



HAL
open science

Discrepancy between form and meaning in Word Formation: the case of over- and under-marking in French

Nabil Hathout, Fiammetta Namer

► To cite this version:

Nabil Hathout, Fiammetta Namer. Discrepancy between form and meaning in Word Formation: the case of over- and under-marking in French. Franz Rainer; Francesco Gardani; Hans Christian Luschützky; Wolfgang U. Dressler. Morphology and Meaning. Selected papers from the 15th International Morphology Meeting, Vienna, February 2012, 327, John Benjamins, pp.177-190, 2014, Current Issues in Linguistic Theory, 9789027248466. 10.1075/cilt.327.12hat . hal-00989259

HAL Id: hal-00989259

<https://hal.science/hal-00989259>

Submitted on 5 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Discrepancy between form and meaning in Word Formation: the case of over- and under-marking in French

Nabil Hathout⁽¹⁾ / Fiammetta Namer⁽²⁾

(1) CLLE-ERSS, CNRS & Université de Toulouse Le Mirail

(2) Université de Lorraine & ATILF - UMR 7118

Abstract

Most word formation (WF) theories assume that derivatives tend in general towards canonicity, that is one-to-one correspondence between form and meaning and thus full morpho-semantic transparency. However, form-meaning mismatches actually have a widespread coverage, both in terms of languages and WF rules. These mainly fall into two types of discrepancy: over-marking and under-marking. In this paper we propose a classification of these deviations with a distinction between a derivational level and a lexical one. We illustrate this classification with examples from French and other languages: English, Italian and Dutch. Then, we sketch a unified analysis of these deviations, within a word-based framework. We propose to analyse the relative importance of canonicity and discrepancies in WF from the perspective of the interaction between the speaker and the hearer.

Introduction¹

In this paper, we focus on what may be called “deviant word formation”, that is Word Formation (WF) departing from the canonical one-to-one correspondence between form and meaning, and thus moving away from morpho-semantic transparency. We will show that form-meaning mismatches have a widespread coverage, both in terms of languages and word formation rules (WFR). Despite an apparent heterogeneity, cases of form-meaning discrepancy can be given a consistent explanation and a unified analysis. Our interest in this type of discrepancy arises from the description of two French derivations, the prefixation in *anti-* (Hathout 2011) and the suffixation in *-aliser* (Namer 2013) illustrated in (1a) and (1b). The meaning of the derivative in (1a) is ‘against parliament’ while the literal interpretation of its form is ‘against parliamentarians’. Similarly, in (1b) *instrumentaliser* means ‘make something become an instrument’ while its form leads to the interpretation ‘make something become instrumental’.

¹This document is under copyright.

Hathout N, Namer F. Discrepancy between form and meaning in word-formation: The case of over- and under-marking in French. In: Rainer F, Gardani F, Luschützky HC, Dressler WU (eds). *Morphology and Meaning: Selected papers from the 15th International Morphology Meeting*, Vienna, February 2012. Amsterdam: John Benjamins Publishing Company; 2014. p. 177-190. [doi.org/10.1075/cilt.327.12hat].

- (1) a. Fr: parlement_N → antiparlementaire_A
 ‘parliament’ ‘against parliament’
 b. Fr: instrument_N → instrumentaliser_V
 ‘instrument’ ‘make something become an instrument’

Both descriptions have been performed on fairly large corpora of derivatives: 3000 for *anti-* and 700 for *-aliser*. These derivatives come from various sources: the French dictionary *Trésor de la Langue Française*; text corpora such as newspapers and corpora collected on the Web; direct interrogation of Web search engines.

We were puzzled in both studies by the sheer number of non-canonical derivatives. These mainly fall into two types of discrepancy: over-marking and under-marking. In this paper we propose a tentative classification of these deviations with a distinction between a derivational level and a lexical one. We illustrate this classification with examples from French and other languages: English, Italian and Dutch. In the third part of the paper, we sketch a unified analysis of these deviations, within a word-based framework.

1. *Morphological discrepancies between meaning and form*

The existence and nature of a mismatch between form and meaning in some derivation depends strongly on the theoretical framework of WF used for its analysis. In the word-based framework adopted in this paper, morphology is relational. Derivatives are connected with their bases by relations on three levels: formal, categorical and semantic. Some of the discrepancies we present hereafter have received specific analyses in morpheme-based WF frameworks. The word-based framework used in this study allows us to propose a unified explanation for the broad range of form-meaning mismatch types we are interested in. This section presents these types, which can be defined in terms of a multi-dimensional distance from a standard: canonicity.

Corbett (2010) among others has addressed the issue of canonicity in derivational morphology. For a derivation to be canonical there must be a regular one-to-one correspondence between the elements of form and the elements of meaning in the derived word. In practice, we only consider the changes associated with the derivation; the meaning and the form of the base can be seen as elementary. Canonicity can be illustrated by the example in (2) where two parts can be identified in the form of *singer*, each of them being in a regular correspondence with one meaning element.

- (2) form: sing → sing er
 meaning: ‘sing’ → ‘sing’ ‘person who Vs’

In (2), the phonological sequence corresponding to *er* matches a meaning element which can be glossed as ‘person who Vs’. Actually, a general principle of derivational morphology is that **derivation is semantically incremental**: the meaning of a derived word is always defined with respect to the entire meaning of its base. This simply follows from the fact that the base is involved in the process as an “integral” word.

Deviation with respect to canonical derivation may be looked at from two points of view: the derivations and the lexicon. We will distinguish the two levels and will say that a deviation is derivational when it is considered at the level of a single derivation and that it is lexical when it is considered at the level of the entire lexicon. Therefore, derivational deviations are local and concern individual base-derivative relations, whereas lexical deviations are global and involve sets of derivations.

1.1 *Form-meaning discrepancy from the point of view of derivations*

Derivative-base relations fall into four types with respect to the meaning-form correspondences. The first is **canonicity** when the meaning elements perfectly match all the form elements and only them as illustrated by (2). The three other types are deviant with respect to derivational canonicity.

The first is **derivational over-marking** where some parts of the form do not correlate with any element of the meaning. More precisely, the derivative (*D*) is over-marked with respect to the base (*B*) if some formal elements do not correlate with any meaning element. In this case, there is a one-to-many correspondence between the meaning elements and the formal marks. An example of over-marking is given in (3) where the formal element *-in-* adds no contribution to the meaning of *paginate*. This phenomenon is known as “empty morpheme” in classical WF theories, (see, e.g. Beard 1995). In Section 3, we show how word-based frameworks enable us to do without such a device.

(3) En: page_N → paginate_V

The second type is **derivational under-marking** where some meaning elements are not formally marked as in the converted verb *salt* in (4). The meaning contribution of the conversion (‘add X to’) has no formal counterpart in the form of the verb. In an under-marked derivation, there is a many-to-one correspondence between the meaning elements and the formal marks. We will see that we do not need the “zero-morpheme” device to account for this phenomenon.

(4) En: salt_N → salt_V

The third type is **simultaneous over- and under-marking**, as in (5): on the one hand the meaning elements marked by *-ism* in *B* have no formal counterpart in the form of *D*; on the other hand, the sequence *-ist-* in *pessimistic* is an extra mark with no contribution to the meaning of the derivative. More specifically, *D* is over- and under-marked with respect to *B* if there is a many-to-many correspondence between the meaning and form elements of *B* and *D*.

(5) En: pessimism_N → pessimistic_A

1.2 Form-meaning discrepancy at the lexicon level

Deviation from one-to-one correspondence may also be observed from a lexical point of view focussing on in relations between bases and derivatives instead of correspondences between meaning and form elements. More specifically, canonicity and deviation from canonicity can be extended to sets of derivations obtained by **the same WF operation**. These sets fall into four types. These are similar to the ones we have just proposed for the meaning-form correspondences at the derivational level. The first type is **lexical canonicity** which corresponds to the cases where one derivation connects one base with one derivative as in (2). The three other types correspond to cases of **lexical deviations** where multiple derivations connect a given base with more than one derivative, a given derivative with more than one base or several bases with several derivatives.

The first of them is **lexical over-marking**. It can be defined as follows: a set of derivations is lexically over-marked if they connect one *B* with several words D_1, \dots, D_n , all derived from *B* by the same operation but have distinct forms. This one-to-many correspondence is a kind of derivational over-abundance (Thornton to appear). Lexical over-marking may be illustrated by (6) where the adjectives *idealist* and *idealistic* derive from the same base *ideal* and are synonyms.

(6) En: ideal_N → idealist_A
ideal_N → idealistic_A

This example shows that lexical deviation is directly related to derivational deviation: *idealistic* is an over-marked derivative. Actually, derivational and lexical deviations grasp the same phenomenon. The only difference between the two is that the former deals with the phenomenon at the word level while the latter considers it at the level of the entire lexicon. This will be made clearer in the remainder of the paper. More generally, it is impossible to have more than one canonical derivative obtained by the same operation from one base or to have one derivative canonically derived from more than one base via the same operation.

The second type of lexical deviation is **lexical under-marking** when more than one base are connected with a derived word by the same operation as *localize* in (7) which has three possible bases. Note that the meaning of *localize* varies according to the chosen base: ‘make X become local’ for the first, ‘limit or confine X to a locality’ for the second and ‘identify the location of X’ for the third. Lexical under-marking characterizes a many-to-one base-derivative correspondence.

(7) En: local_A ↘
 locality_N → localize_v
 location_N ↗

The third type of lexical deviation corresponds to the cases of many-to-many correspondences, namely sets of derivations between more than one base and more than one derivative. This type is illustrated in Section 2.2.3.

2. Derivation and lexicon: two independent levels for the description of morphological discrepancy

In this section, we show that form-meaning discrepancy is ubiquitous in derivational morphology. We present a number of representative examples of the various types of discrepancy that have just been defined. In the following, DD stands for “derivational deviation” and LD for “lexical deviation”. As we will see, these are two independent dimensions, which cannot be reduced one to the other. Both DD and LD ranged over four values we call *deviation values*: 1:1 represents canonicity, 1:n denotes over-marking, n:1 under-marking, and n:n simultaneous over- and under-marking. In each example, we adopt the same typographic code: underlined sequences correspond to the rule exponent (i.e. expected prefix, suffix or compounding element), and strings in boldface to the extra sequences with respect to the 1:1 derivational situation.

2.1 Derivational deviations

2.1.1 DD = 1:n

One-to-many DD is expressed by an extra formal mark on the derivative, with respect to its meaning. Typical examples of such over-marking are the so-called “parasynthetic derived words”, originally defined in Darmesteter (1875: 96-103) (see also Allen 1981; Iacobini 2004; Scalise 1994; Serrano Dolader 1995), that can be found in several European languages, and especially in Romance ones, cf. (8) and (9).

(8) En: bacteria_N → anti**bacterial**_A

- (9) It: mare_N → sottomarino_A
 ‘sea’ ‘submarine’

In these examples, the derived words are obtained by means of various prefixation processes, and an extra suffix mark shows up on each prefixed words. In (8), the meaning of the adjective *antibacterial* is formed on that of the noun *bacteria*, whereas it is formally obtained from the existing adjective *bacterial*. *Anti-* being the exponent of the prefixation, the final *-al* is clearly an extra mark with no contribution to the meaning of the derivative. Similar derived words exist in Italian as in (9) where *sottomarino* ‘under the sea’ displays an extra mark *-in* with respect to the form of its base *mare*.

- (10) Fr: mort_N → mortalité_N
 ‘dead’ ‘mortality’
 nation_N → nationalisme_N
 ‘nation’ ‘nationalism’
- (11) Fr: goutte_N → gouttelette_N
 ‘drop’ ‘droplet’

One-to-many deviation also occurs with suffixation processes, as illustrated in (10) and (11). In (10), the *-al-* extra mark is that of an existing adjective. As Koehl (2009) points out, the noun *mortalité* (resp. *nationalisme*, cf. Roché 2011b) is semantically derived from the noun *mort* (resp. *nation*) since it means ‘rate of dead people’ (resp. ‘ideology that favours the nation’), though its stem is the adjective *mortel* ‘deathly’ (resp. *national*). Example (11) illustrates a variation of this situation (Plénat & Roché, 2004). The suffixed noun *gouttelette* is defined with respect to the noun *goutte*, which surfaces as the stem *gouttel*. However, unlike (10), *gouttel* is not a nominal stem; it does not correspond to any attested word.

2.1.2 DD = n:1

Apart from conversion, cf. (4), derivational under-marking occurs when parts of the base formal material are not present in the derived word. This is typical of word formation process types across languages, as the examples in English, French, Italian, and Dutch show: back-formation (12), blending (13), and cross-formation (14, 15); (on back- and cross-formation, see e.g. Adams 1981; Becker 1993; Nagano 2007; Shimamura 1983; Szymanek 2005; on blending, see Dressler 2000; Fradin 2000; Warren 1990 among others). DD n:1 also concerns derivation from a multi-lexical base, where only one part of this basis surfaces in the derivative (16) (Namer 2013).

- (12) En: mass-production_N → mass-produce_v
 (13) biography_N + picture_N → biopic_N
 (14) It: comunismo_N → comunista_N

- ‘communism’ → ‘communist’
 (15) Nl: Vlissingen_{NPR} → vlissingen_N²
 (16) Fr: bombe_N atomique_A → anti-atomique_A
 ‘atomic bomb’ → ‘anti-atomic’

2.1.3 DD = n:n

Finally, in many-to-many DD, an extra formal mark occurs on both the base and the derived word. In (17), the sequence *-isme* does not show up in *anti-abolitionniste*, which in turn includes the supernumerary mark *-iste* which plays no semantic role in the contribution to the meaning of the derivative with respect to the base *abolitionnisme* (Roché 2011a).

- (17) Fr: abolitionnisme_N → anti-abolitionniste_A
 ‘abolitionism’ → ‘anti-abolitionist’

2.2 Lexical deviations

Unlike DD, lexical deviation (LD) is defined within a given relation between the input and the output of a WFR. It is characterized by a supernumerary base, or, conversely, by too many derived words. LD necessarily involves DD: one of the inputs (and/or outputs) in excess is either over- or under-marked in terms of DD value.

2.2.1 LD = 1:n

In lexical one-to-many deviation, several derived words sharing the same meaning derive from one unique base by means of the same WFR, as illustrated in Table 1. ‘Base’ rows 1 in Table 1 provide four examples of WFR inputs numbered from (a) to (d), each of them being the base of two or three synonymous derivatives. Each base/derivative is either derivationally canonical (rows 1:1) or deviant (rows 1:n, n:1 and n:n).

Examples (a) and (b) are similar. In (a) two adjective formal variants derive from the noun base (Lindsay & Aronoff 2013). One is directly obtained by the *-ic* suffixation. The other one (*historical*) is marked by the additional sequence *al*. With (b), we can see that there can be more than one extra-marked derivative.

In examples (c) and (d), the non-canonical base/derivative relation is either a DD n:n (*piratage* → *antipiratable*) or a DD n:1 (*fraisier* → *fraisie*). Notice also that a deviant output form may take advantage of the morphological complexity of its base. The noun *fraisier* is derived from *fraise*, and the noun of the plantation derived in *-aie* is either *fraiseraie* or *fraisie* (Roché 2011b, 2011c).

² See (Booij 2002) for an analysis of toponyms, ethnonyms and ethnic adjectives in the framework of construction grammars.

base	(a) En: history _N	(b) It: monopolio _N ‘monopoly’
1:1	histori <u>C</u> _A	antimonopolio _A
1:n	histori <u>cal</u> _A	antimonopolista _A antimonopolistico _A
n:1		
n:n		
	‘related to history’	‘against monopoly’
base	(c) Fr: piratage _N ‘piracy’	(d) Fr: (fraise _N ‘strawberry’) → fraisier _N ‘strawberry plant’
1:1	antipiratage _A	fraiseraie _N
1:n		
n:1		fraisai <u>e</u> _N
n:n	antipiratable _A	
	‘against piracy’	‘strawberry plantation’

Table 1: One-to-many lexical deviations

2.2.2 $LD = n:1$

In a both structurally and semantically opposite situation, two or more inputs serve as bases for a single, ambiguous, output form. Table 2 illustrates this systematic construction of polysemous derivatives.

	(a)		(b)	
base	institution _N	institutionnel _A ‘institutional’	mensuel _A ‘monthly’	salair <u>e</u> _N mensuel _A ‘monthly salary’
1:1	institutionnalis <u>e</u> _V		mensualis <u>e</u> _V	
1:n	institutionnalis <u>e</u> _V		mensualis <u>e</u> _V	
n:1	institutionnalis <u>e</u> _V		mensualis <u>e</u> _V	
	‘put (smb) into an institution’	‘make (smth) become institutional’	‘make (smth) become monthly’	‘provide (smb) with a monthly salary’

Table 2 : Many-to-one lexical deviations

Two sorts of lexical under-marking are illustrated in Table 2.

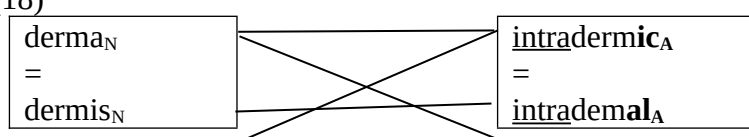
In (a), the output verb (*institutionnalis*) has two readings, depending on the selected base. However the first base stem (*institution*) blends in with the adjective form (*institutionnel*). The resulting base/derivative mismatch is a case of DD 1:n (*institution/institutionnalis*). In (b), one of the bases of

the polysemous verb *mensualiser* is the AN noun phrase *salaire_N mensuel_A*. The missing sequence on the derivative is thus the noun *salaire*.

2.2.3 LD = n:n

The last sort of lexical deviation can be qualified as a many-to-many base/derivative correspondence: a given WFR relates several bases with several derivatives. First, in the English example of *intra-* prefixation in (18), each of the two synonymous variants *derma* and *dermis* is derived into either *intradermic* or *intradermal*, which also are synonymous. Both adjectives are derivationally over-marked (DD 1:n).

(18)



In Figure 1, both French bases *obèse_N* and *obésité_N* derive from the same adjective (*obèse*) and are therefore indirectly related. When based on *obésité*, the derived adjectives are synonymous (*anti-obèse* = *anti-obésité*, ‘against obesity’). However, *anti-obèse* can also be related to *obèse_N*. Therefore it has a second reading: ‘against obese people’.

Figure 1 involves two DD values: canonicity, as far as *obèse/anti-obèse* and *obésité/anti-obésité* pairs are concerned, and n:1 for the *obésité/anti-obèse* relation.

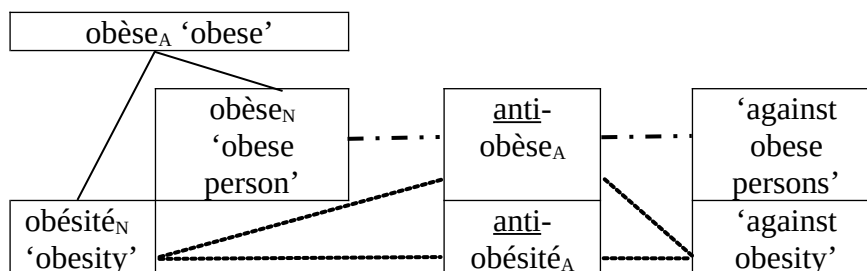


Figure 1: Many-to-many lexical deviation

3. A unified analysis of deviation

The examples in Section 2 show that deviation is wide-spread among languages and how it can be seen from a derivational and a lexical point of view. Several questions arise from this review: why do languages make use of non-canonical derivations and derived words? What role do canonicity

and discrepancy play in WF? Are canonical derivatives “superior” to deviant ones? What mechanisms are involved in the creation of the deviant derivatives?

In this section, we propose a unified analysis of form-meaning deviations in WF. It stems from a word-based framework where the objects relevant to derivational morphology are the lexemes, the derivations being seen as relations between lexemes. In this framework, canonicity and discrepancy result from a generalized optimization process that operates at once on all the morphological types. The optimization aims at the maximization of the efficiency of the speaker-hearer interaction (Section 3.2). It is implemented as a set of partially contradictory constraints (Section 3.1) reminiscent of the ones used in the optimality theory (Prince & Smolensky 1993).

3.1 Constraints at the WF level

The examples presented in Section 2 show that form and meaning are independent components of the lexemes although connected by correspondences. Optimization is performed on each component and each correspondence. The form of a derivative is optimal if it is the best one morpho-phonologically. It is subject to various constraints including the well-known **dissimilative constraints**, illustrated in (10)-(11), where the sequence *al* in *mortalité*, or *el* in *gouttelette*, prevents the consecution of two /t/ in °*mortité* /mɔʁtite/, and in °*gouttette* /gutɛt/ cf. (Plénat 2011) or the **constraints on size** that, in French, favor derivatives with bi-syllabic stems. These constraints tend to reduce the derivational canonicity as in the French example *milanais*_A ‘Milanese’ → *milanissime*_A ‘Milanese to an extreme degree’ where the final *-ais* is truncated so that the length of the stem is closer to two syllables (Plénat 2009).

Lexemes are also subject to constraints on the correspondences between their components. Constraints on the **form-meaning correspondence** favor the lexemes with a meaning perfectly described by the form (e.g. *singer* in (2)) and strongly contribute to derivational canonicity. On the other hand, constraints on **correspondence between form and category** can be detrimental to derivational canonicity. They give preference to the forms that look most like forms of derived word of the same part-of-speech (Corbin 2001, Hathout 2011). The strategy commonly used to satisfy these constraints is to create forms that display an ending typical of derived words of that part of speech, most often a suffix. For instance, *antiparlement*_A, ending in *-ment*, looks more like a noun and *antiparlementaire*_A (1a), ending in *-aire*, more like an adjective. The same explanation applies to similar cases of over-marked prefixed words such as (8) and (9).

A partial conclusion drawn from the four above-mentioned constraints is that canonicity and discrepancy are essentially of the same nature. Both result from the fulfillment of very similar constraints. The term deviation

therefore seems rather misleading: deviations are not errors but merely means used by the speakers to satisfy these constraints. In 3.2, we focus on the relative importance of canonicity and discrepancy in WF and more precisely on the balance between the forces that cause them. We propose to look at the interplay between these forces from the perspective of the interaction between the speaker and the hearer. Both share the same objective, namely the success and the efficiency of their interaction and this common goal has direct effects on WF.

3.2 *The speaker-hearer interaction point of view*

The speaker-hearer interaction has been the subject of many studies. For Zipf (1949), there is a trade-off between two opposing forces: unification which causes lexical and derivational deviation and diversification which results in canonicity. For Levinson (2000), the balance lies between the hearer inferences and the speaker effort. He claims that “inference is cheap, articulation expensive and thus the design requirements are for a [communication] system that maximizes inferences” and therefore disfavors canonicity. More recently, Piantadosi *et al.* (2012) looked at the same question from an information theoretic perspective. They concluded that “ambiguity allows efficient linguistic units to be preferentially re-used decreasing the overall effort needed to use a linguistic system.” In other words, lexical canonicity is expensive and it reduces the efficiency of the interaction. Conversely, the reuse of efficient word forms causes lexical deviation while the reuse of the efficient word form parts induces derivational deviation.

The trade-off between the interests of the speaker and the ones of the hearer can be spelled out in terms of WF optimization towards canonicity or discrepancy. Speakers make use of different strategies in order to achieve optimal derivatives. They all are based on the reuse of formal and semantic properties. The first is full recycling, that is the association of new meanings to existing forms and new forms with existing meanings. Recycling is one of the main causes of lexical and derivational deviation. The association of new meanings yields to lexical under-marking (see (a) and (b) in Table 2, e.g. *institutionnaliser*, or the polysemous *anti-obèse* in Figure 1) while that of new forms yields to lexical over-marking (see examples in Table 1, e.g. *historic* and *historical*). Recycling reduces the number of the memorized forms and meanings; it preserves lexical regularities and strengthens existing series and subseries.

A second strategy is partial borrowing of (large) chunks from close words. The borrowed material is often the stem of the derivative as in *antiparlementaire* (1a) where *parlementaire* is the form of a close member of the family of the base, *parlement* (see also (8)-(11) or (18)). In other cases, borrowing may help a stem obtain an epenthetic consonant as in

English *manga*_N → *mangakist*_N (author of manga), which borrows its stem from *mangaka*, the Japanese word for cartoonist. Sometimes, derivatives may borrow material from non-related words as in Fr: *bambou*_N ‘bamboo’ → *bamboulesque*_A ‘relative to bamboo’. The stem *bamboul-* is borrowed from the unrelated *bamboula*_N ‘party’ which only happens to have a very similar form (Plénat 2011). Borrowing is therefore an alternative to recycling. It increases the similarity of the derivatives with the existing forms, that is, with forms that have proven their efficiency and their fitness.

A third strategy is to give preference to the prominent series and subseries in order to strengthen the global organization of the lexicon. The words inserted in these subseries benefit from the possible interpretative automatism dedicated to their members. The selection of efficient series and subseries enables the reuse of the formal and the semantic features which characterize them. For instance, *institutionnaliser* (Table 2) is preferred to *institutionniser*_V because *-aliser* verbs are a prominent and highly productive subseries within the *-iser* verbs’ series: they are strongly connected to other prominent series (*X-alité*, *X-el*, *X-ellement*, etc.) and represent a large portion of the including series (21% of the *-iser* verbs in the TLF end in *-aliser*).

4. Conclusion

Three main facts have been explored in this paper:

1. The manifestations of discrepancy with respect to a situation regarded as canonical are numerous; the canonical situation being transparent form-meaning correspondence.
2. The effects of discrepancy are either a formal over-marking of the derived words or ambiguity with respect to the form and to the semantic content of the base.
3. The range and frequency of discrepancy across languages clearly show that these form-meaning deviations never penalize the understanding nor the production of the forms.

Form-meaning mismatches are traditionally accounted for with various devices. We have here proposed a unified explanation based on a general strategy applied by the speaker: the reuse of efficient units. It is implemented by borrowing the stems of derived words from other existing forms and by their integration into prominent series. This strategy is guided by an aim common to the speaker and the hearer: optimize their interaction by balancing two communicative pressures, clarity and ease. As Piantadosi *et al.* (2012) point out:

a clear communication system is one in which the intended meaning can be recovered from the signal with high probability by the hearer.

An easy communication system is one which signals are efficiently produced, communicated and processed by the speaker. (p.281)

As we have shown, over-marking does not compromise the ease, nor does ambiguity alter clarity.

The production and understanding of canonical and deviant derived forms make up a coherent whole. The key to explaining the working of the system lies in the clear convergence and the necessary mutual influence between two domains generally studied separately: grammar and communication. Therefore, further directions of our research will include the study of the mechanisms responsible for these interactions.

References

- Adams, Valerie. 2001. *Complex Words in English*. Harlow: Longman.
- Allen, Andrew S. 1981. "The development of prefixal and parasynthetic verbs in Latin and Romance". *Romance Philology* 35.79-88.
- Beard, Robert. 1995. *Lexeme-Morpheme Based Morphology: a General Theory of Inflection and Word Formation*. Albany: SUNY Press.
- Becker, Thomas. 1993. "Back-formation, cross-formation, and 'bracketing paradoxes' in paradigmatic morphology". *Yearbook of Morphology* 1992.1-27.
- Booij, Geert. 2002. *The Morphology of Dutch*. New York: OUP.
- Corbett, Greville. 2010. "Canonical Derivational Morphology". *Word Structure* 3.141-155.
- Corbin, Danièle. 2001. "Du nouveau sur beurre laitier. Note sur une fausse conversion". *Par monts et par vaux. Itinéraires linguistiques et grammaticaux*, ed. by Claude Buridant, Georges Kleiber & Jean-Christophe Pellat, 127-143. Louvain & Paris: Peeters.
- Darmesteter, Arsène. 1875. *Traité de la Formation des mots composés dans la langue française, 2ème édition: 1894*. Paris: Honoré Champion.
- Dressler, Wolfgang U. 2000. "Extragrammatical vs marginal morphology". *Marginal and Extragrammatical Morphology*, ed. by Ursula Doleschal & Anna M. Thornton, 1-10. München: Lincom Europa.
- Fradin, Bernard. 2000. "Combining forms, blends and related phenomena". *Extragrammatical and Marginal Morphology*, ed. by Ursula Doleschal & Anna M. Thornton, 11-59. München: Lincom Europa.
- Hathout, Nabil. 2011. "Une analyse unifiée de la préfixation en *anti-*". Roché, Boyé, Hathout, Lignon & Plénat 2011, 251-318.
- Iacobini, Claudio. 2004. "Parasintesi". *La formazione delle parole in italiano*, ed. by Maria Grossmann & Franz Rainer, 165-188. Tübingen: Niemeyer.
- Koehl, Aurore. 2009. "Are French *-ité* Suffixed Nouns Property Nouns?".

- Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux*, ed. by Fabio Montermini, Gilles Boyé & Jesse Tseng, 95-110. Somerville, Mass.: Cascadilla Proceedings Project.
- Levinson, Steven C. 2000. *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge, Mass.: MIT Press.
- Lindsay, Mark & Mark Aronoff. 2013. "Natural selection in self-organizing morphological systems". *Morphology in Toulouse*, ed. by Nabil Hathout, Fabio Montermini & Jesse Tseng. München: Lincom Europa. 133-153.
- Nagano, Akiko. 2007. "Marchand's analysis of back-formation revisited: back-formation as a type of conversion". *Acta Linguistica Hungarica* 54.33-72.
- Namer, Fiammetta. 2013. "Adjectival bases of French *-aliser* and *-ariser* verbs: syncretism or under-specification?". *Morphology in Toulouse*, ed. by Nabil Hathout, Fabio Montermini and Jesse Tseng. München: Lincom Europa. 185-210.
- Piantadosi, Steven, Harry Tily, & Edward Gibson. 2012. The communicative function of ambiguity in language. *Cognition* 122:3.280-291.
- Plénat, Marc. 2009. "Les contraintes de taille". *Aperçus de Morphologie du français*, ed. by Bernard Fradin, Françoise Kerleroux & Marc Plénat, 47-64. Paris: Presses Universitaires de Vincennes.
- Plénat, Marc & Michel Roché. 2004. "Entre morphologie et phonologie: la suffixation décalée". *Lexique* 16.159-198.
- Prince, Alan & Paul Smolensky. 1993. *Optimality Theory. Constraint Interaction in Generative Grammar*. Piscataway: Rutgers Center for Cognitive Sciences - Rutgers University.
- Roché, Michel. 2011a. "Quel traitement unifié pour les dérivations en *-isme* et en *-iste*". Roché, Boyé, Hathout, Lignon & Plénat 2011, 69-143.
- Roché, Michel. 2011b. "Quelle morphologie?". Roché, Boyé, Hathout, Lignon & Plénat 2011, 15-39.
- Roché, Michel. 2011c. "Pression lexicale et contraintes phonologiques dans la dérivation en *-aie* du français". *Linguistica* 51.5-22.
- Roché, Michel, Gilles Boyé, Nabil Hathout, Stéphanie Lignon & Marc Plénat. 2011. *Des Unités Morphologiques au Lexique*. Paris: Hermès.
- Scalise, Sergio. 1994. *Morfologia*. Bologna: Il Mulino.
- Serrano Dolader, David. 1995. *Las formaciones parasintéticas en español*. Madrid: Arco Libros.
- Shimamura, Reiko. 1983. "Backformation of English Compound Verbs". *Papers from the parasession on the interplay of phonology, morphology and syntax*, ed. by John F. Richardson, Mitchell Marks & Amy Chukerman, 271-282. Chicago: Chicago Linguistic Society.
- Szymanek, Bogdan. 2005. "The latest Trends in English Word-Formation". *Handbook of Word-Formation*, ed. by Pavel Štekauer & Rochelle Lieber, 429-448. Dordrecht: Springer.

- Thornton, Anna M. to appear. "Reduction and maintenance of overabundance. A case study on Italian verb paradigms". *Word Structure* 5.
- Warren, Beatrice. 1990. "The importance of combining forms". *Contemporary Morphology*, ed. by Wolfgang U. Dressler, Hans Ch. Luschützky, Oskar E. Pfeiffer & John R. Rennison, 111-132. Berlin & New York: Mouton - Walter de Gruyter.
- Zipf, George K. 1949. *Human behavior and the principle of least effort*. New York: Addison-Wesley.