Chapter 19

Word formation in LFG-based layered morphology and two-level semantics

Christoph Schwarze

This article treats the problem of how the semantics of word formation can be accounted for in terms of rules and representations. A comprehensive model of multilayered, LFG-based morphology is proposed. It comprises four layers of representation: phonology, constituent structure, functional feature structure and lexical semantics. The meaning of derived words is treated in the framework of two-level semantics. It is assumed that rules of word formation derive underspecified semantic forms, parting from which the actual meanings are construed by recourse to conceptual structure. The model is illustrated on the basis of three morphological processes: French é-prefixation, Italian denominal verbs of removal, and noun-toverb conversion in French. The analyses of é-prefixation and of verbs of removal are taken from the literature; the study on noun-to-verb conversion is original work.

1 Introduction

The hypothesis that the semantics of word formation is an aspect of grammar assumes that the processes of word formation concern both form and meaning. However, actual work on this basis encounters considerable challenges. The data available for the study of a given process of word formation never seem to show a perfect parallelism between form and meaning: forms that stem from a given generative process often have meanings on which it seems to be impossible to form a descriptive generalization. It is the aim of this paper to show how challenges to the semantics of word formation can be dealt with.

I will first address the question of how morphological processes and structures can comprehensively be represented. I will then present three hypotheses concerning the semantics of word formation, namely

- i. The semantic output of morphological rules is underspecified.
- ii. The meanings of the words that constitute the data arise from a sequence of steps.



Christoph Schwarze. Word formation in LFG-based layered morphology and twolevel semantics. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer (eds.), *The lexeme in descriptive and theoretical morphology*, 487– 508. Berlin: Language Science Press. DOI:10.5281/zenodo.1407023 iii. These steps are a. the morphological rule defines an underspecified semantic form, b. semantic form is turned into a specified meaning on the basis of conceptual knowledge, c. the derived word enters the lexicon, and d. the lexicalized word may have its own development, independently from morphology, and its original meaning may thus be changed and its morphological origin be obscured.

2 LFG-based Layered Morphology

LFG-based Layered Morphology (LLM) integrates essential properties of Construction Grammar Morphology,¹ which does justice to the multi-layered nature of the lexicon, and HPSG-based morphology, which elaborates on the features that syntax receives from morphology.²

Notice that LLM is a model, not a theory or a hypothesis. Unlike theories and hypotheses, which can be empirically evaluated with reference to observable data, models can only be evaluated with respect to their usefulness for the progression of knowledge. This kind of usefulness cannot be measured, it can only be shown by actual work on specific phenomena. That is what I will try to do in this study.

Lexicalist models of grammar commonly assume that words are linguistic objects with layered representations, phonological, syntactic and semantic. Accordingly, morphological processes operate simultaneously at various layers or levels of representation.³ In accordance with Lexical Functional Grammar (LFG) the LLM model makes a distinction between the level of constituents, called the c-structure level, and a level of functional features, called the f-structure level.⁴ The latter contains features concerning agreement, tense, mood, inflectional class etc. It also contains grammatical functions and, importantly, predicate features, which are labels of lexical meanings and encode grammatical functions, the syntactic reflex of argument structure.

In addition to these two "syntactic" levels, morphological representations need to comprise a phonological level to account for non-concatenative morphological processes, like German Umlaut; cf. Germ. krank /krank/ 'sick' + $-lich /lx\chi/$ 'ly' \rightarrow kränklich /krɛnklı χ / 'sickly'.

And, of course, there is a semantic level, where the lexical meanings encoded in the lexicon are represented and processed. Resuming, morphological representations and processes are located at

• The level of constituent structure (the c-structure level)

¹See Booij (2010), Booij & Audring (2017).

²For work on French, see Fradin (2005) and Tribout (2010b).

³I fully agree with Aurnague & Plénat (2008: 1)when they say: "Une lexie est un n-tuplet de représentations reliées entre elles, mais relevant chacune d'un niveau linguistique (phonologique, syntaxique, sémantique, etc.) distinct. La description d'un mode de formation lexical productif suppose par conséquent que soient relevées et expliquées les régularités apparaissant à chacun de ces niveaux."

⁴LLM was first presented in a seminar held by the author at the University of Padova in 2008 and subsequently applied to the formation of Italian past and passive participles in Schwarze (2011).

- The functional level (the f-structure level)
- The phonological level (the p-structure level)
- The semantic level (the s-structure level)

Unlike syntax, morphology may manipulate predicates, thus deriving new predicates.

3 A sample analysis: French é-prefixation

I will illustrate LLM on the basis of Namer & Jacquey's $(2003)^5$ article on the French éprefixation, which endevours to give a formalized version of the findings of Aurnague & Plénat (1997). In one of its modalities, é-prefixation turns nouns into transitive verbs that denote events where the referent of the base noun is distanced, removed, or separated from the referent of the direct object, as in (1):⁶

```
(1) FR.
```

a.	\acute{e} + branche \rightarrow \acute{e} brancher (un arbre)
	'branch' 'to prune a tree'
b.	\acute{e} + feuille \rightarrow effeuiller (x)
	'leaf' 'to strip the leaves or petals from x'
c.	$\acute{e} + gorge \rightarrow \acute{e}gorger(x)$
	'throat' 'to cut x's throat'
d.	$\acute{e} + pou \longrightarrow \acute{e}pouiller(x)$
	'louse' 'to delouse x'

Moreover, as has been shown by Aurnague & Plénat (1997, 2007, 2008), the relation that holds between the two dissociated entities must be "usual" and "natural",⁷ or, as Namer & Jacquey put it:

[D]escribing the process consisting in clearing a tree of e.g. the magpies (*pie*) or the cats (*chat*) that colonize it cannot be performed by processes referred to by the ?épier⁸ or ?échatter impossible derived verbs. (Namer & Jacquey 2003: 2)

Table 1 gives the rule that generates verbs like *ébrancher*, *effeuiller*, *égorger* or *épouiller* in the LLM notation.

The c-structure change as formulated in Table 1 should be self-explanatory, whereas a few comments on the f- and s-structure part of the rule will be useful.

⁵In a subsequent article, Namer & Jacquey (2012) proposed a modelization of the N>V vs. V>N derivations within the framework of the Generative Lexicon.

⁶Changes like adding /j/ to the stem as in *épouiller* are idiosyncratic and must be accounted for in the lexicon. ⁷"[...] les dérivés en *é*- expriment la dissociation [...] par un agent intentionnel [...] d'une relation

d'attachement habituel [...] créée naturellement [...] et à laquelle il s'oppose [...]" (Aurnague & Plénat 2008: 28).

⁸Not to be confounded with existing *épier qu*. 'to spy on someone'.

p-structure <no change="" morphologically="" relevant=""> s-structure $\lambda y p(y) \rightarrow \lambda x \lambda y \exists z$</no>	f-structure p-structure	1 0 , 0
--	----------------------------	---------

PRED is a feature attribute, whose value identifies a word's lexical meaning and argument structure. The input, (\uparrow PRED1)='P', contains a predicate variable, *p*, which ranges over the nominal predicates associated with constituent N_{stem}. The up-arrow is an abbreviation for a function that projects the feature to the dominant c-structure node. The output of the semantic change is a new predicate PRED2, which is defined by the rule. It has two arguments, an agent and a theme, realized as the subject and the object respectively. Notice that the prefix, in accordance with Namer & Jacquey (2003), has no direct functional representation, because it has no referential meaning.⁹ As to the s-structure level, the derived predicate, 'dissociate', has three semantic arguments: *x*, which is the subject and refers to the agent, *y*, which is the object and refers to the theme, and *z*, which is incorporated in the verb's meaning and refers to the entity which is dissociated from *y*. The additional predication, repeated as (2), is needed to constrain the range of *y* and *z*:

(2) natural_relationship(y, z)

This part of the representation expresses the fact that the relation between y and z must not be a merely accidental one, as reported above.

Notice that the change in s-structure as expressed in Table 1 does not predict the full actual meanings of the verbs derived by *é*-prefixation: the derived representation is underspecified.¹⁰ In the following section I will give some background for such an assumption.

⁹"Our purpose is ... to represent the verb class obtained by the *é*-prefix derivation. To achieve this, two basic ways are provided: (i) representing the prefix itself or (ii) representing an abstract, parametrized lexical unit describing the output (verbal) class. The motivation for the first choice would be the fact that the affix can be seen as some kind of predicate, operating on and controlling two arguments, the base and the derived word, from a structural, categorial and above all semantic points of view. However, the nature itself of the affix is a counterargument: according to the morphological theory defended here, an affix does not belong to any of the major categories. In addition, we have seen that it bears no referential meaning: consequently, it is not foreseeable to modelize its semantic content, as it has no proper semantic content." (Namer & Jacquey 2003). I follow this argumentation, with the exception that not belonging to a major category does not generally imply the lack of functional or semantic information.

¹⁰This assumption is quite common in the literature, see the survey in Tribout (2010b: 282-284).

4 Two-level semantics

If we assume that lexical morphology is a generative subsystem that feeds the lexicon, then its semantics is part of lexical semantics. Now, the fundamental question is to which extent lexical semantics is an affair of grammar. According to the conception known as two-level-semantics, lexical meaning is represented at two distinct levels, semantic form (SF) and conceptual structure (cs).¹¹

Semantic form is linguistic knowledge. sFs "are systematically connected to, and hence covered by, lexical items and their combinatorial potential to form more complex expressions" (Lang & Maienborn 2011: 711). They "form an integral part of the information cluster represented by the lexical entries of a given language" (Lang & Maienborn 2011: 711). They are "accessibly stored in long-term memory" (ib.). They are underspecified with respect to cs representations (Lang & Maienborn 2011: 713). And, importantly, sF is the level at which two-level semantics endeavors to represent the compositionality of lexical meaning and the grammatical role of lexical decomposition (Lang & Maienborn 2011: 723).

Returning to the semantics of word formation, it is an aspect of grammar, as far as semantic form is concerned. Most of the characteristics of sF that hold for ordinary lexical semantics also hold for the semantics of word formation, with one exception: compositionality is not a general feature of lexical morphology. In fact, non-concatenative processes may be absolutely regular, but cannot be compositional, since compositionality presupposes concatenation.

In order to see whether the semantics of lexical morphology can reach out to phenomena that are situated beyond SF, let us see what two-level semantics means by conceptual structure.

Conceptual structure can be said to be world knowledge (Lang & Maienborn 2011: 711). That does not mean, however, that it has nothing to do with language, actually, it is closely related to SF: CS representations are built upon and enrich SF representations. Thus, semantic representations typically contain both, CS and SF features. This happens in such a way that, for the representation of a given lexical meaning, the CS features specify and enrich SF representations, thus enabling words to denote their referents.^{12,13,14}

¹¹For a critical state-of-the-art overview, see Lang & Maienborn (2011).

¹²In Lang and Maienborn's words: "...for every linguistic expression *e* in language *L* there is a CS representation *c* assignable to it via sF(e), but not vice versa" (Lang & Maienborn 2011: 711); "... CS representations are taken to belong to, or at least to be rooted in, the non-linguistic mental systems based on which linguistic expressions are interpreted and related to their denotations."

¹³This conception has an important consequence: if the features retrieved from cs are combined with or replace sF features, doing lexical semantics does not mean to represent the entire bulk of knowledge and beliefs that we have about the referents of the lexemes under investigation.

¹⁴As to the mental status and processing of CS representations, they are assumed to be "activated and compiled in working memory", contrarily to SF representations, which, as has been said above, are stored in long-term memory (Lang & Maienborn 2011: 712). I am not sure about the mental status of CS: it may safely be assumed that concepts, once they are lexicalized as meaning components, are as stable as SFS.

5 A second sample analysis: Italian denominal verbs of removal

It will be useful to illustrate underspecification and its resolution with an example from derivational morphology. I will briefly present the analysis of the denominal verbs derived by *s*- prefixation in Italian as proposed by von Heusinger & Schwarze (2006).¹⁵

The morphological process generates verb stems from noun stems by prefixing the constituent s- to the noun stem;¹⁶ cf. (3) and (4):

- (3) a. crem(a)_N \rightarrow screma(re)_V 'cream' 'to skim'
 - b. $carcer(e)_N \rightarrow scarcera(re)_V$ 'prison' 'to release from prison'
- (4) a. La mattina, la nonna scremava il latte. the morning the grandmother skimmed_{IMPERFECT} the milk
 'In the morning, Grandmother used to skim the milk'
 - b. Il giudice ha scarcerato Giovanni Rossi.
 the judge has released-from-prison Giovanni Rossi
 'The judge released Giovanni Rossi from prison.'

Both verbs, *scremare* and *scarcerare*, mean 'x removes y from z'. However, they differ with respect to the role of the nominal base in the verbs' meaning. In terms of Leonard Talmy's (1985) lexical typology of motion events, the entity denoted by the base noun may be the Figure or the Ground. In (4a) the cream (*crema*) is the Figure; it is removed from the milk (*latte*), which is the Ground. Inversely, in (3b) the prison (*carcere*) is the Ground, from which Giovanni Rossi, the Figure, is released. Thus the speaker needs to decide on the assignment of Figure and Ground for every single verb generated by $N \rightarrow V$ *s*-prefixation. In a two-level semantics, sF will only state that the verbs under discussion denote caused motion, the role of the incorporated noun being left open. The general semantic form of these verbs may thus be written as (5):¹⁷

(5) $cause(x, become(\neg located(y, z))) \& [N(y) \lor N(z)]$

The first part of representation (5), $cause(x, become(\neglocated(y, z)))$, is the lexical decomposition of the main feature of all verbs of removal, remove(x, y, z). The second part,

¹⁵Giuseppina Todaro (2017) applies the von Heusinger & Schwarze (2006) approach to prefixed deadjectival verbs in Italian.

¹⁶Notice that Italian also has a V \rightarrow V *s*-prefixation, which derives verbs of reversal, see Mayo et al. (1995: 932), among others. This is a different morphological process, which I do not discuss here.

¹⁷In von Heusinger & Schwarze (2006) the representation given here as (5) is not the final version, which uses indices in order to account for the correlation between ambiguity of role assignment and the alternative of quantification. In fact, if the predicate of the base noun is incorporated in the verb, it is only existentially bound by \exists . If it becomes the direct object, it is bound by the λ operator. In (5), quantification is omitted for the sake of easier reading.

 $N(y) \lor N(z)$, expresses the underspecification of *s*-prefixed verbs of removal by a disjunction, where N is the predicate of the base noun.

The ambiguity expressed by this disjunction is resolved at the cs level. According to von Heusinger & Schwarze's (2006) analysis, the resolution of the underspecification passes through the following phases: the concepts associated with the base noun predicates are looked up in cs and checked regarding their aptitude to be a Figure or a Ground in a motion event. Objects that may contain something, are apt to take the role of Ground, objects that may easily perform or undergo motion are apt to be the Figure. Some objects, such as a sheet of paper, may meet both criteria and may consequently motivate derived verbs with two alternative fully specified meanings. Italian *scartare*, derived from *carta* 'paper', is such a case: it may be used as both a Ground verb or a Figure verb; cf. (6):

(6)	Mario scarta	il	regalo.		
	Mario s-paper _{3sg.ind.pres}	, the	e gift		
	'Mario takes the gift ou	t of	the paper'	C	Ground verb
	'Mario takes the paper	off	the gift'		Figure verb

Table 2 gives the rule that derives denominal *s*-prefixed verbs of removal in the LLM format, with the semantic layer formulated in such a way as to generate underspecified sF representations.^{18,19}

Table 2: The rule for deriving Italian denominal s-prefixed verbs

c-structure	$[s]_{V_{prefix}} + N_{stem} \rightarrow [sN_{stem}]_{V_{stem}}$
f-structure	(\[\frac{Pred1}{Pred1}]='P' \rightarrow (\[\frac{Pred2}{Pred2}]='remove (\[\frac{Subj}{Subj}), (\[\frac{T}{OBJ}), (\[\frac{T}{OBJ}))'
-	<no change="" morphologically="" relevant=""> $\pi(y) \rightarrow \text{cause}(x, \text{become}(\neg \text{located}(y, z))) \& [N(y) \lor N(z)]$</no>

6 A third sample analysis: French $N \rightarrow V$ conversion

I will now present a case study of French $N \rightarrow V$ conversion, as exemplified by the pairs in (7):

- (7) a. amidon amidonner 'starch' 'to starch'
 - b. archives archiver 'archives' 'to archive'

¹⁸For easier reading, I do not express here the case-marking of the Oblique, which must be *ne* if its predicate is 'PRO' and must be marked by preposition *da* elsewhere.

¹⁹The [s] vs. [z] realization of the prefix is a matter of post-lexical phonology, hence it is not expressed in the morphological rule.

c. bêche – bêcher 'spade' 'to dig'
d. sucre – sucrer 'sugar' 'to sugar'

The relation between the nouns and the respective verbs in (7) is clearly directed, which does not hold for other noun-verb pairs, as those given in (8):

- (8) a. chant chanter 'song' 'to sing'
 - b. gel geler
 'frost' 'to freeze'
 - c. prêt prêter 'loan' 'to lend'
 - d. vent venter'wind' 'to be windy'

The difference between (7) and (8) is due to the ontological class of the nouns' predicates: whereas the nouns in (7) denote objects or substances and thus are clearly distinct from the respective verbs, those given in (8) denote events or results of events and thus are not clearly distinct from the verbs they relate to. The derivational direction in (7) clearly is N→V, because event predicates may be built upon object or substance predicates, but not inversely.²⁰ On the contrary, the conversion in (8) may be the opposite, $V\rightarrow N$,²¹ or non-directional, N↔V, because the nouns' and the verbs' predicates are identical or very closely related.

As for the semantics of $N \rightarrow V$ conversion, I assume that the rule defines an underspecified semantic form, from which full meanings are derived by a retrieval of conceptual structure.²² To account for actual meanings that are not predicted on this basis, postmorphological processes are taken into account. It is also assumed that there are certain verbs that look like $N \rightarrow V$ converts, but are idiosyncratic items not derived by the rule.

²⁰Cf. the more explicit formulation by Tribout (2010b: 140): "... le recours aux propriétés sémantiques des deux lexèmes pour déterminer l'orientation de la conversion repose, par exemple, sur l'idée que le lexème dérivé est nécessairement défini par le biais de son lexème base, tandis que le lexème base est sémantiquement indépendant de son lexème dérivé. Ainsi pour la paire CLOU~CLOUER, CLOUER est nécessairement défini relativement à CLOU comme 'faire quelque chose avec des clous' tandis que CLOU est défini comme un petit objet pointu, indépendamment de clouer. Cette asymétrie dans la relation sémantique entre les deux lexèmes permet de prédire une orientation de la conversion de nom à verbe."

²¹For a state-of-the-art discussion on the direction of the French N \rightarrow V vs. V \rightarrow N conversion see Tribout (2010a: 348–356).

²²Tribout (2010b: 284–290) criticizes the underspecification approach; instead she proposes and spells out a fully specified semantics, based upon a classification of the output verbs. I am trying to show that an underspecification-based analysis of the French N→V conversion is an achievable goal.

6.1 A database

As a descriptive basis for the study, I established a database of 170 verbs that clearly are $N \rightarrow V$ converts. 19 of these verbs are prefixed and have no lexicalized unprefixed counterpart, such as *emprisonner* 'to imprison'.

I consider including prefixed verbs of this kind as legitimate, because the prefixes involved, *en-*, *dé-* and *re-*, require a verbal base. *Emprisonner*, e.g., thus has the derivational history shown by (9):

(9) $prison_N \rightarrow prisonner_V \rightarrow emprisonner_V$

In addition to the verbs and their base nouns, the database contains the following information:

- The verb's underspecified semantic form (SF), if there is one
- The specified semantic representation (SR)
- · The conceptual class of the base noun
- The presence of a prefix, if there is one
- · Remarks on formal and semantic properties of the derived verbs

6.2 The underspecified semantic forms

Underspecified semantic forms could be construed for 142 of the 170 verbs. The predominant one, which holds for 136 of the 170 verbs contained in the database, states the following:

- The verb denotes an event, which is an action
- · It involves an agent and a theme
- The denotation of the noun from which the verb is derived is a salient component of the action

For an illustration, see example (10):

(10) Le secrétaire a archivé la correspondence. the secretary has archived the correspondence 'The secretary archived the correspondence.'

The sF underlying (10) states that the sentence describes an action. The denotation of the noun *archives* is a salient component of that action. The verb, *archivé*, has two arguments, *le secrétaire* and *la correspondence*, whose roles are agent and theme respectively.

In addition to the predominant sF, two more sFs have been identified; they are closely related to the predominant one, see examples (11) and (12). (11) describes an action, but

unlike (10), the verb has no argument in the role of theme. (12), where the reflexive pronoun is the operator of the middle voice, describes a process, the verb's only argument is in the role of theme.

- (11) Nous avons passé l'après-midi à magasiner. (Canadian Fr.) we have spent the afternoon to window-shopping
 'We spent the afternoon window-shopping.'
- (12) Leurs genoux se sont ankylosés. their knees REFLEXIVE_PRONOUN are ankylosed 'Their knees have become stiff.'

All sFs assumed for the verbs contained in the database are shown in Table 3, which also shows the forms of the semantic predicates involved, the mapping of the arguments onto grammatical functions and the number of verbs for each sF.²³

	Semantic predicate	Grammatical functions	SF	verbs
SF1	P(e, x, y) agent(x) theme(y)	Р (subj), (оbj) х у	x accomplishes an action on y, N is salient in that action.	136
sf2	P(e, x) agent (x)	Р (subj) 	x accomplishes an action. N is salient in that action.	4
sf3	P(e, x, y) theme(x)	x P (subj) x	x undergoes a process. N is salient in that process.	4

Table 3: Underspecified semantic forms of converted denominal verbs

We can now formulate the rule for French $N \rightarrow V$ conversion, see Table 4.²⁴ At the semantic layer, only the predominant sF is given.

²³There are two questions that I cannot address here in detail. First, how productive is the process analyzed here? French is a language that overwhelmingly prefers affixation to conversion. I assume that N→V is fully productive, but that much of its output is blocked by the output of competing rules of affixation. Second, can the non-dominant sFs be derived from the predominent one? Further research is needed here.

²⁴Except the selection of alternative lexicalized stem variants, see fn. 8. In the table I omit quantification again in order to make reading easier.

c-structure f-structure	$N_{\text{stem}} \rightarrow V_{\text{stem}}$, 1 st inflectional class (\uparrow PRED)='P1' \rightarrow (\uparrow PRED)='P2 (\uparrow SUBJ),(\uparrow OBJ)'				
p-structure	<no change="" morphologically="" relevant=""></no>				
s-structure	$p1(z) \rightarrow p2(e, x, y) \wedge agent(x) \wedge theme(y) \wedge$				
	salient_component_of(e) = p1(y)				

Table 4: The layered rule for $N \rightarrow V$ conversion

7 Resolving the underspecified semantic forms

As has already been pointed out, the underspecified semantic forms cannot be used in discourse, because they are unable to refer to the specific actions denoted by the verbs. Hence the underspecification needs to be resolved. This happens by accessing the conceptual knowledge associated to the base nouns. Regarding $N \rightarrow V$ conversion, I assume that the speaker or hearer looks up the conceptual knowledge associated with the noun, inspects the event types in which the noun's denotation is typically involved, and finally creates a new semantic predicate in which one of these event types is, so to speak, incorporated. The noun's meaning is then turned into a feature of the new predicate, a feature that becomes visible by lexical decomposition. I will try to illustrate this idea by means of two examples, the first is (13):

(13) L' orfèvre a ciselé leurs noms sur les alliances. the goldsmith has chiseled their names on the wedding_rings 'The goldsmith engraved their names on the wedding rings.'

The verb contained in (13) has the general, underspecified semantic form listed as sF1 in Table 3, and repeated here as (14):

(14) X accomplishes an action on y; N is salient in that action.

For ciseler 'to chisel, to engrave' we replace N with "a chisel", getting (15):

(15) X accomplishes an action on y, a chisel (Fr. *ciseau*) is salient in that action.

The conceptual knowledge associated with *ciseau* contains, among others, the information given under (16):

(16) A chisel is a tool, used for cutting wood, stone or metals.

The predicate cut(x, y) is the semantic counterpart of the concept of cutting. Going back from conceptual structure to semantic form, the speaker inserts it into the decomposed semantic representation of the new predicate created by the conversion rule. The meaning of the new predicate also contains chisel(x), taken from the base noun. Since, according to (16), a chisel is a tool, i.e. an instrument, the feature will be $instrument_used(e, x, y)=chisel(z)$. Notice that z is not an argument of the new predicate and will not be realized in the sentence. (17) is the assumed semantic representation of *ciseler*, after the resolution of underspecification.

(17) $\exists e \ \lambda x \ \lambda y \ chisel(e, x, y)^{25}$ $event_type(e) = action(e)$ $action_type(e) = cut(x, y)$ agent(e) = xtheme(e) = y instrument_used(e) = chisel(z)

The first line of (17) gives the semantic representation of *ciseler* in the standard notation. The remaining lines give its decomposed meaning in terms of features, written as equations, in the tradition of unification grammars. (This notation mainly shows its usefulness when larger sections of the lexicon are analyzed: it makes it easy to express feature inheritance, and it helps to control the consistency of the features declared.)

The second example I give for the resolution of underspecification is (18):

 (18) Les chasseurs ont huilé leurs fusils. the hunters have oiled their shotguns
 'The hunters oiled their shotguns.'

The verb *huiler* 'to oil' has the same sF as *ciseler*. Applied to the base noun *huile* 'oil' it reads:

(19) X accomplishes an action on y, oil (Fr. huile) is salient in that action.

Accessing the conceptual knowledge associated with *huile*, the speaker gets, among others, the information given under (20):

(20) Oil is a substance used to lubricate a mechanism.

The predicate lubricate(x, y) is the semantic counterpart of the concept of lubricating. The speaker inserts it into the decomposed semantic representation; the meaning of the new predicate also contains oil(x), taken from the base noun. Since, according to (20), oil is a substance, the feature will be $substance_used(e, x, y) = oil(z)$. (21) is the assumed semantic representation of *huiler*:

(21) $\exists e \ \lambda x \ \lambda y \ oil(e, x, y)$ $event_type(e) = action(e)$ $action_type(e) = lubricate(x, y)$ agent(e) = xtheme(e) = y substance used(e) = oil(z)

²⁵For readers not familiar with the French language, I use English to name semantic features, even though this may make the analysis somewhat inaccurate.

7.1 Polysemy in lexical morphology

The conceptual categorization of 'oil' I assumed for the above sample analysis, i.e. that 'oil' is a substance used to lubricate a mechanism, is far from being the only one.²⁶ As we know, oil also is used to preserve wood or iron, to cook and season food, it also is a fuel, and an ingredient of oil paint. As linguists, we do not have scientific methods to find out to what extent knowledge of this kind is contained in the conceptual structure and we have no precise knowledge of how conceptual structure is processed during the resolution of semantic underspecification. However, we can look at the lexicon and see those elements of conceptual structure that show up in the lexical meanings of a given language. Thus we can observe that, in the meaning variation of the French verb *huiler* 'to oil' the following bits of information clearly play a role:

- i. Oil is a lubricant (22), repeated from (18).
- ii. Oil is a preservative (23).
- (22) Les chasseurs ont huilé leurs fusils. the hunters have oiled their shotguns 'The hunters oiled their shotguns.'
- (23) Cette table a besoin d' être huilée. this table has need to be oiled 'This table needs to be oiled.'

As to using oil for preparing or seasoning food, the situation is less clear. According to the reviewer of this article, whom I believe to be a native speaker of French, *huiler* cannot mean 'to season with oil'. I briefly searched the Internet and found out that there were zero hits for *huiler la viande* (*viande* means 'meat') and *huiler les steaks*. There were several hits for *huiler la salade*, but only two of them were from real text (24) and (25), the others being citations from dictionaries.

(24) J'aime faire des vinaigrettes qui ne font pas qu'assaisonner ou huiler la salade mais qui apportent plutôt une valeur ajoutée.²⁷

'I like to make vinaigrettes that do not only season or oil the salad but rather bring an additional value.'

(25) Ne pas huiler la salade, car ainsi suivant son goût chacun fera sa propre vinaigrette, et puis s'il reste de la salade, elle se conservera plus facilement sans vinaigrette.²⁸

'Don't oil the salad, because that's how everyone will make their own vinaigrette to their taste, and then, if some salad is left over, it will be preserved more easily without vinaigrette.'

²⁶I inserted this section as a response to a comment I received from an anonymous reviewer. For the analysis of polysemy in lexical morphology, also see Schwarze (2012).

²⁷ http://brutalimentation.ca/2017/01/14/salade-festive-vinaigrette-digestive [2017-08-29].

²⁸http://ilovecuisine.blogspot.ch/2013/09/ma-salade-de-lete-la-salade-nicoise.html [2017-08-29].

The remaining known uses of oil do not seem to play a role in the meaning variation of French *huiler*. Instead of speculating about why this should be so, let us pass on to a question that immediately arises from what we could observe.

Assuming that the accessible conceptual structure offers competing information for the resolution of the underspecified meaning generated by the morphological process, the full meaning of *huiler* shows the following variants:

- (26) a. 'To lubricate with oil'
 - b. 'To preserve with oil'
 - c. 'To prepare or season with oil'

The question now is: How do speakers pick out the convenient reading in producing or parsing utterances? This is a very general question, not specific to the semantics of word formation. In the case of transitive verbs such as *huiler*, a sort of semantic agreement is at work, which checks the compatibility of the verb's reading with the conceptual class of the direct object.

Regarding the avoidance of *huiler* with a direct object denoting meat, there may be practical reasons or no reason at all; there are phenomena in verbal behavior that are beyond the reach of linguistic analysis.

8 Restrictions on the input

It can easily be seen that many nouns are not fit to be a base in the French $N \rightarrow V$ conversion. In a list of the first 100 non-eventual nouns contained in the *Petit Larousse*, only two are a base of $N \rightarrow V$ converts, and only one of them, *acier* 'steel', is the stem of a verb with a transparent meaning, *aciérer* 'to cover with steel'.²⁹ Notice, however, that this finding rests on a very weak empirical basis. The nouns considered are very few, and the data are limited to strongly lexicalized items. More research is needed to get reliable quantitative results. So I will just characterize the database with respect to the 143 nouns that are the base of verbs with a transparent meaning. Turning these observations into well-founded constraints and disentangling grammatical constraints on the input and conditions for use and lexicalization of the output must be left to further research.

The following semantic characteristics of the base nouns can be gathered from the database:

• Most base nouns denote an instrument (42 items),³⁰ a substance (36 items), a container (seven items), or a body part (nine items); see Tables 8 to 11 in the Appendix.

²⁹The other, *abîme* 'abyss', has *abîmer* 'to damage' as a convert, but that verb has a meaning that does not seem to be derived in a straightforward way from the noun's meaning.

³⁰Cf. "Les verbes converts instrumentaux sont parmi les plus nombreux. Ils sont mentionnés dans toutes les études portant sur la conversion et sont généralement définis comme signifiant 'utiliser N', selon le schéma ... X utiliser Nb" (Tribout 2010b: 263).

- Only one noun, *enfant* 'child', denotes a human being. The derived verb, *enfanter* 'to give birth to', is infrequent and strongly marked as belonging to the literary register.
- Only two nouns denote an animal, *raton* 'young rat', and *zèbre* 'zebra'. *Ratonner* 'to commit a racial attack (*ratonnade*) on North-African immigrants' has no transparent semantic relationship to its base. As to *zèbre*, the derived verb, *zébrer* 'to stripe', is only weakly transparent: rather than to the animal, it refers to a visual pattern, black stripes upon a white ground.

Regarding the formal properties of the base nouns, short words are preferred: most of them are mono- or disyllabic, only three (*ankylose* 'ankylosis', *courbature* 'ache, stiffness', and *magasin* 'store') have three and only one (*photographie* 'photography') has four syllables.

Nouns consisting of one morpheme only are clearly preferred; only *tambourin* 'tambourine' and *photographie* 'photography' may be segmented into morphemes. There are no agent nouns in -(at)eur and no quality nouns in -(i)te in the stems of derived verbs.

9 Reduced or lacking transparency – construed lexemes in time

The database contains several verbs whose relationship with the base noun is not fully transparent or not transparent at all. For none less than 25 of the 170 verbs, no underspecified semantic form could be identified, which means that the meaning of the base noun is not a feature of the derived verb, see the examples in (27):

(27)	a.	fourrager	fourrage
		'to rummage through'	'forage'
	b.	fronder	fronde
		'to satirize'	'slingshot; revolt'
	c.	gueuler	gueule
		'to yell, to bawl'	'mouth'

Ten verbs can be analyzed as having undergone some post-morphological change along one of the familiar paths of semantic change or variation, such as narrowing or widening an original meaning. Examples are shown in Table 5:

A particular kind of incomplete semantic transparency of the converted verb is due the fact that, rather than the verb, the base noun underwent a change after the derived verb entered the mental lexicon. Examples are *échafauder* 'to put up scaffolding' and *mitrailler* 'to machine-gun'. The base noun of *échafauder*, *échafaud*, does not mean 'scaffolding' any longer, it means 'executioner's platform' in modern-day French. The verb's meaning came about when *échafaud* still meant 'scaffolding'. Likewise, *mitrailler* 'to

Ν	English	V	English	kind of change
fer	'iron'	ferrer	 'to shoe (a horse)', 'to strike (a fish)' 'to do some gardening' 'to wall (a prisoner)' 'to shade, to hatch' 'to peel' 'to track down (a game)' 'to pluck (a bird)' 'to decorate (a wall and similar)' 	narrowing
jardin	'garden'	jardiner		narrowing
mur	'wall'	emmurer		narrowing
ombre	'shadow'	ombrer		narrowing
peau	'skin'	peler		narrowing
piste	'trace'	dépister		narrowing
plume	'feather'	plumer		narrowing
tapis	'carpet'	tapisser		widening

Table 5: Post-morphological change along some familiar paths

machine-gun' was created when the noun, *mitraille*, still meant 'machine gun'. Its meaning changed to 'hail of bullets', which lessened the semantic transparency of the derived verb.

The formal transparency may also be obscured, i.e. the noun's stem may differ to some extent from the derived verb's stem.³¹ The variation in such cases mostly is due to morphologization of a phonological variation existing at an earlier stage of the language and may be made less opaque by the existence, in modern French, of other examples that exhibit the same lexical variation. The variation between /o/ and /ɛl/ or /əl/ as in *peau* /po/ 'skin' – /pɛl/ 'peels' and *peler* /pəle/ 'to peel' is such a case. Its transparency is improved by the presence of numerous items like those given in (28):

- (28) a. nouveau /nuvo/ 'new.mas'
 - b. nouvelle /nuvɛl/ 'new.FEM'
 - c. renouveler /rənuvəle/

'to renew'

 d. niveau – niveler /nivo/ /nivəle/ 'level' 'to level'

³¹For a complete list of the kinds of allomorphy involved in N→V conversion see Tribout (2010b: 114f). She argues that even totally opaque pairs such as *pierre* 'stone' and *lapider* 'to stone' may be analyzed as cases of conversion, because they are related by suppletion (Tribout 2010b: 110, 118).

But this is not always the case. See the right-most column in Table 6.

Ν	English	V	English	remarks
ciseau /sizo/	'chisle'	ciseler /sizəle/	'to chisle'	see (25)
faux /fo/	'scythe'	<i>faucher</i> /fo∫e/	'to scythe'	isolated variation
grain /gьɛ̃/	ʻgrain'	engrener /ɑ̃gвəne/	'to engage with'	no transparency
hiver /ivɛʁ/	'winter'	hiverner /ivɛвne/	'to winter'	cf. jour -journée
<i>marteau</i> /maʁto/	'hammer'	<i>marteler</i> /maʁtəle/	'to hammer	see (26)
nœud /nø/	'knot'	nouer /nue/	'to knot'	cf. jeu-jouer
<i>poil</i> /pwal/	'hair'	<i>peler</i> /pəle/	'to peel'	cf. moi-me
<i>poil</i> /pwal/	'hair'	<i>épiler</i> /epile/	'to depilate'	native vs. borrowed
sang /sã/	blood'	saigner /sɛɲe/	'to bleed'	isolated variation

Table 6: Stem variation in N-V pairs

Most of these cases of reduced or lacking transparency have originated from the development of the grammar combined with the effects of lexicalization. $N \rightarrow V$ conversion has been a persistent rule in a changing grammar. It was present at the Latin stage of the language (see Table 7), and endured throughout the centuries up to the present day, while there happened important changes elsewhere in the grammar.

Table 7: $N \rightarrow V$	conversion in Latin
----------------------------	---------------------

N	English	V	English
cor cordis glacies navigium onus oneris pignus pignoris pilum	heart ice vessel, ship cargo, burden, load bet, stake, pledge hair	recordor glaciō navigō onerō pignorō pilō	to call to mind, to remember to freeze to navigate, to sail to load, to burden to pledge to depilate
pugnus sal vēlum	fist salt curtain, sail, covering	pugnō salō vēlō	to depliate to fight to salt to enfold, envelop, veil

When speakers found it useful for communication, the output of the rule entered into usage and was lexicalized. This happened at various periods, when the meaning of the base noun could be different from today's, and when there was a regular phonological variation given up later. But the original forms and meanings could remain in the lexicon.

Moreover, once a construed word has entered the mental lexicon, its meaning may develop freely, which leads to reduced or lacking transparency with respect to the original meaning, founded on some sF and its conceptual resolution. What does that mean for the morphological process as a part of mental grammar? Remember that word formation rules are thought to have a double purpose: they create possible words, and they analyze existent words. Hence the $N \rightarrow V$ conversion rule will not create opaque or semi-transparent forms. However, as a means of learning and understanding construed lexemes, it will also cope with semi-transparent forms, to the extent that suitable variation patterns are present in the lexicon. Thus speakers will presumably be able to relate *ciseler* /sizəle/ to *ciseau* /sizo/ or *marteler* /maʁtəle/ to *marteau* /maʁto/, because these pairs show a variation pattern that is also present elsewhere in the lexicon. In addition, a clear semantic relationship between the noun and the verb certainly is a strong support to transparency. It would be interesting to see experimental research on this point.

Acknowledgments

I am most grateful to Fabio Montermini, who thouroughly read the present text and gave valuable comments that helped me improve its form and content. I also am indebted to an anonymous reviewer, who discovered several remaining errors and made most constructive comments.

Appendix

The Appendix contains some tables that would have disturbed the reading process of the main text.

V	English	Ν	English
ciller	to blink	cil	eyelash
enculer	to sodomize	cul	arse
doigter	to use one's fingers correctly on a piano and similar	doigt	finger
griffer	to scratch	griffe	claw
gueuler	to yell	gueule	mouth
manier	to handle	main	hand
peler	to peel	реаи	skin
plumer	to pluck (a bird)	plume	feather
dépiler	to depilate	poil	hair
sourciller	to raise one's eyebrows	sourcil	eyebrow
talonner	to follow someone's heels	talon	heel
zyeuter	to take a look at	yeux	eyes

Table 8: Verbs derived from nouns that denote a body part

V	English	Ν	English
ancrer	to anchor	ancre	anchor
arquer	to curve	arc	bow
basculer	to topple over	bascule	seesaw
bêcher	to dig with a spade	bêche	spade
boulonner	to bolt	boulon	bolt
brosser	to brush	brosse	brush
ceinturer	to surround	ceinture	belt
ciseler	to chisle	ciseau	chisle
claironner	to shout from the rouftops	clairon	bugle
clouer	to nail	clou	nail
cravacher	to whip	cravache	whip
crocheter	to pick (a door, a lock)	crochet	picklock
chaîner	to put on snow chains	chaîne	chain
faucher	to scythe	faux	scythe
filtrer	to filter	filtre	filter
flinguer	to blow away, to shoot	flingue	gun
flûter	to produce a flute-like sound	flûte	flute
fouetter	to flog, to whip	fouet	whip
fourcher	to split	fourche	fork
freiner	to brake	frein	brake
fronder	to satirize	fronde	sling, revolt
fusiller	to shoot (a condemned person)	fusil	rifle
hacher	to chop	hache	ax
griller	to grill	gril	grill
limer	to file	lime	file
marteler	to beat, to pound	marteau	hammer
menotter	to handcuff	menottes	handcuffs
miner	to mine, to sap	mine	mine
mitrailler	to machine-gun	mitraille	hail of bullets
peigner	to comb	peigne	comb
photographier	to photograph	photographie	photography
pilonner	to bombard, to grind	pilon	pestle
poignarder	to stab	poignard	dagger
raboter	to plane	rabot	plane
sabrer	to cut down	sabre	sword
scier	to saw	scie	saw
tambouriner	to hammer, to drum	tambourin	tambourin
tamiser	to sieve	tamis	sieve
téléphoner	to phone	téléphone	phone
se tirebouchonner	to be twisted, to be wrinkled	tirebouchon	corkscrew
visser	to screw on	vis	screw
vriller	to bore, to pierce	vrille	spiral
vi illoi	to bole, to pierce	vi 1110	spirar

Table 9: Verbs derived from nouns that denote an instrument

V	English	Ν	English
aérer	to air	air	air
amidonner	to starch	amidon	starch
argenter	to silver	argent	silver
bétonner	to concrete	béton	concrete
beurrer	to butter	beurre	butter
bitumer	to asphalt, to tarmac	bitume	bitumen
charbonner	to blacken	charbon	coal
chiffonner	to crumple	chiffon	mousseline, rag
cimenter	to cement	ciment	cement
cirer	to polish (shoes, the floor)	cire	wax
crotter	to muddy	crotte	dropping
cuivrer	to bronze, to copper	cuivre	copper
émailler	to enamel	émail	enamel
fariner	to flour	farine	flour
ferrer	to shoe (a horse)	fer	iron
feutrer	to felt	feutre	felt
enfieller	to fill with bile	fiel	bile, venom
enfumer	to fill with smoke	fumée	smoke
gazer	to gas	gaz	gas
givrer	to frost over	givre	frost
goudronner	to tarmac	goudron	tar
graisser	to grease	graisse	grease
huiler	to oil	huile	oil
larder	to lard	lard	fat streaky baco
pimenter	to put chillies in	piment	hot pepper
plastiquer	to carry out a bomb attack on	plastic	plastic explosive
plâtrer	to plaster	plâtre	plaster
plomber	to fill (a tooth), to seal	plomb	lead
poivrer	to pepper	poivre	pepper
poudrer	to powder	poudre	powder
rouiller	to rust	rouille	rust
sabler	to sandblast	sable	sand
saigner	to bleed	sang	blood
savonner	to rub soap on	savon	soap
saler	to salt	sel	salt
sucrer	to put sugar in	sucre	sugar

Table 10: Verbs derived from nouns that denote a substance

V	English	Ν	English
cuver engainer engranger emmagasiner emprisonner	to archive to ferment to put into its sheath to gather in, to store to store to imprison to register	archives cuve gaine grange magasin prison registre	archive tank sheath barn store prison register

Table 11: Verbs derived from nouns that denote a container

References

- Aurnague, Michel & Marc Plénat. 1997. Manifestations morphologiques de la relation d'attachement habituel. In Danielle Corbin, Bernard Fradin, Benoît Habert, Françoise Kerleroux, Marc Plénat & SILEX (eds.), *Mots possibles et mots existants* (Silexicales), 15–24. Villeneuve d'Ascq: URA 382 - CNRS & Université de Lille 3.
- Aurnague, Michel & Marc Plénat. 2007. *Contraintes sémantiques et dérivation en é-: attachement habituel, naturalité et dissociation intentionnelle*. Carnets de Grammaire 16. CLLE-ERSS.
- Aurnague, Michel & Marc Plénat. 2008. Sémantique de l'espace et morphologie : le cas de la préfixation en *é*-. *Bulletin de la Société de Linguistique de Paris* CIII. 201–236.
- Booij, Geert. 2010. Construction morphology. Oxford: Oxford University Press.
- Booij, Geert & Jenny Audring. 2017. Construction morphology and the parallel architecture of grammar. *Cognitive science* 41(S2). 277–302.
- Fradin, Bernard. 2005. On a semantically grounded difference between derivation and compounding. In Wolfgang U. Dressler, Dieter Kastovsky, Oskar E. Pfeiffer & Franz Rainer (eds.), *Morphology and its demarcations*, 161–182. Amsterdam: John Benjamins.
- Lang, Ewald & Claudia Maienborn. 2011. Two-level semantics: Semantic form and conceptual structure. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, 709–740. Berlin: De Gruyter Mouton.
- Mayo, Bruce, Marie-Therese Schepping, Christoph Schwarze & Angela Zaffanella. 1995. Semantic in the derivational morphology of Italian: Implications for the structure of the lexicon. *Linguistics* 33. 883–938.
- Namer, Fiammetta & Evelyne Jacquey. 2003. Lexical semantics and derivational morphology: The case of the popular 'é-' prefixation in French. In 2nd International Workshop on Generative Approaches to the Lexicon, 115–123.
- Namer, Fiammetta & Evelyne Jacquey. 2012. Word formation rules and the generative lexicon: Representing noun-to-verb versus verb-to-noun conversion. *Advances in Generative Lexicon Theory*. 385–414.

- Schwarze, Christoph. 2011. Wie syntaktisch ist die Morphologie? Der vierte Verbstamm des Italienischen in einem Schichtenmodell. In Was grammatische Kategorien miteinander machen: Form und Funktion in romanischen Sprachen von Morphosyntax bis Pragmatik: Festschriftø für Ulrich Wandruszka, 57–78. Tübingen: Narr.
- Schwarze, Christoph. 2012. Le vert, les jeunes, l'important: Aspects sémantiques de la conversion adjectif-nom. Scolia 26.
- Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure and lexical form. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 3, 57–150. Cambridge: Cambridge University Press.
- Todaro, Giuseppina. 2017. Nomi (e aggettivi) che diventano verbi tramite prefissazione: quel che resta della parasintesi. Università Roma Tre & Université Toulouse Jean-Jaurès dissertation.
- Tribout, Delphine. 2010a. How many conversions from verb to noun are there in French? In *Proceedings of the HPSG 2010 conference*, 341–357. Stanford: CSLI Publications.
- Tribout, Delphine. 2010b. *Les conversions de nom à verbe et de verbe à nom en français*. Université Paris Diderot - Paris 7 Thèse de doctorat.
- von Heusinger, Klaus & Christoph Schwarze. 2006. Underspecification in the semantics of word-formation: The case of denominal verbs of removal in Italian. *Linguistics* 44(6). 1165–1194.