.)

'χ', 'χ_1', etc.: for any properly constructed 13-tuple $\langle f_1, \ldots, \varepsilon \rangle$; *example*: arguments of word-formation processes in idiolect systems S

The variables used in the Format Formula are given an interpretation in the two lists of variables that minimizes linguistic content: as far as possible, only formal aspects are covered of the various word-formation processes that were discussed in Secs 3 to 5. In those Sections a number of terms were informally introduced for reference to word-formation processes in S; the choice of these terms was linguistically motivated: 'basic triple', 'form-change function', etc. The terminology will now be generalized to arbitrary 13-tuples χ.

6.1 c. Process terminology

Suppose, then, that $\chi = \langle f_1, \ldots, \varepsilon \rangle$. The sixtuple $\langle f_1, \ldots, b_2 \rangle$ is the *basis* of χ, with χ's *basic triple* $\langle f_1, J_1, b_1 \rangle$, *basic unit* f_1, *basic category set* or *basic categorization* J_1, and *basic concept* or *basic meaning* b_1; and $\langle f_2, J_2, b_2 \rangle$ is the *added triple* of χ, with the *addition* f_2, *added category set* J_2, and *added concept* b_2 of χ.[45]

$\langle \beta_{11}, \ldots, \varepsilon \rangle$ is the *construction mode* of χ. γ is the *arrangement function* of χ; β_{11} is its *basic shortening*, β_{12} its *basic form-change*; β_{21} its *added shortening*, and β_{22} its *added form-change*; δ is the *category change* of χ, and ε its *meaning change*.—$\langle f, J, b \rangle$ is the *result for* or *of* χ.

Having this formal framework at our disposal, we are in a position to render explicit the notion of word-formation process in S.

[45] If χ is an argument of a word-formation process in S, the basic triple must be a fully specified unit, allowing for basic triples in quoting (not treated in this essay) and onomatopoeia as possible exceptions; the added triple need not be a fully specified unit, hence, the more neutral terminology for the added triple. For either triple, some S_1 other than S must be permitted to account for borrowing (see Sec. 5.4 c).

6.2 Word-formation processes in S

6.2 a. The word-formation base (Definitions 1 and 2, Axiom 1)

We first use the general explanations in Sec. 6.1 for introducing a concept of *potential* word-formation process in S (Definition 1). It is then assumed (Axiom 1) that any idiolect system S has exactly one component, the *word-formation base of* S (Definition 2), that is a set of potential word-formation processes in S.

Definition 1

φ is a ***potential word-formation process in*** S iff φ is non-empty and for every χ, f, J, b, f_1, J_1, and b_1, if χ is an argument of φ and φ(χ) = ⟨f, J, b⟩ and ⟨f_1, J_1, b_1⟩ is the basic triple of χ, then (a) and (b):

a. either (i) or (ii):
   (i) ⟨f, J, b⟩ is a fully specified stem form of S,
   (ii) ⟨f, J, b⟩ is a fully specified word form of S;

b. $b \neq b_1$, or $b = b_1$ and J and J_1 are non-inflective in S.

Informally, the potential processes are the functions that have a format as characterized in Sec. 6.1, exclude inflection (see Sec. 4.5 d), and have fully specified stem forms or fully specified word forms as their values.—The following axiom serves to assign word-formation processes in S a place in the idiolect system, it 'anchors' them:

Axiom 1 (Anchor Axiom)

Any idiolect system S has exactly one component that is a set M of potential word-formation processes in S.

Remark on Axiom 1. As an alternative, a single potential word-formation process in S instead of a set M of such processes might have been proposed, eventually making the processes of compounding, derivation and conversion in S specific subcases of this process. I am not following such a route because it overemphasizes the shared features of the three word-formation processes, blurring their relationship. There is no reason for uniting the three processes into a single one over and beyond the fact that they are functions of a single formal type; and the fact that the three processes can be arranged in a 'process cline' should be treated as a basic fact about idiolect systems S, rather than something that happens to be true of certain subcases of a more general potential word-formation process, a process for whose status as a component of S there is no *linguistic* motivation.

Definition 2

The word-formation base of S ***(wfb***$_S$***)*** = the M that satisfies Axiom 1 for S.

Axiom 1 allows us to speak of the word-formation base of any given idiolect system S (which may be empty—an unlikely case). This is a morphosyntactic component of S.

Remark on D1, Axiom 1, and D2. The following modifications may be considered; they would allow us to anchor inflection processes and word-formation processes in a single component of the idiolect system.

(i) No reference is made in the definiens of D1 to f_1, J_1, and b_1, and condition (b) is dropped; in the definiendum, 'or inflection' is added after 'word-formation'.

(ii) In Axiom 1, 'or inflection' is added after 'word-formation'.

(iii) In D2, ‘and inflection’ is added after ‘word-formation’, and ‘wbf_S’ is replaced by ‘wib_S’.

(iv) In D3, below, ‘wfb_S’ is replaced by ‘wib_S’, and the definiens is supplemented by reintroducing condition (b) from D1 as follows: ‘and for every χ, f, J, b, f_1, J_1, and b_1, if χ is an argument of φ and $\varphi(\chi) = \langle f, J, b \rangle$ and $\langle f_1, J_1, b_1 \rangle$ is the basic triple of χ, then: $b \neq b_1$, or $b = b_1$ and J and J_1 are non-inflective in S’.

I am not yet certain, though, that inflection processes should indeed be anchored in this way; D1, Axiom 1, and D2 are therefore formulated as above, and D3 is given as below.

6.2 b. Word-formation processes in S: basic definitions (Definitions 3 to 7, Definition Schema 1)

The *basic* word-formation processes are simply the elements of the word-formation base (Definition 3). The *word-formation processes in* S (Definition 4) are the non-empty (proper or improper) subsets of the basic word-formation processes, and the processes that are not basic are *derived* (Definition 5). *Stem-form/word-form processes* are the ones that result in stem forms / in word forms (Definitions 6 and 7). For all these terms, the relativization to S can be lifted (Definition Schema 1).

Definition 3

φ is a ***basic word-formation process in*** S iff $\varphi \in wfb_S$.

Definition 4

φ is a ***word-formation process in*** S iff:

a. φ is non-empty;
b. there is a φ_1 such that:
   (i) φ_1 is a basic word-formation process in S,
   (ii) $\varphi \subseteq \varphi_1$.

Definition 5

φ is a ***derived word-formation process in*** S iff:

a. φ is a word-formation process in S;
b. φ is not a basic word-formation process in S.

Definition 6

φ is a ***stem-form process*** (a ***morphological word-formation process***) ***in*** S iff:

a. φ is a word-formation process in S;
b. all values of φ are fully specified stem forms of S.

Definition 7

φ is a ***word-form process*** (a ***syntactic word-formation process***) ***in*** S iff:

a. φ is a word-formation process in S;
b. all values of φ are fully specified word forms of S.

Remark on D6 and D7. It has been a typical feature of the Process Model from its beginnings that morphological and syntactic word-formation processes are distinguished in terms of process results not process bases (similarly, Nolda 2012b, also Manova 2011: Ch. 3): a process whose application results in a stem form (morphological process) may well start from a word form or phrase, fully specified, either as a basic triple or as an added triple (syntactic basis). Attempts are frequently made in the literature to distinguish 'morphological' from 'syntactic' word-formation processes in terms of the bases rather than the results; it may be doubted that such attempts have ever been entirely successful. Using the process results instead, we are then free to additionally distinguish subcases by the morphological or syntactic nature of the basis (below, Sec. 6.4 b).

We may wish to speak of word-formation processes per se, i.e. without referring to any S. This is made possible by the following definition schema:

Definition Schema 1

Let t be any of the italicized parts of the definienda in D3 to D7 minus '*in*'.
α is a ***t per se*** iff for some S, α_S is a t in S.

For example, a word-formation process per se is a word-formation process in *some* S.

6.2 c. Subcases and parts of word-formation processes in S (Definitions 8 to 11)

There are a number of additional terms that prove useful in speaking about word-formation processes in any S, specifically, terms by which we may refer to 'subcases' and 'parts' of such processes.

Definition 8

Let φ be a word-formation process in S.
φ_1 is a ***subcase of*** φ iff the domain of $\varphi_1 \subseteq$ the domain of φ.

Remarks on D8. Because of the interpretation of the variables in the Second List of Variables, it is sufficient in D8 to refer to the *domain* of a function (the set of its arguments). The empty subcase is allowed. A word-formation process φ in S is of course a subcase of itself.—Word-formation processes φ in S are functions whose arguments are 13-tuples χ and whose values are triples $\langle f, J, b\rangle$. Any linguistic aspect of linguistic units that is relevant to the way the units behave in the application of word-formation processes is directly represented either through a component of χ or through f, J, or b. This guarantees maximal flexibility in using the concept of subcase: proper subcases may result directly from imposing conditions on any individual component or components of χ, or on f, J, or b, in any combination. It is mainly for this effect on subcases that the notion of potential word-formation process was introduced (D1) without imposing any further set theoretic structure on the arguments of such processes, leaving them simply as thirteen-tuples. (The notions of derived word-formation process in D5 and of subcase in D8 are at least as powerful as the conception of subschema used by Booij in his Construction Grammar framework, as in 2010: Ch. 3.)

Given a word-formation process in S, we distinguish its *morphological part* from its *syntactic* part—the part that leads to stem forms from the one that leads to word forms:

Definitions 9 and 10

Let φ be a word-formation process in S.
The stem-form part / word-form part of φ ***in*** S = the φ_1 whose domain is the set of all $\langle f_1, \ldots, \varepsilon\rangle$ such that:

a. $\langle f_1, \ldots, \varepsilon\rangle \in$ the domain of φ;
b. $\varphi_1(f_1, \ldots, \varepsilon)$ is a fully specified stem form / a fully specified word form of S.

Either part, but not both, may be empty. Non-empty parts are obviously word-formation processes in S, and are stem-form / word-form processes in S. The parts are also subcases of φ, in the sense of D8. Indeed, they are the *major* subcases:

Definition 11

Let φ be a word-formation process in S.
φ_1 is a ***major subcase of*** φ iff (a) or (b):
a. φ_1 is the stem-form part of φ in S;
b. φ_1 is the word-form part of φ in S.

Having outlined the general framework, let us now consider the three word-formation processes in S that are meant to be basic: the compounding, derivation, and conversion processes in S. We are going to concentrate on the processes per se, certain functions α that apply to idiolect systems S.

6.3 The compounding process, derivation process, and conversion process

6.3 a. The notion of type-n word-formation process in S (Definition 12, Axiom 2)

Consider, once again, the conception characterized in Secs 3 to 5 for the compounding process in S, the derivation process in S, and the conversion process in S. The three processes are distinguished from one another by each having a characteristic property that is determined by a requirement on the added triples. There are also properties that are *shared* by the three processes; each is to be a basic word-formation process in S, and is to be a unique process of its kind.

We account for the relevant properties, shared and non-shared, by a notion of 'type-n word-formation process', for n = 1, 2, or 3, and an Axiom of Uniqueness ('ad(χ)' for 'the added triple of χ'):

Definition 12

φ is a ***type***-n ***word-formation process in*** S iff:
a. φ is a basic word-formation process in S;
b. or c. or d.:
b. (i) n = 1;
(ii) for any argument χ of φ,
α. ad(χ) is a fully specified unit of some S_1,
β. the third component of $ad(\chi) \neq b^0$;
c. (i) n = 2;
(ii) for any argument χ of φ,
α. ad(χ) is a fully specified unit of some S_1,
β. the third component of $ad(\chi) = b^0$;
d. (i) n = 3;
(ii) for any argument χ of φ, $ad(\chi) = \langle f^0, \varnothing, b^0\rangle$.

Informally, a *type*-n *word-formation process in* S is a basic process in S such that either (n = 1) the added triples are fully specified units of some S_1 with a *non-empty* lexical meaning, or (n = 2) the added triples are fully specified units of some S_1 with an *empty* lexical meaning, or (n = 3) the added triples have only empty components. It follows from the definition that there are no type-n word-formation processes in S for n > 3. For any S, at most one basic word-formation process of each kind is allowed:

Axiom 2 (Axiom of Uniqueness)

For any S and n = 1, 2, 3, there is at most one type-n word-formation process in S.

It is this axiom that serves as a basis for the definitions that are to follow, definitions for the compounding, derivation and conversion processes per se.

6.3 b. The compounding process, derivation process, and conversion process defined (Definitions 13 to 15, Theorem 1)

The arguments of the compounding process per se (comp) are to be arbitrary idiolect systems S, and its values functions φ. The function comp may apply vacuously to a given S, i.e., may assign the empty set to S. But suppose comp does not apply vacuously. Then $comp_S = \varphi$ must be a type-1 word-formation process in S. The functions der and conv are related in the same way to type-2 and type-3 word-formation processes in S:

Definition 13

The compounding process per se (comp) = the α such that for any S, (a) or (b):

a. there is a type-1 word-formation process φ in S, and $\alpha_S = \varphi$;
b. there is no type-1 word-formation process φ in S, and $\alpha_S = \emptyset$.

Definition 14

The derivation process per se (der) = the α such that for any S, (a) or (b):

a. there is a type-2 word-formation process φ in S, and $\alpha_S = \varphi$;
b. there is no type-2 word-formation process φ in S, and $\alpha_S = \emptyset$.

Definition 15

The conversion process per se (conv) = the α such that for any S, (a) or (b):

a. there is a type-3 word-formation process φ in S, and $\alpha_S = \varphi$;
b. there is no type-3 word-formation process φ in S, and $\alpha_S = \emptyset$.

Terminological remark on D13 to D15. Each one of the three terms may be used, in the sense defined, without its ‘per se’-part: ‘the compounding process’, ‘the derivation process’, and ‘the conversion process’.

There is an obvious consequence of D12 to D15:

Theorem 1

For any S and each $\alpha \in$ {comp, der, conv}, if α_S is non-empty, then α_S is a basic word-formation process in S.

The three definitions clarify the sense in which the three terms ‘compounding’, ‘derivation’, and ‘conversion’, construed as names of word-formation processes α in a general sense and abbreviated as ‘comp’, ‘der’, and ‘conv’, apply to individual idiolect systems, where $comp_S$ (= comp(S)), der_S, and $conv_S$ either are basic word-formation processes in S, each of a certain type, or else are identical to the empty set.

6.3 c. The axioms of existence and completeness (Axioms 3 and 4, Theorems 2 and 3)

So far it has not been guaranteed that the compounding process, the derivation process, and the conversion process, construed as functions α, each assign a basic word-formation process in S to at least one idiolect system, instead of assigning the empty set everywhere. A separate axiom is needed:

Axiom 3 (Axiom of Existence)

For each $\alpha \in$ {comp, der, conv}, there is an S such that $\alpha_S \neq \emptyset$.

It then follows (compare Theorem 1):

Theorem 2

For each $\alpha \in$ {comp, der, conv}, there is an S such that α_S is a basic word-formation process in S.

The axiom and the theorem represent the weakest existence claim that can be made in relation to comp, der, and conv: the three processes each appear in the word-formation component of *some* S—not necessarily of the same S. We might have required that there should be *some* S in whose word-formation component *all three* are represented; and indeed, this should be true, among others, of all English and German idiolect systems S, i.e., $comp_S$, der_S, and $conv_S$ are basic word-formation processes in S. A corresponding assumption for arbitrary idiolect systems would probably be false (see Zamponi 2009: 592-593, on languages without compound words or without derived words).

Axiom 3 and Theorem 2 do not yet exclude basic word-formation processes other than $comp_S$, der_S, or $conv_S$, in some S. I have argued in Secs 3 to 5 that such processes need not be admitted, but ultimately this remains an empirical question. The following axiom, which excludes the alternative processes, is therefore tentative; the processes involved in various ‘minor’ types of word formation remain to be reconstructed; blending and reduplication processes in S—as processes involved in word-formation not inflection—should turn out to be subcases of the compounding process in S, where even partial reduplication can be taken into account by means of form-change functions.[46]

Axiom 4 (Axiom of Completeness) (tentative)

For any S and φ, if φ is a basic word-formation process in S, then there is an $\alpha \in$ {comp, der, conv} such that $\varphi = \alpha_S$.

Using Axiom 4 and Theorem 1, we obtain:

Theorem 3

For any S and φ, φ is a basic word-formation process in S iff there is an $\alpha \in$ {comp, der, conv} such that $\varphi = \alpha_S$ and α_S is non-empty.

Informally, the basic word-formation processes in any S are exactly the non-empty φ assigned to S by one of the three functions comp, der, and conv. This theorem is tentative, due to the tentative nature of the Axiom of Completeness.

---

[46] Assuming identity of the basic and the added triple in reduplication, plus form-change or shortening, appears to bypass the problem of ‘base-dependence’ (see Haugen & Hicks Kennard 2011).

6.4 Interrelating word-formation processes

6.4 a. The notions of α-subcase, stem-form-α, and word-form-α (Definitions 16 to 18)

Using the Axiom of Existence, it follows that the compounding process, the derivation process, and the conversion process per se are indeed *word-formation processes per se* in the sense of Definition Schema 1 (Sec. 6.2 b). For subcases in S that are associated with a specific word-formation process per se (basic *or derived*), we introduce the following term:

Definition 16

Suppose that α is a word-formation process per se.
φ is an α-***subcase in*** S iff:
a. α_S is a word-formation process in S;
b. φ is a subcase of α_S.

We may now speak of the *comp-subcases in* S, or the *compounding-process subcases in* S (the *compounding subcases in* S, for short), etc., in the plural.

The *major subcases* (above, D11) of non-empty $comp_S$, der_S, and $conv_S$ are their stem-form and word-form parts (D9 and D10). In the informal part of this essay I have been using special terms for the major subcases, such as '***st-comp***$_S$', to be read as '***the stem-form compounding process in*** S', abbreviated as '***stem-form compounding in*** S'. The terminology used is now introduced in a general way:

Definitions 17 and 18

Suppose that α is a word-formation process per se, and α_S is a word-formation process in S.
(***st***-α)(S) / (***wf***-α)(S) = the stem-form part / the word-form part of α_S.

The relationship between compounding subcases in S, derivation subcases in S, and conversion subcases in S can now be characterized as a non-continuous, three-step cline.

6.4 b. The process cline (Theorem 4)

Consider word-formation processes φ in S that are, respectively, compounding, derivation, or conversion subcases in S. The relationship between the three types of processes is characterized by the following theorem, which follows from axioms and definitions in Secs 6.2 and 6.3 plus an assumption made in the WLP model on fully specified units (their first two components must be non-empty):

Theorem 4 (Process Cline Theorem)

For any S, non-empty φ, argument χ of φ, f_2, J_2 and b_2:
a. if φ is a comp-subcase in S and $ad(\chi) = \langle f_2, J_2, b_2\rangle$, then f_2 and J_2 are non-empty and $b_2 \neq b^0$;
b. if φ is a der-subcase in S and $ad(\chi) = \langle f_2, J_2, b_2\rangle$, then f_2 and J_2 are non-empty and $b_2 = b^0$;
c. if φ is a conv-subcase in S and $ad(\chi) = \langle f_2, J_2, b_2\rangle$, then f_2 and J_2 are empty and $b_2 = b^0$.

(ad(χ) = the added triple of χ.) It was a guiding principle used in developing the Process Model that such a theorem should be possible. The reason is this.

At a heuristic, pretheoretic level, a word-formation process in its most explicit form consists in adding a linguistic entity to another linguistic entity, both of them semantically specific, so as to obtain a third linguistic entity, also with a specific meaning: this is, informally, definitional for the *compounding process* and its subcases in a linguistic system. One step down from this level, it is linguistic entities with an empty lexical meaning that are added; in the area of stem forms this is typically achieved by use of affixes (which, it is generally agreed, have no lexical meaning, or whose lexical meaning is 'empty'). Taking the empty-meaning requirement as definitional, we arrive at the *derivation process* and its subcases. Still one step down, we remove all 'content' from the added linguistic entity, leaving nothing but an empty shell: this takes us to the *conversion process* and its subcases.

There is one complication, though: how to deal with 'internal changes', changes that do not consist in adding any linguistic material but consist in modifying, or subtracting from, what is given. Such changes may be concomitant with adding a linguistic entity and may then be treated as part of the process of addition. However, if only an empty shell is added—or nothing at all, depending on one's theoretical orientation—then there is a problem. This is solved by taking the following position: internal changes may occur whenever a word-formation process is applied; they *may* but *need not* be due to adding a linguistic entity.

There is strong empirical evidence from language development for the Process Cline Theorem. (i) Affix forms or particle forms used in the *derivation process* in S typically develop from stem forms or word forms added in the *compounding process* in some earlier system S_1. (ii) The *loss* of affix forms or particle forms used in the *derivation process* in some S_1 is a typical source of subcases of the *conversion process* in some later S: by giving rise to empty shells as added triples while retaining the 'internal change' of the basic unit after the form that caused it was lost, a conversion subcase comes into existence.

6.5 Completing the Process Model

6.5 a. Extending the theory of word-formation processes

So far we have been dealing only with that part of the Process Model of Word Formation which renders explicit the conception of word-formation processes. Even so, our presentation has been incomplete, for at least three reasons.

First, for each one of the three basic word-formation processes there are further requirements, *non-definitional*, that the process *must* meet in any S. The definitional requirements all concern added triples, but there are also non-definitional general conditions such as: the *arrangement function* must be a subfunction of concatenation when the conversion process in any S is applied; and apparently, only fully specified *syntactic* units are allowed as basic triples when the major *syntactic* subcase of one of the three basic word-formation processes is applied. General conditions of this type require additional axioms.

Second, there are further important general distinctions between word-formation processes. In particular, a distinction may be drawn, non-disjunctively, between ***word-formation processes in S with a syntactic basis*** (the basic triple, the added triple, or both, are fully specified syntactic units) and ***word-formation processes in S with a morphologi-***

cal basis (the basic triple, the added triple, or both, are fully specified morphological units).

Having a morphological basis is different from being a morphological word-formation process in S (a stem-form process in S); analogously, having a syntactic basis and being a syntactic word-formation process in S (a word-form process in S): stem-form and word-form processes in S are distinguished by their *results*, not by their basic or added triples; see *Remark on D6 and D7*, above. A morphological word-formation process may well have a syntactic basis: by the analyses in Nolda (2012b: Part C), 'nominalization of infinitives' in German systems S uses *stem-form* conversion in S with a *syntactic* basis: the basic triple is a fully specified *syntactic* unit (word form or phrase) of S, the result is a fully specified stem form, a *morphological* unit. This is opposed by Nolda to the 'nominalization of adjectives', which uses *word-form* conversion in S, necessarily with a *syntactic* basis (an adjective form or an adjectival group) but resulting in a fully specified word form, i.e., in a *syntactic* unit.

Third, the word-formation literature provides a wealth of empirical information on individual subcases of word-formation processes in idiolect systems S. In the present Section 6, only *major* subcases have been considered. Secs 3 to 5 also characterize some other subcases, such as processes of determinative compounding in S and processes of copulative compounding in S; these are not yet reconstructed in Sec. 6. The claim has been, though, that a complete theory of word-formation processes—to the extent that such a theory will ever be possible—can be developed, in full empirical detail, using as a starting point the part of the Process Model of Word Formation that has now been characterized.

The Model would still be incomplete if it were to be restricted to a theory of word-formation *processes*: the Model is to cover *the formation of non-basic lexical words in their non-inflectional aspect*; word-formation processes are the basis of, but not identical to, lexical-word formation. I briefly indicate what must still be added.

6.5 b. Defining 'word formation': item formation, word formation, stem formation (Definitions 19 to 21)

Word formation in S was identified in (1e) with a certain relation (not: a function); by (1f), the relation is assigned to S by a function called *word formation* ('word formation per se'). What is still missing is the theoretical basis for conceiving word formation per se in this way.

In dealing with word formation in an idiolect system S we are also dealing with the formation of stem lexemes: stem-form processes in S yield fully specified stem forms of S that are related to stem lexemes in the same way as fully specified word forms are related to lexical words. Word formation in S may therefore be construed as a subrelation of a more basic relation, '(lexical-)item formation in S', with 'stem-lexeme formation in S' as another subrelation. We define accordingly:[47]

[47] Nolda's 'Wortbildungsrelationen' (Nolda 2012b: Sec. 6.2.3) are accounted for in Lieb (2011/2012) by the relation 'word formation in S', except that Nolda also includes stem lexemes in his 'word-formation relations'. I now assume the lexical-item formation in S to cover both word-formation and stem-formation, going beyond Lieb (2011/2012).

Definition 19

[*lexical-*]*item formation* (*itf*) = the function that assigns to any S the set of all quadruples $\langle\langle P, b\rangle, \langle P_1, b_1\rangle, \langle P_2, b_2\rangle, \varphi\rangle$ such that:

a. $\langle P, b\rangle$ is a lexical word of S or is a stem lexeme of S;

b. φ is a word-formation process in S;

c. '$\langle P, b\rangle$ is properly related in S to $\langle P_1, b_1\rangle$ and $\langle P_2, b_2\rangle$ by φ'.

Definitions 20 and 21

***word formation* (*wf*) / *stem*[*-lexeme*] *formation* (*stf*)** = the function that assigns to any S the set of all quadruples $\langle\langle P, b\rangle, \langle P_1, b_1\rangle, \langle P_2, b_2\rangle, \varphi\rangle$ such that:

a. $\langle\langle P, b\rangle, \langle P_1, b_1\rangle, \langle P_2, b_2\rangle, \varphi\rangle \in \text{itf}_S$;

b. $\langle P, b\rangle$ is a lexical word of S / a stem lexeme of S.

Condition (c) in D19 remains to be explained.

6.5 c. Being 'properly related'

Consider any element $\langle\langle P, b\rangle, \langle P_1, b_1\rangle, \langle P_2, b_2\rangle, \varphi\rangle$ of the item formation in any given S. The first pair in the element is either a lexical word or a stem lexeme of S; the other two are lexical words, stem or affix lexemes, or pseudo-words (see Sec. 2.6 b, above, assuming that pseudo-lexemes are not allowed). For a given lexical word, there may but need not be a corresponding stem lexeme; there is either none or exactly one. There are no stem lexemes for pseudo-words.

The fourth component of the quadruple is a relation φ that is a word-formation process in S, i.e., is a non-empty proper or improper subcase of one of the three basic word-formation processes in S, the compounding, derivation, and conversion processes in S. Each one of the three basic processes has two major subcases, its stem-form part and its word-form part; these are word-formation processes unless empty. The stem-form process and its subcases yield fully specified stem forms of S but need not apply to such forms; the word-form process and its subcases yield fully specified word forms of S but do not apply to fully specified morphological units. In this way, φ partly determines if a relation can be established directly between the three pairs $\langle P, b\rangle$, $\langle P_1, b_1\rangle$, and $\langle P_2, b_2\rangle$, or if their stem lexemes (in case there are such) must intervene. If φ is a conversion subcase, added triples must be 'empty'.

Keeping all this in mind, we may interpret condition (c) in D19 as follows, retaining the variables in conditions (a) and (b):

c. there are f, J, f_1, J_1, f_2, J_2, b' and χ such that:

(i) $\langle f, J\rangle \in P$, or $\langle f, J\rangle$ is an element of the paradigm of some stem lexeme of $\langle P, b\rangle$,

(ii) $\langle f_1, J_1\rangle \in P_1$, or $\langle f_1, J_1\rangle$ is an element of the paradigm of some stem lexeme of $\langle P_1, b_1\rangle$,

(iii) $\langle f_2, J_2\rangle \in P_2$, or $\langle f_2, J_2\rangle$ is an element of the paradigm of some stem lexeme of $\langle P_2, b_2\rangle$, or $\langle f_2, J_2\rangle = \langle f^0, \varnothing\rangle$,

(iv) $\langle f_1, J_1, b_1\rangle$ is the basic triple of χ,

(v) $b' = b_2$, or $b' = b^0$,

(vi) $\langle f_2, J_2, b'\rangle$ is the added triple of χ,

(vii) $\varphi(\chi) = \langle f, J, b\rangle$.

Given the general notion of item formation in S to cover both word formation and stem-lexeme formation in S, we may also generalize the notion of category assignment in S that was previously restricted to word formation: we now consider the 'lexical-category assignment' in S irrespective of the categories being stem-lexeme or lexical-word categories, as follows.[48]

6.5 d. Lexical-category assignment, word-category assignment and stem-category assignment in S (Axiom 5, Definitions 22 to 24)

We tentatively assume the *lexical-category assignment* as a second morphosyntactic component of any idiolect system S, in addition to its word-formation base. This is to be a function ψ, i.e., a function whose arguments have the ontological status of quadruples $\langle\langle P, b\rangle, \langle P_1, b_1\rangle, \langle P_2, b_2\rangle, \varphi\rangle$ and whose values are sets O of sets L of pairs ⟨P, b⟩ that have the ontological status of lexical words and lexemes:

Axiom 5 (Axiom of the Lexical-Category Assignment) (tentative)

For any S, there is exactly one function ψ such that:

a. ψ is a component of S;
b. the domain of ψ is the lexical-item formation in S;
c. for each argument $\langle\langle P, b\rangle, \langle P_1, b_1\rangle, \langle P_2, b_2\rangle, \varphi\rangle$ of ψ, the value of ψ for the argument is a set O of lexical categories L of S such that ⟨P, b⟩ belongs to each L in O.

ψ may be empty. Since ⟨P, b⟩ is either a stem lexeme or a lexical word and lexical categories 'do not mix item types', it follows from (c) that O is either a set of stem-lexeme categories or a set of lexical-word categories.—The axiom is the basis for the following definitions:

Definition 22

The lexical-category assignment of S = the ψ that satisfies A5 for S.

Definitions 23 and 24

The word-category assignment of S ***/ the stem-category assignment of*** S = the function ψ such that:

a. the domain of ψ is the set of all arguments $\langle\langle P, b\rangle, \langle P_1, b_1\rangle, \langle P_2, b_2\rangle, \varphi\rangle$ of the lexical category assignment of S such that ⟨P, b⟩ is a lexical word of S / is a stem-lexeme of S;
b. for each argument $\langle\langle P, b\rangle, \langle P_1, b_1\rangle, \langle P_2, b_2\rangle, \varphi\rangle$ of ψ, the value of ψ for the argument is the value of the lexical-category assignment of S for the argument.

Axiom 5 is tentative for the following reason. Both the lexical-category and the word-category assignment for S are to identify, via their values, only such lexical categories for a given stem lexeme or lexical word ⟨P, b⟩ that are word-formation related. Following tradition in word-formation studies, essentially the following kinds of categories qualify as candidates if ⟨P, b⟩ is a lexical word: part of speech, grammatical gender of nouns (in case there are such genders), government (both quantitative and qualitative), and inflection class; analogously, for stem-lexeme categories. It is, however, an open question if all rele-

[48] This step was not yet taken in Lieb (2011/2012).

vant categories are already determined for ⟨P, b⟩ by the requirement that ⟨⟨P, b⟩, $\langle P_1, b_1\rangle$, $\langle P_2, b_2\rangle$, φ⟩ should be an argument of word formation (of stem formation) in S, or even more strongly, determined by φ alone. If they are, the lexical-category assignment of S need not be introduced as a separate component of the idiolect system but may be based directly on its lexical-item formation, rendering Axiom 5 superfluous.

Given the notion of word formation in S as defined in D20, a number of traditional concepts relating to 'word-formation types' can be made more precise.

6.5 e. Types of word formation in S (Definitions 25 to 28)

Referring to specific functions φ, *types of word formation* are distinguished as in the following example:

Definition 25

derivation in S $=_{df}$ the set of all ⟨⟨P, b⟩, $\langle P_1, b_1\rangle$, $\langle P_2, b_2\rangle$, φ⟩ such that:

a. ⟨⟨P, b⟩, $\langle P_1, b_1\rangle$, $\langle P_2, b_2\rangle$, φ⟩ ∈ wf(S);

b. φ is a subcase of the derivation process in S.[49]

Terms like 'derivation in S', 'compounding in S', and 'conversion in S' that denote word-formation *relations* in S must not be confused with *process* terms like 'the derivation process in S', 'the derivation process per se', 'the derivation process', 'stem-form derivation in S', where parts like 'derivation' are inseparable from the rest of the term and have no independent meaning.

We may also define traditional terms of the type '⟨P, b⟩ is derived' (meaning: obtained through derivation) or '⟨P, b⟩ is derived from $\langle P_1, b_1\rangle$', relativizing them to idiolect systems S:

Definition 26

⟨P, b⟩ is ***derived in*** S ***from*** $\langle P_1, b_1\rangle$ ***using*** $\langle P_2, b_2\rangle$ iff for some φ, ⟨⟨P, b⟩, $\langle P_1, b_1\rangle$, $\langle P_2, b_2\rangle$, φ⟩ ∈ derivation in S.

Definition 27

⟨P, b⟩ is ***derived in*** S ***from*** $\langle P_1, b_1\rangle$ iff for some $\langle P_2, b_2\rangle$, ⟨P, b⟩ is derived in S from $\langle P_1, b_1\rangle$ using $\langle P_2, b_2\rangle$.

Definition 28

⟨P, b⟩ is ***derived*** (is a ***derived word***) ***in*** S iff for some $\langle P_1, b_1\rangle$, ⟨P, b⟩ is derived in S from $\langle P_1, b_1\rangle$.

Definitions 25 to 28 reconstruct one frequent usage of word-formation terms. In particular, it can now be seen from the definitions how word formation in S imposes a certain *relational structure* (D25 to D27) and a *class structure* (D28) on the lexicon. These structures are derivative, they are not, on our account, components of the lexicon.

The Process Model as it has now been outlined is part of a theory of language. Let us briefly consider how the model may be used in actual language description. Strictly

[49] This exemplifies how a word-formation *relation* of S may be determined by a word-formation *process* in S, a feature already present in earlier versions of the Process Model, and also independently introduced by Nolda into his PR model (2012b: Sec. 6.2.3).

speaking, we are now proceeding from a theory of language to a theory of grammars (language descriptions), construed as a sister theory of the theory of language.[50]

[50] The following sketch outlines a use of the Process Model in language description that differs from the use emphasized by Nolda (2012b) for his own word-formation theory: using it for explanations in a classical, Hempel-Oppenheim sense. This is of course not excluded for the Process Model either.

7 Using the Process Model in word-formation description

7.1 Introduction

Suppose that we want to describe word-formation in the idiolect systems of a given language or language variety D, construing D as a set of idiolects that each have a system; this is tantamount to describing ‘word-formation in D’. For this purpose we require a grammar of D with a word-formation part. The grammar again is a theory: a theory of D; in our case, what is to be used in the grammar of D, is the theory of language containing the Process Model of Word Formation. The theory of language must be available in the grammar without being a literal part of it; we are confronted here with a problem of ‘theory integration’. In Lieb (1983: Part G), such problems are studied extensively, and various kinds of theory integration are distinguished in relation to linguistic theories. The type that is relevant here is *presupposition*: the grammar *presupposes*—in the sense of Lieb (1983: 425)—the theory of language that contains the Process Model of Word Formation.

Without going into details, this implies that the terminology and the axioms, theorems, and definitions of the theory of language can be used freely in those sentences of the grammar which are to describe word formation in the idiolect systems of D. In particular, where terms from the Process Model appear in sentences of the grammar, the terms are to be understood as defined in the theory of language; they are not newly defined in the grammar. Moreover, the terms, axioms, theorems, and definitions of the Process Model are available for deductions in the grammar where the grammar deals with word-formation in D. At the same time, the grammar of D must contain sentences that isolate properties specific to D, in particular, sentences that identify word-formation and word-formation processes in the idiolect systems of D. Such sentences do not belong to the presupposed theory of language or its Process Model part but are sentences solely of the grammar that formulate empirical assumptions on D; the sentences are *empirically true or false*. This holds of *any* grammar that presupposes our theory of language; it makes for maximal comparability of word-formation studies, regardless of language, and also allows us to separate the general from the specific in describing word-formation in a given language.

For exemplification, consider the *identification of basic word-formation processes* in the idiolect systems of a language or language variety, arguably the most important task in describing word-formation.

7.2 Process identification

7.2 a. Identification sentences

Suppose that we are dealing in a grammar G with the basic word-formation processes in the idiolect systems S of a language or language variety D. These processes, say, der_S, may then be identified by G through a sentence of G of the following form (relevant 13-tuples denoted in abbreviated form, as ‘$\langle f_1, \ldots, \varepsilon\rangle$’):

Format of identification sentences (informal example)

For any idiolect system S of D and any $f_1, \ldots, \varepsilon$, $der_S(f_1, \ldots, \varepsilon) = \langle f, J, b\rangle$ iff: one of the following conditions is satisfied for $\langle f_1, \ldots, \varepsilon\rangle$: ….., and $\langle f_1, \ldots, \varepsilon\rangle$ determines $\langle f, J, b\rangle$.

(For ‘determines’, see Sec. 6.1 b, above.) The ‘conditions’, indicated by the dots after the colon, must be non-logical and must be *specific to* D. This means: it does not yet follow from sentences of the presupposed theory of language—its axioms, theorems, or definitions—that the conditions are satisfied. In particular, this does not follow from the definition of the process name (‘der’). Nor does it follow from sentences of the presupposed theory of language and a sentence of the grammar of D stating that D is (a variety of) a language. Consider the following (incomplete) example.

We choose Standard British English (SBE) as D. Let G be a description, or grammar, of SBE that presupposes a theory of language with a Process Model part and uses this part in its description W of word-formation in SBE. Let the description W be exhaustive. W then contains an *identification sentence for the derivation process in the idiolect systems of SBE*, i.e., contains an axiom or theorem of the following kind, or an equivalent sentence:

Identification sentence in W for the derivation process in SBE systems

For all S, if S is an idiolect system of SBE, then for any $f_1, \ldots, \varepsilon$, f, J, and b:
$der_S(f_1, \ldots, \varepsilon) = \langle f, J, b\rangle$ if and only if (I) and (II):

I. $\langle f_1, \ldots, \varepsilon\rangle$ satisfies one of the following conditions (a_1) to (a_n):

a_1. $\langle f_1, J_1, b_1\rangle$ is a ‘suitable’ fully specified unit of S
and J_1 = {Pres-St, Inf-St}
and $\langle f_2, J_2, b_2\rangle$ is a fully specified affix form of S
and f_2 = un_1
and $\beta_{11} = \beta_{21}$ = zero-shortening$_S$
and $\beta_{12} = \beta_{22}$ = identity
and γ = inverse concatenation
and δ = identity
and ε = annulling;

a_2. …
.
.
a_n. …

II. $\langle f_1, \ldots, \varepsilon\rangle$ determines $\langle f, J, b\rangle$.

Condition (a_1) covers *one* possibility for ‘derivation by *un*$_1$’; cf. Sec. 4.1, above. ‘Suitable’ accounts for the fact that there are restrictions on the basic triples, in particular, semantic ones, in addition to the condition formulated for J_1; these are not spelled out here. For the identification sentence to make sense, it must be implied by each (a_i) in (I) that $\langle f_1, \ldots, \varepsilon\rangle$ is ‘properly constructed’; this is a precondition for using ‘determines’ in (II) (see Sec. 6.1 b, above).

7.2 b. Discussion

I do not wish to give the impression that an identification sentence for der_S in the idiolect systems of SBE must have exactly this form; in an actual description, such a sentence may be formulated more formally or less formally than above, and may be split up and distributed over the descriptive text. Still, an identification sentence of this kind will be implied by any exhaustive description of the derivation process in the SBE idiolect systems.

Three points should be emphasized with respect to identification sentences in a grammar:

(i) They are not definitions of the relevant terms; for example, 'der' is defined, once for all, in the presupposed theory of language (Definition 14 in Sec. 6.3 b), and is used in the above identification sentence in its defined sense.

(ii) Formally, identification sentences are non-logical axioms or theorems of the grammar that are not yet available in the presupposed theory of language and cannot be obtained from sentences of the theory of language plus a sentence of the grammar stating that its object is a language, or variety of a language.

(iii) Because of (ii), an identification sentence is a sentence newly introduced by the grammar and is a sentence that is—in contrast to a definition—empirically true or false.

The identification sentences for word-formation processes in the systems of different languages never make identical claims; typically, they will proceed by case distinctions even when dealing with a single language or language variety, as indicated above. This does *not* mean, though, that we cannot adequately define terms like 'der' or 'the derivation process' in a theory of language, as names of functions that apply to arbitrary idiolect systems S and assign to each system a function φ (possibly empty). I submit that the definitions proposed in the present essay are indeed adequate.

Not being a definition, an identification sentence in a grammar does not define any one of the terms it uses but presupposes that the terms are already understood, mostly as defined or definable in a presupposed theory of language; if undefined, the term is used in the grammar as a basic term of the theory of language, where it must still be interpreted. Insisting on meaningful identification sentences but denying the possibility of general definitions—a deep-rooted tendency in linguistics, largely due to a confused view of 'definition'—is, in many cases, self-contradictory.

Obviously, the number *n* of cases that must be distinguished in an exhaustive identification sentence for the derivation process in SBE systems is quite large. It is in this context that a notion of *derivation rule*—more generally, *word-formation rule*—may be introduced.

7.3 Word-formation rules

Word-formation rules are to be sentences of a grammar that serve to partially identify word-formation processes in the idiolect systems of the language of which the grammar is a theory (word-formation rules in this sense are not 'rules for forming words' but 'grammar rules for identifying processes involved in word formation'). Such rules can be given roughly the following form, exemplified by derivation rules:

Format of word-formation rules (informal example)

For any idiolect system S of D and any f_1, …, ε, f, J, b, if

a. ⟨f_1, …, ε⟩ satisfies condition ….., and

b. ⟨f_1, …, ε⟩ determines ⟨f, J, b⟩,

then

c. $\langle f, J, b \rangle = der_S(f_1, \ldots, \varepsilon)$.

Again, (a) must imply that $\langle f_1, \ldots, \varepsilon \rangle$ is 'properly constructed' for (b) to make sense.

Let us use the condition (a_1) from the identification sentence in Sec. 7.2 a to fill in the gap left in (a) after 'condition'; take D to be Standard British English. We then obtain the following example:

Example of a word-formation rule

For all S, if S is an idiolect system of SBE, then for all $f_1, \ldots, \varepsilon$, f, J, b, if

a. $\langle f_1, \ldots, \varepsilon \rangle$ satisfies condition (a_1), and

b. $\langle f_1, \ldots, \varepsilon \rangle$ determines $\langle f, J, b \rangle$,

then

c. $\langle f, J, b \rangle = der_S(f_1, \ldots, \varepsilon)$.

From a *formal* point of view, this is either an axiom or a theorem of the SBE grammar. A comparison of the rule with the above identification sentence shows an obvious relationship: the rule is a logical consequence of the identification sentence. This can be generalized: in a grammar, an identification sentence for a word-formation process—an axiom or theorem of the grammar—should precede corresponding rules; these may then be derived as theorems from the identification sentence. (Rules need not be as restrictive as the one above.)

From a *methodological* point of view, however, the order is reversed. Rules for a word-formation process in idiolect systems will be usually formulated before the identification sentence has been established: we work our way up from the many specific cases to the identification sentence. As long as no identification sentence has been formulated from which the rule would follow, the rule has the methodological status of an individual empirical hypothesis, of limited range and subject to revision or eventual rejection. Only when we feel certain that all relevant cases have been found will we attempt to formulate an identification sentence, using the rules to obtain a disjunction of conditions as in (I), above; it is essentially this disjunction, representing a single condition, that allows us to identify the word-formation process. From a methodological point of view, the identification sentence continues to have the status of an empirical hypothesis, less restrictive than a rule but still subject to revision or rejection.

This concludes the present outline of the Process Model of Word Formation. Typically, several competing theories are possible for any given empirical domain, not necessarily equivalent in predictive power; and even if two theories are equivalent in this respect, they may differ in adequacy, in more than one sense. Indeed, there is a model of word formation, closely related to the Process Model, with which the Process Model may be confronted: the *Pattern and Restriction* model developed by Andreas Nolda.

8 Alternatives: remarks on a daughter model

8.1 Introduction

An earlier version of the Process Model of Word Formation was used by Andreas Nolda to develop his own *Pattern and Restriction* (PR) model of word formation, which retains essential features of the Process Model (for a first impression, there is a brief English outline in Nolda 2012a: Sec. 1; a carefully worked out axiomatic formulation in German is included as Appendix B in Nolda 2012b; I am going to translate Nolda's German terms into English). Nolda uses different variables from the ones employed in the Process Model, or uses the same variables differently (for example,'S' stands for any 'linguistic system': any system of an idiolect, *or system for a variety or language*), and the structure of his axiomatic system differs from the structure of the system outlined in Sec. 6. There are major deviations from the Process Model; still, many details of the Process Model reappear in Nolda's theory, and the basic orientation is by and large the same, despite a difference in the role assigned to word-formation processes: in the Process Model, word-formation processes in S serve to directly identify the non-basic lexical words of S in their non-inflectional aspect whereas in the PR model the relation is less direct.

This is not the place for a detailed comparison of the two models. However, essential differences between them can be specified as *five major changes* made in proceeding from the Process Model, in the form it takes in the present essay, to the PR model. (There are other changes, such as, in the PR model, relaxing an important condition on paradigms in order to deal with syncretism, or basing meaning-change functions directly on functions that change concept *intensions* rather than proceeding from *extensions*; but the five changes appear to be the major ones.) The changes will be characterized informally, indicating for each one why it is rejected in the Process Model. Characterizing the changes also serves to throw additional light on the Process Model.

8.2 Change 1: anchoring word-formation processes in S differently

8.2 a. The change

In the ***Process Model****, the word-formation processes in S are the non-empty (proper or improper) subcases of at most three potential word-formation processes in S: the compounding process in S, the derivation process in S, and the conversion process in S, distinguished by properties of the added triples; these are the elements of the word-formation component of S (its 'word-formation base').*

In the ***PR model****, the word-formation component of S is construed as a set of functions* F^n *that are roughly analogous to our potential word-formation processes in S; for each n,* $n > 0$*, there is at most one such function in the set. The* ***n-place word-formation processes in S****, roughly analogous to word-formation processes in S in our sense, are functions* ρ^n*,* $n > 0$*, that are ultimately based on the functions* F^n *in the word-formation component of S, making use of these functions in a specific way; on our approach, n would be the number of triples in the basis excluding affix triples and 'empty' triples but still allowing for more than one triple in the basis. The notion of an n-place word-formation process in S is then* *used for a definition schema that defines n terms of '**n-place compounding in S**' to mean, simply, the n-place word-formation processes in S such that n > 1; similarly, '**derivation in S in a broad sense**' for the one-place word-formation process in S, '**derivation in S in a narrow sense**' for that subfunction of derivation in S in a broad sense that involves segmental change, and '**conversion in S**' for that subfunction of derivation in S in a broad sense that does not involve segmental change.*

8.2 b. Rejecting the change

The Process Model and the PR model agree in anchoring word-formation processes in a specific morphosyntactic component of the system S, different from both the morphological and the syntactic components. Assuming such a component corresponds to one traditional treatment of word formation by which word formation is dealt with separately from, and in addition to, morphology and syntax.

In the earlier versions of the Process Model, the morphological processes and the syntactic word-formation processes in S were kept separate and represented in, respectively, a component of the morphological and a component of the syntactic subsystem of S. This has now been changed; in this respect, the gap between the Process Model and the PR model has narrowed: we are now assuming *three basic processes*, each one with two major subcases, one morphological, the other syntactic, assigning the three basic processes to a single morphosyntactic component of S.

Still, important differences remain between the ways in which the Process Model and the PR model anchor word-formation processes in S in the system: in the Process Model but not in the PR model, only three basic processes in S are currently recognized (any number of compounding processes is allowed in the PR model), and the basic processes are represented directly in the word-formation component; the Process Model is, in this respect, simpler, and ties in more directly with the basic types of word formation traditionally distinguished in linguistics. In summary, the anchoring approach followed in the PR model continues to be partly rejected in the Process Model.

8.3 Change 2: using different 'basic objects'

8.3 a. The change

*In the **Process Model**, fully specified units ⟨f, J, b⟩ of some idiolect system—informally, '**basic objects**'—supply the first three components and, typically, the second three components of the basis of an argument when a word-formation process in S is applied; the result is a fully specified unit of S.*

*In the **PR model**, fully specified units are replaced as basic objects by 'word-formation instances in S', quadruples ⟨f, P, L, c⟩ where ⟨f, P⟩ is essentially like ⟨f, J⟩ in a fully specified unit ⟨f, J, b⟩: ⟨f, P⟩ is an element of the paradigm of some presupposed lexical item of S (a 'lexical stem' or a lexical word, affixes are excluded in this context), and c is a concept said to be 'compatible with' the meaning of the lexical item; L is a set of lexical categories that is a (partial) categorization of the presupposed lexical item. There are no basic objects to correspond to fully specified affix forms or to triples with 'empty' components.*

8.3 b. Rejecting the change

In the Process Model, fully specified units do not include a categorization L of a presupposed lexical item, mainly for the following reason. Fully specified units are *given* in S independently of lexical items; in particular, fully specified stem forms or word forms, including categorizations J consisting of *word-form* categories, are ultimately given through a paradigm basis; in such a basis, *lexical-word* categories do not appear. Word-formation processes in S that do not involve lexical-word categories can be used in the *identification* of stem lexemes and lexical words. Generally, word-formation in S as it is construed in the Process Model (wf_S) is to account for the *major aspect of word formation in a traditional sense*, viz. identification of non-basic lexical words excluding inflection. The *minor aspect*—partial determination of a word's membership in lexical categories—is accounted for only after the word has been identified (see Secs 1.5 c, 3.3 b and 6.5 d, above, for details).

Moreover, in the Process Model the basic objects in word-formation processes are related to lexemes and lexical words directly, not indirectly as in the PR model where 'underspecification' is allowed in setting up basic objects: each component of a basic object may have to be made more specific before it turns into something that helps to directly identify a lexeme or lexical word; this is specifically true of the quadruple that *results from* applying a word-formation process.

In particular, applying a word-formation process in the Process Model directly identifies lexeme and word meanings b in contrast to the PR model, where the semantic component c in a result quadruple need only be 'compatible with' the lexical meaning of a corresponding paradigm form: the meaning itself may be 'underspecified' by c. This is an unnecessary complication since the lexical meanings must still be identified; excluding their actual determination as 'idiosyncratic' from word-formation processes, as Nolda does, is a highly dubious move. It seems to be motivated by a failure to directly attack the apparent semantic indeterminacy of determinative compounds (see Sec. 3.4 b, above, *Remark on spec*), and by not representing (Nolda 2012b: Ch. 7, esp. fn. 12) the place-number of result meanings (such as a two-place verbal concept vs. a three-place one) directly in semantic functions, thus rejecting, for no good reason, *families* of semantic functions that differ only in this respect, whereas such families are allowed and used in the Process Model.

Generally, 'underspecifying' the components of process results clashes with the direct-identification aim adopted in the Process Model, and is therefore rejected.

A seemingly innocuous step taken in the PR model has vast consequences: this is the exclusion of basic objects that correspond to fully specified affix forms, or correspond to triples with 'empty' components. Due to this step, there is no basic object in the PR model to correspond to an added triple in dealing with derivation or conversion, whereas there *are* such objects in the case of compounding processes in S. As a consequence it is hard, if not impossible, to assign derivation processes what should be their proper place in a process cline (Secs 5.1, 6.4 b, above): intermediate between compounding processes and conversion processes. On the conception adopted in the Process Model, triples with empty components are allowed as long as the empty set is not introduced as a linguistic unit.

In summary, Change 2 is rejected in the Process Model because it clashes with the aim of direct word-identification by weakening the close connection between the results of word-formation processes on the one hand and lexical items on the other, and because it obscures the process cline.

8.4 Change 3: introducing n-place word-formation processes

8.4 a. The change

*In the **Process Model**, all word-formation processes in S have arguments with a bipartite basis: the first three components form the basic triple, the second three the added triple; at least the basic triple is (normally) a fully specified unit. In this sense, all word-formation processes in S are bipartite, or 'two-place'.*

*In the **PR model**, a basis in our sense is replaced by a basic object ⟨f, P, L, c⟩, or by an n-tuple of basic objects for n > 1; for the sake of brevity: a basis in our sense is replaced by an n-tuple of basic objects for n > 0, allowing for 1-tuples. The n-tuples are introduced as elements of sets called 'n-place base restrictions'. Let us therefore call the n-tuples '**restriction tuples**'. Only indirectly does a word-formation process in S apply to a restriction tuple, still, it applies (cf. below, Change 4). A restriction tuple has n components (consists of n basic objects), for n > 0. Since n may vary, it is necessary in the PR model to relativize word-formation processes in S to n, the number of components of the restriction tuples to which a process is meant to indirectly apply: no longer are all word-formation processes in S uniformly 'two-place', are, simply, word-formation processes in S; instead, we are now dealing with **n-place word-formation processes in S**, for n > 0.*

8.4 b. Rejecting the change

Change 3 is rejected in the Process Model for the following reason. The change would vastly complicate the theory, as demonstrated by the PR model; in order to account for the change, a large part of the theoretical vocabulary in the PR model has to be relativized to n, too, resulting in families t^n of terms rather than simply terms t, with a corresponding increase in the types of entities that the theory must deal with.

The relativization to n in the PR model appears to have been generated by two theoretical decisions and one empirical assumption. The *theoretical decisions* are: exclusion of basic objects that correspond to fully specified affix forms, and exclusion of basic objects with 'empty' components (see above, Change 2). The first decision excludes standard derivation functions as being uniformly 'two-place', and the second, conversion functions. However, neither decision *has* to be taken. The *empirical assumption* is this: there must be, or should be, functions of copulative compounding in S that apply, directly or indirectly, to n-tuples of basic objects such that n > 2. This assumption appears to be false (see Sec. 3.6, above). If it is, n-place word-formation processes are empirically unmotivated in the PR-model.

Since the theoretical decisions need not be taken and the empirical assumption may well be false, Change 3 is unnecessary, and is not made in the Process Model due to the complications it creates.

8.5 Change 4: reconceiving the arguments of word-formation processes

8.5 a. The change

*In the **Process Model**, an argument of a word-formation process in S is a 13-tuple that comprises both a basis (components 1 to 6, forming two basic objects) and a construction mode (components 7 to 13: seven functions).*

*In the **PR model**, an argument of an n-place word-formation process in S, say $comp^nS$, is a quadruple of four n-place functions, called a '**word-formation pattern**', that corresponds to our construction mode alone, without the basis. The four functions operate on, respectively, n-tuples of forms f, n-tuples of categorizations P, n-tuples of categorizations L, and n-tuples of concepts c. An n-place restriction tuple of n basic objects that corresponds to a basis in our sense is introduced only after another step has been taken. First, the word-formation process, $comp^nS$, takes the 'word-formation pattern' of four functions as an argument and assigns to it another, intermediate function, also n-place, that is called '**the operation specified by**' the pattern. This intermediate function then takes the **n-place restriction tuple** as an argument and assigns to it a basic object ⟨f, P, L, c⟩ by having the four functions in the word-formation pattern operate on the relevant components of the n basic objects in the restriction tuple: operate on, respectively, f_1 to f_n, P_1 to P_n, L_1 to L_n, and c_1 to c_n. The **quadruple ⟨f, P, L, c⟩** so obtained corresponds to the fully specified unit ⟨f, J, b⟩ that would be a result triple in the Process Model.*

8.5 b. Rejecting the change

Change 4 is rejected in the Process Model for the following reasons. *First*, restricting the arguments of a word formation process in S to the construction mode and then introducing an intermediate function to take care of the basis is simply not necessary on either empirical or theoretical grounds; the resulting formal und conceptual increase in complexity, obvious from the PR model, is therefore unnecessary, too.[51] *Second*, in characterizing arbitrary subcases of a word-formation process in S, it is preferable to have *direct and simultaneous access* in an argument to both the components of the basis and the components of the construction mode, which excludes making the change (see Sec. 6.2 c, *Remarks on D8*).

8.6 Change 5: using different functions that manipulate forms, generally and for conversion

8.6 a. The change

*In the **Process Model**, functions of shortening, form change, and arrangement are separately represented in the construction mode and are applied to forms in this order. When a **conversion process in S** is applied, the basic unit may turn into a shorter sequence due to the basic shortening function, and its members may differ from the original ones, segmentally or suprasegmentally, due to the basic form-change function.*

*In the **PR model**, functions of shortening, form change and arrangement are not individually represented. Instead, the first component of a 'word-formation pattern'—call it '**the form manipulator**'—is construed as an n-place function assigning a single form to the n forms of a given n-place restriction tuple, accounting jointly for the effects of shortening, form-change and arrangement functions in our sense, using the notion of functional product to obtain the single form manipulator.*

[51] This is indirectly admitted by Nolda (2012b: Sec. 6.2.2), who claims of the two above ways of construing word-formation processes that 'in a sense, these are notational variants'. However, no definite advantage of Nolda's version is forthcoming to make up for the added complexity.

Conversion in S *is construed as that subfunction of (one-place) derivation in S in a broad sense for which the form manipulator leaves the form f* ***segmentally unchanged*** *when the derivation function applies to a given quadruple ⟨f, P, L, c⟩; suprasegmental changes are allowed.* ***Derivation in S in a narrow sense*** *is the analogous subfunction for which the form manipulator does not leave the form f segmentally unchanged.*

8.6 b. Rejecting the change

There are *four major aspects of manipulating forms* by a word-formation process that may be distinguished: addition, shortening, segmental and suprasegmental change, and combination. In the Process Model, addition is accounted for by the universal presence of an added triple. The other three aspects are kept apart and represented separately by functions that appear as individual components of the arguments of word-formation processes in S and always apply in the same fixed order. Due to this approach different word-formation processes in S, and different applications of a single word-formation process, can be compared with respect to all three aspects in a way that is both *direct and maximally explicit*. This is not possible in the PR model, due to the fusional nature of its form manipulators; they are to account even for the aspect of addition, excepting the compounding process. The PR account of form manipulation (inspired by Beard's conception of spelling operations: Beard 1995: Ch. 3) is therefore rejected.

The reasons for rejecting the PR view of conversion in S are given in Secs 5.4 e (Alternative 2) and 2.1 f (rejection of conversion forms and categories); they need not be repeated here.

Concluding remark. The confrontation of the Process Model with the PR model, critical of the latter, is to explain and justify decisions that are taken in the Process Model but are not made, or made differently, in the PR model. This may give the impression of the PR model having been evaluated negatively—wrongly so: the Pattern and Restriction model is an important contribution to word-formation theory, more advanced than most, providing welcome support to adopting word-formation theories that are process-driven. Taken in conjunction, the two models demonstrate how two related theories of word formation that partly differ in dealing with their domain, may throw light on it by their very differences.

References

Aikhenvald, Alexandra & Pieter Muysken (eds.). 2010. *Multi-verb constructions: a view from the Americas*. Amsterdam: Brill.

Aikhenvald, Alexandra Y. & Robert M. W. Dixon (eds.). 2006. *Serial verb constructions: A cross-linguistic typology*. 2 vols. Oxford : Oxford University Press.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. 2005. *Aspects of the theory of clitics* (Oxford Studies in Theoretical Linguistics 11). Oxford: Oxford University Press.

Arcodia, Giorgio F., Nicola Grandi & Bernhard Wälchli. 2010. Coordination in compounding. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 177-198. Amsterdam/Philadelphia: Benjamins.

Aronoff, Mark. 1976. *Word formation in Generative Grammar* (Linguistic Inquiry Monographs 1). Cambridge, Mass.: MIT Press.

Aronoff, Mark. 1994. *Morphology by itself. Stems and inflectional classes* (Linguistic Inquiry Monograph 22.) Cambridge, Massachusetts; London, England: The MIT Press.

Aronoff, Mark. 2000. Generative Grammar. In Geert Booij & Christian Lehmann & Joachim Mugdan (eds.), *Morphologie. Morphology. An international handbook on inflection and word-formation* (Handbücher zur Sprach- und Kommunikationswissenschaft 17.1), 194-209. Berlin/New York: de Gruyter.

Aronoff, Mark. 2007. In the beginning was the word. *Language* 83(4). 803-830.

Aronoff, Mark. 2012. Morphological stems: what William of Ockham really said. *Word Structure* 5(1). 28-51.

Baermann, Matthew & Greville C. Corbett. 2012. Stem alternations and multiple exponence. *Word Structure* 5(1). 52-68.

Balteiro, Isabel. 2007a. *A contribution to the study of conversion in English*. Münster: Waxmann.

Balteiro, Isabel. 2007b. *The directionality of conversion in English. A dia-synchronic study* (Linguistic Insights. Studies in Language and Communication 59). Bern etc.: Peter Lang.

Barz, Irmhild. 2005. Die Wortbildung. In Dudenredaktion (eds), *Die Grammatik* (Duden Vol. 4), 7th ed., completely new and enlarged, 641-772. Bibliographisches Institut: Mannheim.

Beard, Robert. 1995. *Lexeme-Morpheme Base Morphology. A general theory of inflection and word formation* (SUNY Series in Linguistics). Albany: State University of New York Press.

Becker, Thomas. 1993. Morphologische Ersetzungsbildungen im Deutschen. *Zeitschrift für Sprachwissenschaft* 12. 185-217.

Bell, Melanie J. & Ingo Plag. 2012. Informativeness is a determinant of compound stress in English. *Journal of Linguistics* 48 (3). 485-520.

Berg, Thomas. 2011. The clipping of common and proper nouns. *Word Structure* 4(1). 1-20.

Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 42. 531-573.

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.

Booij, Geert. 2007. *The grammar of words. An introduction to linguistic morphology*. 2nd edn. Oxford: Oxford University Press.

Booij, Geert. 2010. *Construction Morphology.* Oxford: Oxford University Press.

Booij, Geert. 2012. Construction Morphology, a brief introduction. *Morphology* 22. 343-346.

Boye, Kasper & Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88(1). 1-44.

Budde, Monika. (2012). Wörter und Sätze. Anmerkungen zur disziplinären Struktur der Sprachwissenschaft. In Ernest W. B. Hess-Lüttich (ed.), *Sign Culture Zeichen Kultur*. Würzburg: Königshausen & Neumann. 51-87.

Budde, Monika. 2000. *Wortarten. Definition und Identifikation*. Ph. D. dissertation. Berlin: Freie Universität Berlin. [microfiche]

Carston, Robyn. 2006. Relevance Theory the saying/implicating distinction. In Laurence R. Horn & Gregory Ward (eds.), *The handbook of pragmatics* (Blackwell Handbooks in Linguistics), 633-656. Oxford: Blackwell. (First published in 2004.)

Dressler, Wolfgang U. 2011. Foreword. In Stela Manova, *Understanding morphological rules. With special emphasis on conversion and subtraction in Bulgarian, Russian and Serbo-Croation* (Studies in Morphology 1), v-vii. Dordrecht: Springer.

Drude, Sebastian. 2004. *Wörterbuchinterpretation. Integrative Lexikographie am Beispiel des Guaraní* (Lexicographica, Series maior 120). Tübingen: Max Niemeyer.

Drude, Sebastian. 2010. 'Derivational verbs' and other multiple-verb constructions in Awetí and Tupí-Guaraní. In Alexandra Aikhenvald & Pieter Muysken (eds.), *Multi-verb constructions: a view from the Americas*, 259-304. Amsterdam: Brill.

Eins, Wieland. 2008. *Muster und Konstituenten der Lehnwortbildung. Das Konfix-Konzept und seine Grenzen* (Germanistische Linguistik, Monographien 23). Hildesheim etc.: Georg Olms Verlag.

Eisenberg, Peter. 1998. *Grundriß der deutschen Grammatik.* Vol. 1: *Das Wort*. Stuttgart/Weimar: Metzler.

Eisenberg, Peter. 2006. *Grundriß der deutschen Grammatik.* Vol. 1: *Das Wort.* 3rd, corr. ed., Stuttgart/Weimar: Metzler.

Eisenberg, Peter. 2012. *Das Fremdwort im Deutschen*. (De-Gruyter-Studium). 2nd, rev. ed., Berlin etc.: de Gruyter.

Embick, David & Morris Halle. 2005. On the status of *stems* in morphological theory. In Twan Geerts & Ivo van Ginneken & Haike Jacobs (eds.), *Romance languages and linguistic theory 2003*, 1-31. Amsterdam: Benjamins.

Eschenlohr, Stefanie. 1999. *Vom Nomen zum Verb. Konversion, Präfigierung und Rückbildung im Deutschen* (Germanistische Linguistik, Monographien 3). Hildesheim etc.: Olms.

Etoré, Murielle. 2007. *Etude comparative des préverbes allemands et des postverbes anglais dans le langage contemporain* (Berkeley Insights in Linguistics and Semiotics 65). New York etc.: Peter Lang.

Fuhrhop, Nanna. 1998. *Grenzfälle morphologischer Einheiten*. (Studien zur deutschen Grammatik 57.) Tübingen: Stauffenburg.

Giegerich, Heinz J. 2009. The English compound stress myth. *Word Structure* 2(1). 1-17.

Haugen, Jason D. & Cathy Hicks Kennard. 2011. Base-dependence in reduplication. *Morphology* 21. 1-29.

Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10. 210-234.

Hoeksema, Jack & Richard D. Janda. 1988. Implications of process morphology for Categorial Grammar. In Richard T. Oehrle & Emmon Bach & Deirdre Wheeler (eds.), *Categorial Grammars and natural language structures* (Studies in Linguistics and Philosophy 32), 199-247. Dordrecht: Reidel.

Hoeksema, Jack. 1985. *Categorial morphology* (Outstanding Dissertations in Linguistics). New York/London: Garland Publishing.

Jackendoff, Ray. 2002a. *Foundations of language.* Oxford: Oxford University Press.

Jackendoff, Ray. 2002b. English particle constructions, the lexicon, and the autonomy of syntax. In Nicole Dehé & Ray Jackendoff & Andrew McIntyre & Silke Urban (eds.), *Verb-particle explorations* (Interface Explorations 1), 67-94. Berlin/New York: Mouton de Gruyter.

Jackendoff, Ray. 2011. What is the human language faculty? Two views. *Language* 87(3). 586-624.

Jacobs, Joachim. 2008. Wozu Konstruktionen? *Linguistische Berichte* 213. 3-44.

Lieb, Hans-Heinrich. 1976. Grammars as theories: The case for axiomatic grammar (Part II). *Theoretical Linguistics* 3. 1-98.

Lieb, Hans-Heinrich. 1980a. Words as syntactic paradigms. In Gunter Brettschneider & Christian Lehmann (eds.), *Wege zur Universalienforschung. Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*, 115-123. Tübingen: Narr.

Lieb, Hans-Heinrich. 1980b. Wortbedeutung: Argumente für eine psychologische Konzeption. *Lingua* 52: 151-182.

Lieb, Hans-Heinrich. 1983. *Integrational linguistics. Vol. I: General Outline* (Current Issues in Linguistic Theory 17). Amsterdam/Philadelphia: Benjamins.

Lieb, Hans-Heinrich. 1992a. Integrational semantics: An integrative view of linguistic meaning. In Maxim Stamenow (ed.), *Current advances in semantic theory* (Current Issues in Linguistic Theory 73), 239-268. Amsterdam/Philadelphia: Benjamins.

Lieb, Hans-Heinrich. 1992b. Die Polyfunktionalität des deutschen Vorgangspassivs. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 45(2). 178-188.

Lieb, Hans-Heinrich. 1993a. *Linguistic variables: Towards a unified theory of linguistic variation* (Current Issues in Linguistic Theory 108). Amsterdam/Philadelphia: Benjamins.

Lieb, Hans-Heinrich. 1993b. Integrational Linguistics. In Joachim Jacobs & Arnim von Stechow & Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax. Ein internationales Handbuch zeitgenössischer Forschung/An international handbook of contemporary research* (Handbücher zur Sprach- und Kommunikationswissenschaft 9.1), 430-468. Berlin/New York: de Gruyter.

Lieb, Hans-Heinrich. 1999. Was ist Wortakzent? Eine Untersuchung am Beispiel des Deutschen. In Wolfgang Schindler & Jürgen Untermann (eds), *Grippe, Kamm und Eulenspiegel: Festschrift für Elmar Seebold zum 65. Geburtstag*, 225–261. Berlin; New York: de Gruyter.

Lieb, Hans-Heinrich. 2005. Notions of paradigm in grammar. In D. Alan Cruse & Franz Hundsnurscher & Michael Job & Peter Rolf Lutzeier (eds.), *Lexikologie. Lexicology. An international handbook on the nature and structure of words and vocabularies* (Handbücher zur Sprach- und Kommunikationswissenschaft 21.2), 1613-1646. Berlin/New York: de Gruyter.

Lieb, Hans-Heinrich. 2006. Wortbildung auf morphologischer und Wortbildung auf syntaktischer Grundlage. Lecture given at the 28th Annual Meeting of the Deutsche Gesellschaft für Sprachwissenschaf at Bielefeld, February 23. Handout.

Lieb, Hans-Heinrich. 2008. The case for two-level phonology: German obstruent tensing and nasal alternation in French. In Robin Sackmann (ed.), *Explorations in Integrational Linguistics. Four essays on German, French, and Guaraní* (Current Issues in

Linguistic Theory 285, Studies in Integrational Linguistics 1), 21-96. Amsterdam/Philadelphia: Benjamins.

Lieb, Hans-Heinrich. 2011. Syntaktische Funktionen integrativ: Eine Skizze (2009). In Xiaoqin Su, *Reflexivität im Chinesischen: Eine integrative Analyse. Mit zwei Anhängen von Hans-Heinrich Lieb*, Appendix A, 249-256. Frankfurt am Main: Peter Lang.

Lieb, Hans-Heinrich. 2011/2012. The Process Model of Word Formation: the conception, and outline of a theory. Ms. (unpubl.)

Lieb, Hans-Heinrich. (in prep.). *The Process Model of Word Formation and Inflection.* Amsterdam/Philadelphia: Benjamins.

Lieber, Rochelle & Pavol Štekauer (eds.). 2009. *The Oxford handbook of compounding*. Oxford: Oxford University Press.

Lieber, Rochelle. 2009. IE, Germanic: English. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 357-369. Mannheim etc.: Dudenverlag.

Lüdeling, Anke. 2001. *On particle verbs and similar constructions in German* (Dissertations in Linguistics). Stanford, Ca: CSLI Publications.

Manova, Stela & Wolfgang U. Dressler. 2005. The morphological technique of conversion in the inflecting-fusional type. In Laurie Bauer & Salvador H. Valera (eds.), *Approaches to conversion/zero-derivation*, 67-101. Münster: Waxmann.

Manova, Stela. 2011. *Understanding morphological rules. With special emphasis on conversion and subtraction in Bulgarian, Russian and Serbo-Croation* (Studies in Morphology 1). Dordrecht: Springer.

Masini, Francesca & Valentina Benigni. 2012. Phrasal lexemes and shortening strategies in Russian: the case for constructions. *Morphology* 22. 417-451.

Matthews, Peter H. 1965. Some concepts in Word-and-Paradigm morphology. *Foundations of Language* 1. 268-289.

Matthews, Peter H. 1972. *Inflectional morphology. A theoretical study based on aspects of Latin verb conjugation*. Cambridge: Cambridge University Press.

Matthews, Peter H. 1974. *Morphology. An introduction to the theory of word-structure*. Cambridge: Cambridge University Press.

Müller, Stefan. 2002. *Complex predicates: verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constrained-Based Lexicalization 13). Stanford, Ca: CSLI Publications.

Müller, Stefan. 2006. Phrasal or lexical constructions? *Language* 82(4). 850-883.

Nolda, Andreas. 2007. *Die Thema-Integration. Syntax und Semantik der 'gespaltenen Topikalisierung' im Deutschen* (Studien zur deutschen Grammatik 72). Tübingen: Stauffenburg Verlag.

Nolda, Andreas. 2012a. Definition and classification of compound types. A semantic approach. (Conference handout at "Universals and Typology in Word-Formation II", Košice, August 27, 2012).

Nolda, Andreas. 2012b. *Konversion im Deutschen: Muster und Beschränkungen. Mit einem Grundriß einer allgemeinen Theorie der Wortbildung.* Thesis, submitted at the Humboldt Universität zu Berlin in partial fulfilment of the Habilitation requirements, December 2012. (Ms). [Revised version to appear.]

Plag, Ingo & Gero Kunter. 2010. Constituent family size and compound stress assignment in English. In Susan Olsen (ed.), *New impulses in word-formation* (Linguistische Berichte, Sonderheft 17), 349-382. Hamburg: Buske.

Plank, Frans. 2010. Variable direction in zero-derivation and the unity of polysemous lexical items. *Word Structure* 3(1). 82-97.

Rainer, Franz. 2012. Morphological metaphysics: virtual, potential, and actual words. *Word Structure* 5(2). 165-182.

Rich, Georg A. 2003. *Partikelverben in der deutschen Gegenwartssprache mit* durch-, über-, unter-, ab-, an-. Frankfurt am Main etc.: Peter Lang.

Richter, Heide. 1988. *Indexikalität. Ihre Behandlung in Philosophie und Sprachwissenschaft* (Linguistische Arbeiten 217). Tübingen: Max Niemeyer Verlag.

Roach, Peter, James Hartman & Jane Setter (eds.). 2003. *Daniel Jones: English pronouncing dictionary*. 16th edn. Cambridge: Cambridge University Press.

Sackmann, Robin. 2006. Integrational Linguistics. In Keith Brown (ed.-in-chief), *Encyclopedia of language & linguistics*. 14 vols. 2nd edn, 704-713. Oxford: Elsevier.

Sackmann, Robin. 2008. An introduction to Integrational Linguistics. In Robin Sackmann (ed.), *Explorations in Integrational Linguistics. Four essays on German, French, and Guaraní* (Current Issues in Linguistic Theory 285, Studies in Integrational Linguistics 1), 1-20. Amsterdam/Philadelphia: Benjamins.

Scalise, Sergio & Antonietta Bisetto. 2009. The classification of compounds. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 34-53. Oxford: Oxford University Press.

Scalise, Sergio & Antonio Fábregas. 2010. The head in compounding. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding* (Current Issues in Linguistic Theory 311), 109-125. Amsterdam/Philadelphia: Benjamins.

Schlücking, Barbara & Matthias Hüning. 2009. Introduction. *Word Structure* 2(2). 149-154.

Spencer, Andrew. 2012. Identifying stems. *Word Structure* 5(1). 88-108.

Štekauer, Pavol. 2005. Onomasiological approach to word-formation. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation* (Studies in Natural Language and Linguistic Theory 64), 207-232. Dordrecht: Springer.

Štekauer, Pavol, Salvador Valera & Lívia Körtvélyessy. 2012. *Word-formation in the world's languages: a typological survey*. Cambridge: Cambridge University Press.

Stokhof, Martin & Michiel van Lambalgen. 2011. Abstractions and idealisations: The construction of modern Linguistics. *Theoretical Linguistics* 37 (1/2). 1-26.

Stump, Gregory T. 2001. *Inflectional morphology. A theory of paradigm structure* (Cambridge Studies in Linguistics 93). Cambridge: Cambridge University Press.

Stump, Gregory T. 2002. Morphological and syntactic paradigms: Arguments for a theory of paradigm linkage. *Yearbook of Morphology* 2001. 147-180.

Stump, Gregory T. 2005. Word-formation and inflectional morphology. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation* (Studies in Natural Language and Linguistic Theory 64), 49-71. Dordrecht: Springer.

Stump, Gregory T. 2010. The derivation of compound ordinal numerals: Implications for morphological theory. *Word Structure* 3(2). 205-233.

Viguier, Marie-Hélène. 2013. *Tempussemantik: Das französische Tempussystem. Eine integrative Analyse* (Beihefte zur Zeitschrift für romanische Philologie 366). Berlin: de Gruyter.

Zamponi, Raoul. 2009. Arawakan: Maipure-Yavitero. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 584-593. Oxford: Oxford University Press.