The meaning of Italian VN compounds in a diachronic perspective: the development of a word-formation rule and its semantic instruction

Pavel Štichauer
Charles University in Prague
pavel.stichauer@ff.cuni.cz

1. Introduction: the premises and the aim of the study[1]

The VN compounding (of the type *portalettere*) in Italian (and, in general, in the Romance languages, cf. Gather 2001, Bauer 2010, among many others) is a well-studied and well-described phenomenon. However, recent research has shown that at least two questions are worth being further developed. First, it is the quantitative aspect of the productivity and of the overall distribution of the VN compounds in large – synchronic – corpora (see Ricca 2010). Second, it is the diachronic development of the pattern in question that has come to the foreground (see Moyna 2011 for Spanish).

The aim of this paper is to combine both aspects and to offer a *partial corpus-based diachronic picture* covering the period that goes from the 16th to the 19th century, based on selected data drawn from the LIZ 4.0; and, at the same time, a *partial dictionary-based diachronic description* of the pattern covering, in a complementary way, the same period, based on four editions of the *Vocabolario degli Accademici della Crusca* (ranging from 1612 to 1738) and on the dictionary *il Tommaseo* (2004) by Niccolò Tommaseo and Bernardo Bellini (1857-1879) (for a history of the Italian lexicography see Marazzini 2009).

This study is supposed to show that VN compounding, though being always *available*, has become more *profitable* – in the sense of Corbin's (1987:177) distinction between *disponibilité* and *rentabilité* – only over the period under investigation. It will also be shown that VN compounding has always been endowed with the same semantic instruction inherent in the word-formation rule (WFR), exhibiting thus all the known semantic outputs to be described below.

2. VN compounds in present-day Italian

As is well known, VN compounding is a well-settled productive pattern in present-day Italian (see, e.g., Tollemache 1945; Scalise 1994: 134-136; Bisetto 2004: 45-47; Radimský 2006: 95-101; Scalise & Bisetto 2008: 133-135; Dardano 2009: 190-198; Ricca 2010, among others).

The pattern consists in a combination of a verbal base with a noun which is usually its internal argument: *portalettere* 'postman', *guardaboschi* 'forester', *lavapiatti* 'dishwasher', *contagiri* 'tachometer', etc. (e.g., Scalise 1994: 134). However, not every imaginable verbal base is found in the attested compounds; the restrictions, as summarised by Ricca (2010: 243-246), seem to be phonological (preference for bisyllabic bases), morphological (avoidance of verbs with *-isc-*), syntactic (limitation to transitive / agentive verb bases), and probably also semantic, as the number of really productive verbal bases is rather low. In fact, corpus-based

and dictionary-based studies (Ricca 2010, 2005; Radimský 2006; Gather 2001) show that productive leading bases (Ricca 2010: 247) are mainly *porta-* 'bring', *salva-* 'save', *mangia-* 'eat', *acchiappa-* 'take', *ammazza-* 'kill', *copri-* 'cover', *taglia-* 'cut', *para-* 'block', *spacca* 'break' and *trita-* 'grind'.

Semantically, these compounds may be *agents, instruments, events, locations* (the last two being marginal), but they can also function as adjectival modifiers (qualifying/relational). Also, it was assumed that there was a primary meaning – that of agent – from which the other meanings could be diacronically derived, but recent research (see, especially, Moyna 2011, for Spanish) suggests that all semantic outputs tend to coexist since the early attestations.

3. VN compounds in diachrony

Interestingly, VN compounding in Romance is not a pattern with a stable diachronic productivity. In fact, as Moyna (2011: 206) puts it, its "beginnings are modest, less robust than those of many other compounding patterns, (...)."

Moyna (2011: 206) demonstrates that a real quantitative spread of the pattern can be traced back to as late as the 19th century: 70% of all VN compounds are attested precisely from the 19th century onwards.

In order to obtain a similar picture for Italian, the data from the 16th to the 19th century will be presented on a limited (but sufficiently representative) sample of 15 leading bases: *porta-* 'bring', *copri-* 'cover', *salva-* 'save', *lava-* 'wash', *conta-* 'count', *para-* 'block', *passa-* 'pass', *reggi-* 'bear', *rompi-* 'break', *apri-* 'open', *batti-* 'beat', *guarda-* 'watch', *mangia-* 'eat', *ammazza-* 'kill', *spacca* 'break', *taglia-* 'cut'.

4. Lexicographic data

The lexicographic evidence drawn from the major historical dictionaries shows clearly that the representation of the VN compounding is really weak; the overall number of lemmas (containing only the 15 leading bases introduced above) is captured in Table 1.

Table 1: VN compounds attested in historical dictionaries.

Dictionary	Period	Lemmas
Vocabolario degli Accademici della Crusca (1st edition)	1612	11
Vocabolario degli Accademici della Crusca (2nd edition)	1623	11
Vocabolario degli Accademici della Crusca (3rd edition)	1691	16
Vocabolario degli Accademici della Crusca (4th edition)	1729-1738	27
il Tommaseo	1857-1879	160

In the first two editions of *La Crusca*, the pattern is really marginal: there are only two bases (*batti-* 'beat', *guarda-* 'watch') which exhibit more than one formation; besides, they cannot really qualify for core formations (e.g., *batticuore* 'palpitations'); moreover, some VN compounds are even not given a separate lemma being used only inside of the definition (e.g., *parasole* turns up only as a synonym for the lemma *solecchio* 'sunshade').

In the 3rd and 4th editions, the newly entered compounds are mainly those with *porta-* 'bring' and *guarda-* 'watch' (*portamantello* 'clothes rack', *guardadonna* 'midwife'). It is only in the major 19th century dictionary *il Tommaseo* that a significant number of lemmas can be found; nevertheless, out of 160 compounds, 85 are covered by just three leading bases (28 *porta-* 'bring', 30 *guarda-* 'watch', 27 *mangia-* 'eat').

5. Corpus data

As far as the corpus data are concerned, it is important to note that they are drawn from a sample of Italian texts LIZ 4.0 (*Letteratura Italiana Zanichelli*, 2001). The selected texts, covering the time span under investigation (and divided into centuries according to the traditional periodization), represent the core of what is usually referred to as *italiano letterario.*[2]

However, the corpus data – within the subcorpora corresponding to the four centuries – are not directly comparable: the 16th and 19th century subcorpora do reveal an increasing tendency (they are roughly of the same size); the 17th and 18th century subcorpora are not sufficiently large but contain nevertheless some interesting examples to be discussed below. The overall situation is captured in Table 2.

Table 2: VN compounds attested across four subcorpora based on LIZ 4.0.

Period/subcorpus	Corpus size in tokens	Number of texts	Number of types
16th century	10 459 937	221	45
17th century	3 028 291	49	15
18th century	4 600 141	237	18
19th century	10 285 433	217	91

Qualitatively, the corpus situation is much similar to the lexicographic one. In the 16th century, the most nouns are those with *porta-* 'bring', *guarda-* 'watch' and *mangia-* 'eat' (e.g. *portanovelle* 'news-reporter', *guardadonna* 'midwife', *mangiaguadagni* 'profit-eater').

The 17th and 18th centuries (due to the reduced size of the corpus) register, on the whole, only those nouns already encountered. Still, some new formations (what is more, in the adjectival function) do turn up here (e.g. *guardacoste* 'coastguard' in the expression *vascello da guerra guardacoste* 'coastguard warship').

In the 19th century, the expansion concerns, as expected, mainly the leading bases *porta-* 'bring', *guarda-* 'watch', *mangia-* 'eat' (these three covering 43 formations out of 91), but also *para-* 'block' (there are some neologisms in Leopardi's *Zibaldone* along with the author's interesting comment on the pattern).

6. Meaning distribution in diachrony and the semantics of the WFR

The diachronic data, though limited qualitatively and quantitavely, permit us to answer some of the questions already alluded to.

First, there is the question of the semantic outputs. Are all earlier formations agents (with the typical disparaging connotation)? The answer is clearly no, as there are some instrument nouns from the very beginning: beside *lavamano* 'washstand', *guardaroba* 'wardrobe' and *salvadanaio* 'moneybox', attested as *salvadenaro* even from as early as the 14th century, we find for example *parasole* 'sunshade' (16th cent.), *guardastelle* (for *telescopio* 'telescope'). There are some locations as well: *battifolle* 'belfry', *battifreddo* 'defending tower', though these are presumably not formed directly by a WFR (*battifreddo* being remodelled on French *beffroi* 'belfry'). Finally, there are also some event nouns, though they are very rare (in line with Ricca's claim 2010: 251): *batticuore* 'heartbeat', *battibecco* 'squabble' (nouns surviving

---

[2] The textual database LIZ 4.0 contains also texts that cannot be considered representative of the literary language (which was to become the basis of the national language after the unification after 1861). The individual subcorpora have been created on the basis of precise criteria, see Štichauer (forthc.).

well into the present-day Italian), *lavacapo* 'brain-washing' (cf. for similar conclusions on Spanish Moyna 2011: 210-211).

Second, the important question of the output category (raised recently by Ricca 2005; 2010) can be answered only to a modest extent, as the corpus data from the 18th and 19th centuries show only one sure example: "(...) e ce lo mandò a bordo il vascello da guerra guardacoste, che sta sempre ancorato a quattro miglia di distanza." ('and it was sent to us on board of a coast-guard warship which is always anchored at the distance of four miles'), (Algarotti, *Viaggi di Russia*, 1764). It is true that *il Tommaseo-Bellini* registers several formations as nouns/adjectives (e.g. *mangiacarne* 'meat-eater'), but the adjectival function is hardly found in the corpora (so it can be concluded that it is probably a recent tendency yet to be investigated thoroughly).

Third, can any more precise diachronic tendencies (within the general quantitative increase) be established? Indeed, within the class of the most frequent leading bases (*porta-*, *mangia-*), it is the instrument meaning that emerges.

Finally, the question of the semantic instruction inherent in the WFR that creates VN compounds can only be reformulated as follows. The diachronic data suggest that whatever definition of the WFR one wishes to adopt (Aronoff 1976; Corbin 1987), it should be sufficiently general and should, as Ricca (2010: 254) puts it, "(...) leave the output category unspecified and basically context-dependent."

7. Conclusions

If we leave aside the precise formulation of the WFR in question, at least two conclusions can be drawn. First, the diachronic data presented so far confirm the view according to which all semantic types have always been present (being thus, in a way or another, inherent in the WFR). Second, the fact that instruments tend to prevail over the past two centuries (a fact to be ascribed to extralinguistic reality) suggests that this semantic output exploits the WFR, increasing thus the *profitability* of the always *available* word-formation pattern.

In this sense, the increasing productivity of the VN compounding is not to be taken as a phenomenon rooted in a deep structural change in the word-formation patterns, but should be considered instead as a *fortuitous exploitation* of an already present morphological procedure.

References

Aronoff, M. (1976) *Word formation in Generative Grammar*. Cambridge, MA: MIT Press.

Bisetto, A. (2004) Composizione con elementi italiani. In M. Grossmann & F. Rainer (Eds.), *La formazione delle parole in italiano*. Tübingen: Niemeyer, 33-51, 53-55.

Bauer, B. (2010) Word formation. In M. Maiden, J. C. Smith & A. Ledgeway (Eds.), *The Cambridge history of the Romance languages. Volume I. Structures.* Cambridge: Cambridge University Press, 532-563.

Corbin, D. (1987) *Morphologie dérivationnelle et structuration du lexique*. 2 voll. Tübingen: Niemeyer.

Dardano, M. (2008) *Costruire parole. La morfologia derivativa dell'italiano*. Bologna: il Mulino.

Gather, A. (2001) *Romanische Verb-Nomen Komposita. Wortbildung zwischen Lexikon, Morphologie und Syntax*. Tübingen: Narr.

il Tommaseo (2004) = *Dizionario della lingua italiana di Niccolò Tommaseo e Bernardo Bellini (1857-1879).* CD-ROM. Bologna: Zanichelli.

LIZ 4.0 (2001) = *Letteratura italiana Zanichelli*. CD-ROM. Edited by P. Stoppelli and E. Picchi. Bologna: Zanichelli

Marazzini, C. (2009) *L'ordine delle parole. Storia di vocabolari italiani.* Bologna: il Mulino.

Moyna, M. I. (2011) *Compound words in Spanish. Theory and history*. Amsterdam/Philadelphia: John Benjamins.

Radimský, J. (2006) *Les composés italiens actuels*. Paris: Cellule de Recherche en Linguistique.

Ricca, D. (2005) Al limite tra sintassi e morfologia: i composti aggettivali V-N nell'italiano contemporaneo. In M. Grossmann & A. M. Thornton (Eds.), *La formazione delle parole. Atti del XXVII Congresso*

Internazionale di Studi della Società di Linguistica Italiana. L'Aquila, 25-27 settembre 2003. Roma: Bulzoni, 465-486.

Ricca, D. (2010) Corpus data and theoretical implications. With special reference to Italian V-N compounds. In S. Scalise & I. Vogel (Eds.), *Cross-disciplinary issues in compounding*. Amsterdam/Philadelphia: John Benjamins, 237-254.

Scalise, S. (1994) *Morfologia*. Bologna: il Mulino.

Scalise, S. & A. Bisetto (2008) *La struttura delle parole*. Bologna: il Mulino.

Štichauer, P. (forthc.) *La formazione delle parole in diacronia. Studi di morfologia derivativa dell'italiano tra il Cinquecento e l'Ottocento*. Prague: Charles University Press.

Tollemache, F. (1945) *Le parole composte nella lingua italiana*. Roma: Rores.