

Idiomatic verb-clitic constructions: lexicalization and productivity

Francesca Masini
Alma Mater Studiorum – University of Bologna
francesca.masini@unibo.it

1. Introduction

The Idiomatic Verb-Clitic Construction (IVCC) is a special verbal complex found in Italian, as well as other Romance languages, that consists of a verb plus one or more pronominal clitics and displays a very specialized or totally idiomatic meaning. These constructions are known as *verbi procomplementari* ‘pro-complement verbs’ in the Italian lexicographic tradition (De Mauro 1999, GRADIT) and have received little attention until very recently (see Viviani 2006 and Russi 2008, 2011). The following examples contain Italian IVCCs based on the verb *prendere* ‘to take’ (1).

- (1) Italian
- a. *prendere-se-la*
take.INF-SI¹-CL.3.SG.F.OBJ.ACC
‘to become upset, to take offence’
 - b. *prendere-le*
take.INF-CL.3.PL.F.OBJ.ACC
‘to be beaten’
 - c. *prendere-ci*
take.INF-CL.LOC
‘to make the right choice, to guess/get it right’

Sometimes, the V+CL complex is part of a larger multiword expression, or complex predicate, as illustrated in (2).

- (2) Italian
- a. *metter-ce-la* *tutta*
put.INF-CL.LOC-CL.3.SG.F.OBJ.ACC all.F
‘to try one’s best’
 - b. *prendere-ci* *gusto*
take.INF-CL.LOC taste
‘to get to like something, to enjoy something’
 - c. *dar-se-la* *a gambe*
give.INF-SI-CL.3.SG.F.OBJ.ACC to legs
‘to run away (very quickly)’

IVCCs are also found in other Romance languages, such as Catalan and Spanish, as illustrated in (3)-(4). Espinal (2009), whose paper is based precisely on these languages, speaks of *inherent clitics* (or *i-clitics*).

¹ We use SI as an abbreviation for the polyfunctional pronominal clitic *si* in Italian (see Section 2.2).

- (3) Catalan (adapted from Espinal 2009: 1222)

ballar-la
take.INF-CL.3.SG.F.OBJ.ACC
'to suffer'

- (4) Spanish (adapted from Espinal 2009: 1224)

pasar lo bien
pass.INF CL.3.SG.NTR.ACC well
'to have a good time'

IVCCs in Romance languages are somehow similar to idioms that present the neuter weak pronoun *it* in English. Some of them are even equivalent to Italian IVCCs, as illustrated in (5).

- | | | | | |
|-----|-----------------------|---|-------------------------------|---------------------------|
| (5) | English | | Italian | |
| a. | <i>give it a rest</i> | → | <i>far-la</i> | <i>finita</i> |
| | | | make.INF-CL.3.SG.F.OBJ.ACC | ended.SG.F |
| | | | <i>smetter-la</i> | |
| | | | stop.INF-CL.3.SG.F.OBJ.ACC | |
| | | | <i>finir-la</i> | |
| | | | end.INF-CL.3.SG.F.OBJ.ACC | |
| b. | <i>take it easy</i> | → | <i>prender-se-la</i> | <i>comoda²</i> |
| | | | take-SI-CL.3.SG.F.OBJ.ACC | comfortable |
| c. | <i>make it</i> | → | <i>far-ce-la</i> | |
| | | | make-CL-LOC-CL.3.SG.F.OBJ.ACC | |

2. Italian IVCCs: what's special about them?

Italian IVCCs display an array of special semantic (Section 2.1) and morphosyntactic (Section 2.2) properties, as also noticed by Russi (2008) and Espinal (2009).

2.1 Semantic properties

Semantically speaking, the clitic is not a true anaphoric pronoun that refers to some other element in the discourse, but an element that is incorporated into the V and has a semantic impact on it. This impact may be different in both *strength* and *nature*.

As for strength, the IVCC and the clitic-less V may be more or less related semantically. In other words, the impact of the clitic goes from mild (6a) to fair (6b-c) to radical (6d), thus forming either fairly interpretable expressions (*encoding idioms* à la Makkai 1972) or totally idiomatic expressions (*decoding idioms*).

- | | | | |
|-----|-----------------|---|--|
| (6) | | | |
| a. | <i>smettere</i> | > | <i>smetter-la</i> |
| | stop.INF | | stop.INF-CL.3.SG.F.OBJ.ACC |
| | 'to stop' | | 'to stop (especially doing something annoying), to give it a rest' |
| b. | <i>sentire</i> | > | <i>sentir-se-la</i> |
| | feel.INF | | feel.INF-SI-CL.3.SG.F.OBJ.ACC |
| | 'to feel' | | 'to feel like doing something (difficult/unpleasant)' |

² This is just one of the many possible translations of *take it easy* into Italian.

- c. *sentire* > *sentir-se-lo*
 feel.INF feel.INF-SI-CL.3.SG.M.OBJ.ACC
 ‘to feel’ ‘to have a premonition about something’
- d. *correre* > *correre-ce-ne*
 run.INF run.INF-CL.LOC-CL.GEN
 ‘to run’ ‘there to be a great difference’

As for nature, the impact of the clitic may be either more lexical or more grammatical. A lexical kind of impact is exemplified by (7a), where the V+CL complex simply acquires a new, idiomatic meaning. In (7b-c), instead, the impact of the clitic is more grammatical, since the IVCC is close in meaning to the base V, but the clitic brings about new nuances. Russi (2008: 180-193) proposes that *la* in (7b) be analyzed as a subjectivity and *possibly* an aspectual marker, whereas *ci* in (7c) as a stativity and intransitive marker (Russi 2008: 169).

(7)

- a. *stare* > *star-ci*
 stay.INF stay.INF-CL.LOC
 ‘to stay’ ‘to agree to do something, be condescending’
- b. *finire* > *finir-la*
 end.INF end.INF-CL.3.SG.F.OBJ.ACC
 ‘to end’ ‘to stop (especially doing something annoying), to give it a rest’
- c. *vedere* > *veder-ci*
 see.INF see.CL.LOC
 ‘to see’ ‘to be able to see’

2.2 Morphosyntactic properties

As regards their position within the verbal complex, inherent clitics behave as normal clitic pronouns,³ as illustrated in (8)-(9). However, IVCCs are peculiar in more than one sense.

(8)

- a. *prender-le* [infinitival form]
 take.INF-CL.3.PL.F.OBJ.ACC
 ‘to be beaten’
- b. *le prendo* [finite form]
 CL.3.PL.F.OBJ.ACC take.1.SG
 ‘I’m beaten’

(9)

- a. *prender-se-la* (**prender-la-si*) [clitic cluster order]
 take.INF-SI-CL.3.SG.F.OBJ.ACC
 ‘to become upset, to take offence’
- b. *esser-se-la presa* (**prender-la-si*) [clitic climbing]
 be.INF-SI-CL.3.SG.F.OBJ.ACC taken.F
 ‘to have become upset, to have taken offence’

³ Cf. Cordin & Calabrese (1988: 589), quoted in Schwarze & Cimaglia (2010), for restrictions on clitic clusters in Italian.

First of all, the range of clitics used in IVCCs is rather limited compared to the full set of existing Italian clitics and clitic combinations. Indeed, only 6 pronominal clitics (including *si*) are used (Russi 2008; Jezek 2011), out of a total of 13 possibilities (Schwarze & Cimaglia 2010: Tab.1):

(10)

- a. *ci* → locative
- b. *la / le / lo* → accusative object (F.SG, F.PL, M.SG respectively)
- c. *ne* → genitive/partitive
- d. *si* → polyfunctional clitic pronoun and grammatical marker (see below)

Moreover, only 6 clitic clusters⁴ are attested:⁵

(11)

- a. *cela* (ci+la), *cene* (ci+ne)
- b. *sela* (si+la), *sele* (si+le), *selo* (si+lo), *sene* (si+ne)

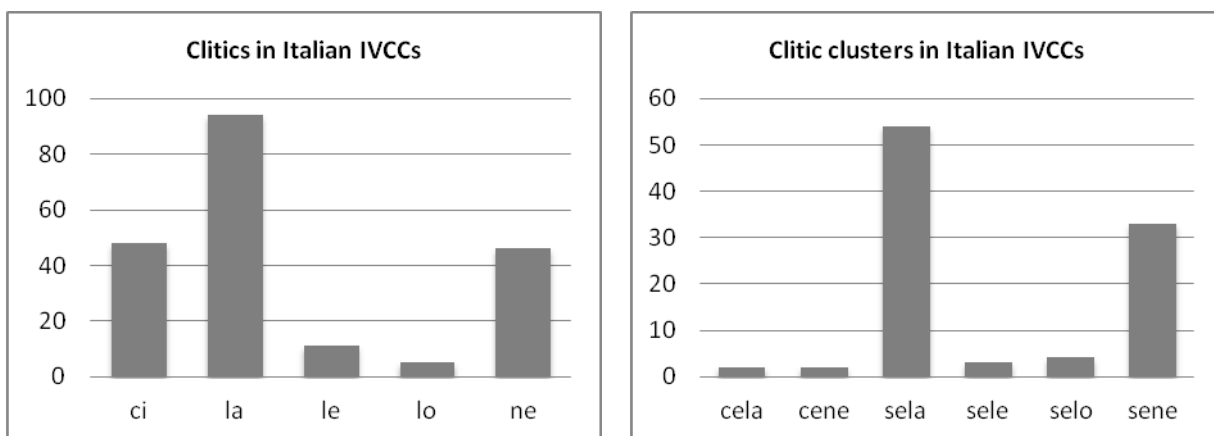
The precise role of *si* in IVCCs is not always easy to single out, given its many functions, e.g.: reflexive, reciprocal, impersonal, passivizing, aspectual, middle (cf. Jezek 2003, 2005). For instance, (12) clearly contains a reciprocal *si*, but the role of *si* is not clear in more idiomatic cases like (9) above.

(12) *dar-se-le*

dare-RECP-CL.3.PL.F.OBJ.ACC
‘to beat each other, to fight’

Apart from formal restrictions, we also have quantitative differences in usage among clitics and clitic clusters, which are illustrated in Figure 1. The data are elaborated from a database of approximately 200 (non-multiword) IVCCs that I collected from different sources (previous literature, dictionaries, corpora and my own intuition).

Figure 1: A quantitative sketch of clitics and clitic clusters in Italian IVCCs.



As shown in Figure 1, *la* (CL.3.SG.F.OBJ.ACC) is the most common clitic found in IVCCs, followed by *ci* (CL.LOC) and *ne* (CL.GEN), whereas *le* (CL.3.PL.F.OBJ.ACC) and *lo*

⁴ But see also *dar-glie-la su* (give.INF-CL.3.SG.M.DAT-CL.3.SG.F.ACC su) ‘to give up’.

⁵ In clitic clusters, *ci* and *si* appear in their allomorphic variants *ce* and *se*, respectively.

(CL.3.SG.M.OBJ.ACC) apparently have a marginal role. As for clusters, *sela* and *sene* are by far the most frequent.

Another special morphosyntactic feature of IVCCs is that inherent clitics are fixed. If we change the clitic, the IVCC (see e.g. (13a)) vanishes and we obtain a different construct, viz. either another IVCC (13b) or a “regular” clitic construction where the clitic is an anaphoric pronoun (13c).

(13)

- a. *prender-le*
take.INF-CL.3.PL.F.OBJ.ACC
‘to be beaten’
- b. *prender-ci*
take.INF-CL.LOC
‘to make the right choice, to get it right’
- c. *prender-la*
take.INF-CL.3.SG.F.OBJ.ACC
‘I take her (*to be beaten)’

Finally, inherent clitics are *mostly* obligatorily realized, i.e. in most cases they cannot be just dropped if we want to retain the same semantics (see (14)).

- (14) *prender-le* vs. *prendere*
take.INF-CL.3.PL.F.OBJ.ACC take.INF
‘to be beaten’ ‘to take (*to be beaten)’

Russi (2008: 180ff.) claims that in some IVCCs the clitic is not strictly obligatory (grammaticalized), because under certain circumstances we can drop it and end up with a grammatical sentence: see e.g. *finirla* vs. *finire* (7b) and *smetterla* vs. *smettere* (6a). Indeed, the meaning of the clitic-less V in these cases is very similar to the meaning of the IVCC (as Russi 2008: 180ff. claims, *la* acts as a sort of a subjectivity marker here), thus dropping the clitic may have no dramatic consequences on the semantics of the sentence in the “right” context, like (15):

(15)

- a. *Devi smettere di fumare*
‘You must quit smoking’
- b. *Devi smetterla di fumare*
‘You must stop smoking / with all this smoking’

However, even in these cases the clitic is not really optional, because it cannot be dropped *arbitrarily*, i.e. the two versions are not interchangeable in any context. See for instance (16a) and the agrammaticality of the clitic-less version (16b), or the different entailments of sentences (17a) and (17b): (17a) is an objective, neutral statement, whereas (17b) generally entails some kind of subjective disapproval about the fact that Sara doesn’t want to stop working.

(16)

- a. *È ora di finirla con questa storia*
‘It’s time to put an end to this stuff / I’ve had enough of this stuff’
- b. **È ora di finire con questa storia*

(17)

- a. *Sara non vuole smettere di lavorare*
'Sara doesn't want to stop working'
- b. *Sara non vuole smetterla di lavorare*
'Sara doesn't want to stop working (but she should!)

To sum up, IVCCs have a regular morphosyntactic behavior, but are morphologically restricted and have an unpredictable meaning. Even when the semantics of the IVCC is close to that of the base V, it's not predictable on the basis of its form. So, we are dealing with a case of morphosyntax-semantics mismatch.

3. Previous analyses and present issues

Espinal (2009) – whose analysis is mainly based on Spanish and Catalan – claims that inherent clitics are “arguments, whose binding-theoretic status is that of free variables incorporated into the verb stem” (Espinal 2009: 1257) and which denote “abstract semantic objects”. The mismatch is explained by saying that the constructions are regularly derived in the syntax, and then the clitic incorporates into the V at LF, thus saturating the argument of the V and giving rise to a complex predicate with a non-compositional semantics.

Russi (2008) – who carries out a descriptive, more diachronically oriented analysis of Italian IVCCs – claims that their emergence is the result of both grammaticalization (since the clitic becomes an obligatory marker, although not all IVCCs have reached a full stage of grammaticalization) and lexicalization (since the union of V+CL becomes a complex predicate with an idiomatic meaning). Russi (2008, 2011) also claims that the morphosyntactic status of inherent clitics (clitic vs. affix, inflectional vs. derivational) is all but clear, and that in fact inherent clitics seem to form a category of their own. Finally, Russi (2011: 400) states that “it should be desirable to determine how (or if) current theoretical frameworks are able to account for the phenomenon of incorporated clitics [...], especially without taking into consideration a diachronic dimension” (2011: 400).

In Section 4 I will address precisely this question: which kind of morphological theory would be suitable to account for the the synchronic and diachronic behavior of IVCCs?

In Section 5, instead, I will investigate in more detail another question: are IVCCs just a “dead” area of the Italian lexicon, or are they “alive” to any extent? In other words: is there any trace of productivity in the IVCC domain?

4. A constructionist account of IVCCs

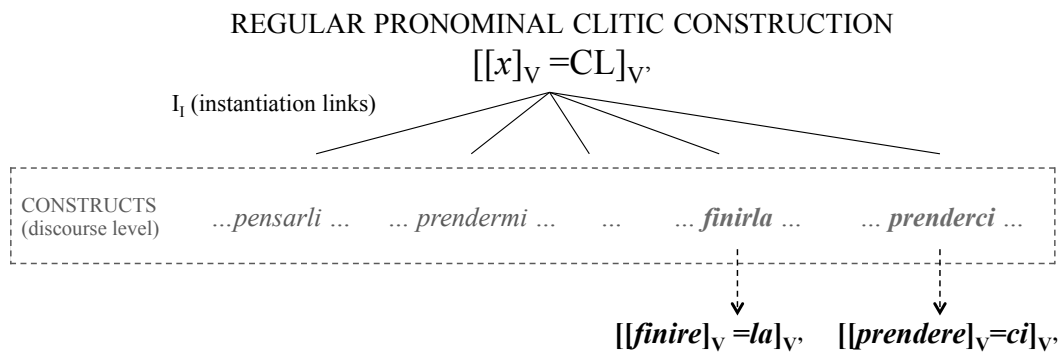
Most facts about IVCCs can find a suitable explanation by adopting a theory that: i) allows in principle for mismatches between form and meaning; ii) accounts for both regularities and idiosyncrasies; iii) fruitfully integrates synchrony and diachrony, formalization and usage. I propose that one such model is Construction Morphology (CxM; Booij 2010), which is framed within the more general theory of Construction Grammar (CxG; Hoffmann & Trousdale 2013).

As is well-known, in a constructionist perspective morphology and syntax are not independent, sequential components. The basic unit of analysis is the construction, i.e. a conventionalized association of a form and a meaning that may have different degrees of schematicity and complexity. Two basic mechanisms are used to form new expressions: instantiation, i.e. the creation of more specific instances from more general constructions, and unification, i.e. the merging of two or more constructions together. Constructions are

organized into a hierarchical lexicon and may be connected to one another via inheritance links of different sorts. The mechanism of default inheritance accommodates both regular and idiosyncratic facts, including form-meaning mismatches. Hybrid morphosyntactic status is therefore an expected, possible output, rather than an exception. Despite having a formalizing vocation, CxM is sensitive to (and equipped to account for) usage and diachronic facts, along the lines of CxG.

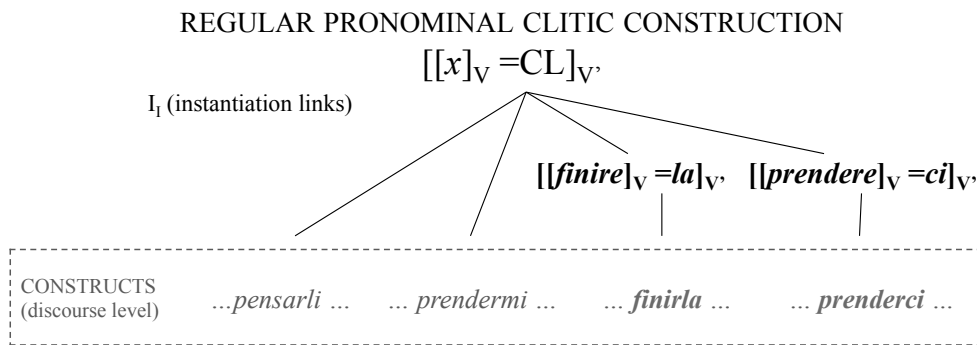
Let us see how some of the IVCC facts can be analyzed within the framework of CxM. Figure 2 describes the diachrony of IVCCs: some instances (constructs) that originate from the regular clitic construction (e.g. *finirla* ‘to give it a rest’ and *prenderci* ‘to guess it right’) become more and more entrenched and acquire new meanings, by virtue of their occurrence in specific contexts (cf. Russi 2008). As a consequence, new constructions are created and stored in the constructicon.

Figure 2: Diachronic picture for IVCCs.



These new constructions (see Figure 3) are lexically fully specified (i.e., the components are fixed) and semantically idiomatic, thus overriding the function of the regular pronominal clitic construction. At the same time, they retain some link to the regular pronominal clitic construction, thus inheriting its morphosyntactic properties (remember that IVCC have a regular morphosyntactic behavior).

Figure 3: Synchronic picture for IVCCs



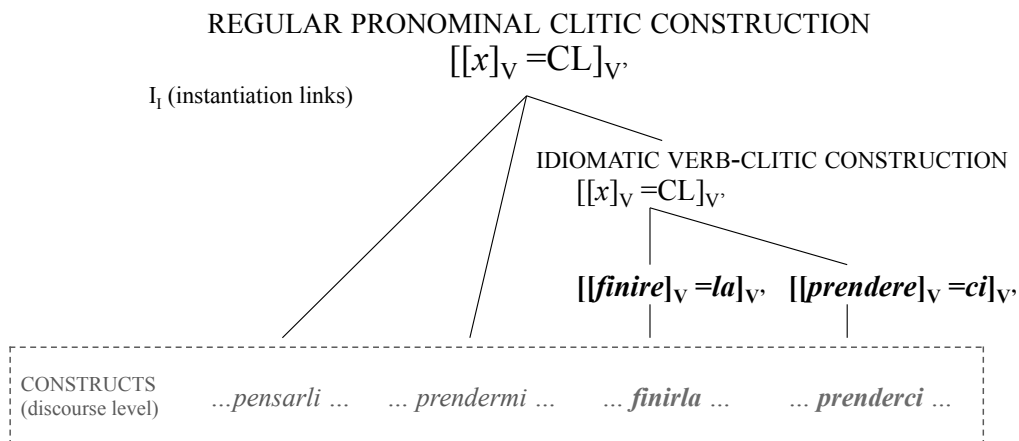
We therefore obtain a mismatch, like the one illustrated in Figure 4.

Figure 4: Mismatch between form and meaning in IVCCs.

| REGULAR PRONOMINAL CLITIC CONSTRUCTION | | | IDIOMATIC VERB-CLITIC CONSTRUCTIONS | |
|--|--------------------------------------|---|-------------------------------------|--|
| Form | $[[x]_{vk} = CL_i]_V$ | → | Form | $[[prendere]_V = ci]_V$ |
| Meaning | $\langle PRED_k (Arg Adj)_i \rangle$ | ↔ | Meaning | $\langle \text{to guess it right} \rangle$ |

An alternative synchronic analysis of IVCCs is given in Figure 5. The difference with Figure 3 lies in the positing of an intermediate abstract IVCC between the regular pronominal clitic construction and the single lexically specified IVCCs. This solution has the advantage of generalizing over the several instances of IVCCs existing in Italian. However, the generalization is pretty loose: the clitics have different roles and functions in the various IVCCs, which normally have a history of their own (but see Section 5). Therefore, the only information we can inscribe into the abstract IVCC is that the clitic is not anaphoric and the meaning is non-compositional. In other words, the abstract IVCC, if posited, should not be taken as a template for the formation of new IVCCs, but as a static generalization.

Figure 5: Alternative synchronic picture for IVCCs: positing an abstract IVCC



In conclusion of this section, two other cases are discussed: IVCCs with *si* (cf. (11)-(12), Section 2.2) and IVCCs within multiword expressions (cf. (2), Section 1).

As for the former, cases like (12), where we have a synchronically identifiable *si* within an IVCC, could be analyzed as cases of unification between the *si* construction (in this case, the reciprocal *si* construction) and the regular pronominal clitic construction, whose output is then lexicalized. Inheriting from the reciprocal *si* construction, too, guarantees that the semantic contribution of *si* is correctly interpreted within the IVCC. This analysis obviously does not apply to cases where the contribution of *si* within the IVCC is not synchronically recognizable.

As for the latter case, the fact that IVCCs frequently occur within larger multiword expressions is not unexpected in a constructionist perspective. Since morphology and syntax are not independent, sequential components, but levels of analysis that closely interact with each other, multiword expressions are expected to occur, alongside morphological words. Therefore, complex predicates are lexical constructions with their own properties and argument structure, just like non-multiword verbs, and as such they are equally eligible to develop into an IVCC.

5. From lexicalization to productivity

At first sight, there are no signs of true productivity in the domain of IVCCs, which seem to be the result of unpredictable diachronic processes. We can only find traces of analogy. For instance, in several cases, IVCCs with different bases and same clitic (18) or same base and different clitics (19) have the very same meaning. This seems to hint that some IVCCs are created along the lines of already existing ones via analogy.

(18)

- | | | | |
|----|--|-----|---|
| a. | <i>finir-la</i> end.INF-CL.3.SG.F.OBJ.ACC 'to stop, to give it a rest' | vs. | <i>smetter-la</i> stop.INF-CL.3.SG.F.OBJ.ACC |
| b. | <i>impiegare-ci</i> employ-CL.LOC 'to take (time to do something)' | vs. | <i>metter-ci</i> put-CL.LOC |

(19)

- | | | | |
|----|---|-----|--|
| a. | <i>cercar-se-la</i> search-SI-CL.3.SG.F.OBJ.ACC 'to deserve something, to ask for it' | vs. | <i>cercar-se-le</i> search-SI-CL.3.PL.F.OBJ.ACC |
| b. | <i>azzeccar-la</i> guess_right-CL.3.SG.F.OBJ.ACC 'to get it right' | vs. | <i>azzeccar-ci</i> guess_right-CL.LOC |

Take for instance the *finirla* vs. *smetterla* couple. As Russi (2008: 180ff.) notes, whereas *finire* 'to end' and *smettere* 'to stop' are not complete synonyms (20), *finirla* and *smetterla* actually are (21).

(20)

- | | | | |
|----|---|-----|---|
| a. | <i>finire un libro</i> 'to finish a book' | vs. | * <i>smettere un libro</i> |
| b. | <i>finire di fumare</i> 'to finish smoking (*to quit smoking)' | vs. | <i>smettere di fumare</i> 'to finish / quit smoking' |

(21)

- | | | | |
|----|--|---|-----------------------------|
| a. | <i>finiscila!</i> 'Stop it! / Give it a break!' | = | <i>smettila!</i> |
| b. | <i>finiscila di parlare!</i> 'Stop that talking!' | = | <i>smettila di parlare!</i> |

If we check when these items were coined (cf. GRADIT), we find out that *finirla* dates back to the 16th c., whereas *smetterla* is much more recent (19th c.). Therefore, it is reasonable to hypothesize that *smetterla* was coined by analogy with *finirla*.

However, there is one case in which things seem to have taken a different course.

5.1 A case study: *andarsene*

One of the most frequent clitic clusters in IVCCs is *sene* (cf. Figure 1), formed by *si* (in its allomorphic shape *se*) and the genitive/partitive clitic *ne*. One of the most often cited

examples of lexicalized IVCCs with *sene* is *andarsene*, which is very frequent in contemporary Italian and whose meaning is usually that in (22):

(22)

- a. *andar-se-ne*
go.INF-SI-CL.GEN
'to go away, to leave'
- b. [...] *salutò e se ne andò*
'He said goodbye and left'

In this case we should probably speak of *sene* as an inseparable clitic cluster, since **andarsi* would be agrammatical in contemporary Italian and *andarne* is an IVCC with a completely different meaning ('be at stake').⁶

At the same time, *andarsene* may also have a different use, which is definitely more marginal and usually requires an adverbial. See the examples in (23):⁷

(23)

- a. *avendo un paio di ore libere, me ne sono andata a spasso sotto una pioggia fitta e insistente*
'having a couple of free hours, I went walking around [somehow emotionally involved] under a heavy and persistent rain'
- b. [...] *se ne andava in giro con il sorriso stampato sulle labbra*
'(she) was going around [somehow emotionally involved] with a smile fixed on her face'

In these cases, there is no "away" semantics. Rather, *sene* has a sort of "emphatic" function, which results in the subject (or some other participant in the event or in the enunciation) being somehow affected by the event denoted by the verb. Since the overall effect of *sene* on the verb is not easily translatable, for convenience it will be marked with [+EMPHASIS] / [+E] hereafter.

What is interesting is that, as cursorily noted by Russi (2008) and Masini (2008, 2012), this "emphatic" *sene* seems to occur not only with *andare* 'to go', but also with other motion verbs, such as directed motion verbs (24) and locative/stative verbs (25)-(26).⁸

(24)

- a. *uscir-sene* exit.INF-SENE 'to exit, to go out [+EMPHASIS]'
- b. *Mentre Sepp, Fritz e Valz se ne uscivano piuttosto depressi dalla riunione [...]*
'While Sepp, Fritz and Valz left the meeting [+E] in a quite depressing mood...'

(25)

- a. *star-sene* exit.INF-SENE 'to stay, to remain [+EMPHASIS]'
- b. *La incontravo a colazione e raramente a pranzo e cena, perchè lei se ne stava in camera*
'I met her during breakfast and, rarely, dinner and supper, since she would stay [+E] in her room'

⁶ However, the presence of *si* might play a covert role here, since its presence triggers the BE auxiliary (*essere*), thus making the whole construction more compatible with telic situations.

⁷ All examples hereafter are taken from the *ITenTen10* web corpus, available on the SketchEngine. Cf. also Section 5.2.

⁸ For simplicity, I will use the term "motion verbs" or "verbs of motion" for verbs conveying both motion (e.g. *go*, *come*) and location (e.g. *stay*).

(26)

- a. *restar-sene* stay.INF-SENE ‘to remain [+EMPHASIS]’
 b. *Andava ancora alle feste, ma se ne restava in disparte*
 ‘(He) still went to parties, but he would stay [+E] on his own’

For the sake of completeness, all these examples would be grammatical even without *sene*:

(27) *(me ne) sono andata a spasso sotto una pioggia fitta e insistente*
 ‘I went walking around under a heavy and persistent rain’

(28) *lei (se ne) stava in camera*
 ‘she stayed in her room’

However, the semantics would slightly change: the *sene* event is not presented in an objective, neutral, detached way ([-EMPHASIS]); rather, the subject (or another participant) is somehow affected ([+EMPHASIS]). In (27), for instance, it would be *inappropriate* to use the *sene* variant if the participant were not emotionally involved (which is not entailed by the corresponding example without *sene*, cf. (23a)). Similarly, in (28) it would be *inappropriate* to use the *sene* variant if the participant were staying in her room all day long unwillingly (which, again, is not entailed by the corresponding example without *sene*, cf. (25b)).

5.2 A corpus search for motion verbs with *sene*

In order to confirm the observation that *sene* can in fact be used with verbs of motion (and to what extent), I carried out a corpus search, based on the 3.1 billion tokens web corpus *ItTenTen10*, available on the SketchEngine (www.sketchengine.co.uk).

It turned out that *sene* can be attached to a large array of verbs of motion, albeit their token frequency is pretty low. Clearly, it is a marginal construction, which may even sound awkward to some native speakers, but whose use is undeniably attested. The *sene* construction appears to be compatible with virtually all verbs of motion, namely:

- PATH VERBS OF MOTION: *andare* ‘to go’, *venire* ‘to come’, *partire* ‘to leave’, *arrivare* ‘to arrive’, *fuggire* ‘to escape’, *tornare* ‘to come back’, *ritornare* ‘to come back’, *uscire* ‘to exit’, *entrare* ‘to enter’, *salire* ‘to go up, climb’, *scendere* ‘to go down, descend’, etc.
- MANNER VERBS OF MOTION: *correre* ‘to run’, *cadere* ‘to fall’, *saltare* ‘to jump’, *volare* ‘to fly’, *scivolare* ‘to slide, slip’, *camminare* ‘to walk’, *gironzolare* ‘to wander around’, *passaggiare* ‘to stroll’, *rotolare* ‘to roll’, etc.
- LOCATIVE/STATIVE VERBS: *restare* ‘to stay’, *stare* ‘to stay’, *giacere* ‘to lie’, etc.
- PHRASAL VERBS with motion/locative meaning (cf. Masini 2005; Iacobini & Masini 2007).

For each category, some examples are given below (cf. (29)-(32)).

(29) PATH VERBS OF MOTION

- a. *Addio Bocca di rosa con te se ne parte la primavera* (F. De André song)
 ‘Goodbye Bocca di rosa, the spring **leaves** [+E] with you’
 b. [...] *egli uccise il mostro. Poi se ne fuggì con Arianna*
 ‘he killed the monster. Then he **fled** [+E] with Ariadne’
 c. *dopo qualche strillo se ne sale in camera incazzata nera*
 ‘after some screaming, (she) **went up** [+E] to her room, really pissed off’

- d. *Ecco a voi [...] Logan che se ne entra al ristorante con il suo maritino*
‘And there you have [...] Logan **entering** [+E] the restaurant with her dear husband’
- e. *In quel momento se ne arriva la baby sitter con i tre pargoli*
‘At that point the babysitter **arrives** [+E] with the three kids’
- f. *Impostata la velocità di 70 Km/h l'aereo se ne scende tranquillo*
‘Once a 70 km/h speed has been set, the airplane **goes down** [+E] easily’

(30) MANNER VERBS OF MOTION

- a. *tutto il materiale [...] se ne rotolava bel bello di qua e di là*
‘all the material [...] was **rolling** [+E] here and there’
- b. *[...] vidi il teppistello corrersene in skateboard lungo la strada*
‘[...] I saw the little thug **running** [+E] on a skateboard along the street’
- c. *come i denti da latte che se ne cadono da soli*
‘like milk teeth that **fall out** [+E] spontaneously’
- d. *Un ragazzo se ne passeggia nel giardino*
‘A boy **wanders** [+E] in the garden’
- e. *[...] mentre l'Uomo Ragno se ne saltava da una parte all'altra*
‘[...] while Spiderman was **jumping** [+E] from side to side’
- f. *Cominciai a invidiare il dottore che [...] se ne camminava nella fresca ombra degli alberi*
‘I started envying the doctor who [...] was **walking** [+E] in the refreshing shadow of the trees’

(31) LOCATIVE/STATIVE VERBS

- a. *[...] il celebre scienziato che se ne giaceva tutto solo*
‘[...] the famous scientist who was **lying** [+E] there all alone’

(32) PHRASAL VERBS with motion/locative meaning

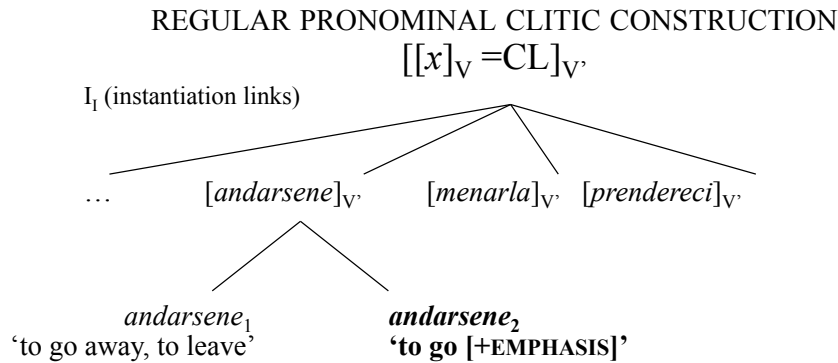
- a. *Paula si alzava di notte e se ne scendeva giù in cucina*
‘Paula used to get up at night and **go down** [+E] to the kitchen’
- b. *La farfalla, tutta rossa per la vergogna, se ne volò via*
‘The butterfly, crimson with embarrassment, **flew away** [+E]’
- c. *[...] il mantello caduto a Plasson nella fretta di corrersene via*
‘the cape that Plasson dropped while he **ran away** [+E] in a rush’
- d. *[...] la luna che se ne viene su*
‘the moon that **rises** (lit. comes up) [+E]’

5.3 Constructionalization

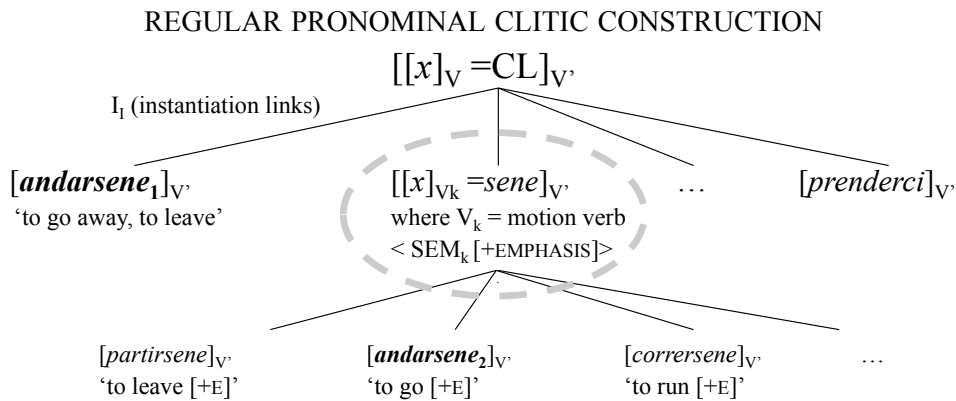
The IVCCs with *sene* in Section 5.2 are not encoded in dictionaries and seem to be recent innovations (20th c.), with the exception of the more widely used *sene* verbs: the verb *andarsene* is the oldest attested (1294, cf. GRADIT), followed by *ritornarsene* (17th c., cf. GRADIT) and *tornarsene* (1879, cf. GRADIT).

Therefore, it seems we are dealing with an exemplar-based constructionalization process, namely the emergence of a (semi-specified) construction with some degree of productivity that originates from one or more lexicalized instances (specific IVCCs) with very high token frequency.

This process is outlined in Figure 6 and Figure 7.

Figure 6: Diachronic picture for the *sene* construction

We first see the instantiation of the lexically specified construction *andarsene*₂ ‘to go [+EMPHASIS]’ (Figure 6): due to the entrenchment of this exemplar, new expressions are created by analogy (e.g. *ritornarsene*, *tornarsene*). Eventually, a new semi-specified construction emerges in the construction with the following properties (see Figure 7): a (fixed) *sene* clitic cluster, an open slot for motion verbs, and an emphatic meaning. Since the V slot is open, new fully specified expressions with *sene* are instantiated.

Figure 7: Synchronic picture for the *sene* construction.

The emerging *sene* construction is limited by input constraints. First, only intransitive verbs of motions seem to be allowed. Transitive motion verbs such as *raggiungere* ‘to reach’ don’t occur with *sene*:

- (33) *raggiungere* (qualcosa/qualcuno) > **raggiunger-sene*
‘to reach (something/someone)’ reach.INF-SENE

Also verbs already containing an inherent *si* such as *arrampicarsi* ‘to climb (up)’ are excluded, probably due to a formal clash:

- (34) *arrampicarsi* > **arrampicarsi-sene*
‘to climb’ climb.INF-SI-SENE

5.4 Beyond motion: expanding the *sene* construction

At a closer look, the *sene* construction seems to go beyond the domain of motion/location.

First of all, the emphatic *sene* is also found with motion verbs used *metaphorically*, as shown in (35)-(36), and with some non-motion verbs, as illustrated in (37)-(38).

(35)

- a. ***cadere a pezzi*** (lit. fall to pieces) ‘to fall to pieces, to crumble away’
- b. *E tutto ormai se ne cade a pezzi*
‘And everything is falling into pieces [+E] by now’

(36)

- a. ***volare via*** (lit. fly away) ‘to pass quickly’
- b. *e così quel breve pomeriggio se ne volò via*
‘and so that brief afternoon passed quickly [+E]’

(37)

- a. ***vivere*** ‘to live’
- b. *Il cane se ne viveva per suo conto*
‘The dog would live [+E] on its own’

(38)

- a. ***nascere*** ‘to be born’
- b. [...] *Leonard Bernstein, che se ne nacque povero in un posto infame*
‘Leonard Bernstein, who had been born [+E] poor in a miserable place’

Also, *sene* is found with *stare* ‘stay’ / *restare* ‘stay’ / *andare* ‘go’ when these are not used as motion verbs, but as functional verbs in aspectual periphrases (see (39)-(41)). The *sene* construction is very common with *stare* ‘stay’ / *restare* ‘stay’.

(39)

- a. ***(re)stare + a + V_{INFINITIVE}***
- b. *era il suo giorno libero, poteva benissimo starsene a dormire*
‘that was his day off, he could well keep on sleeping [+E]’

(40)

- a. ***andare + V_{GERUND}***
- b. *Un attimo dopo Ellis se ne andava camminando nel buio*
‘A minute later Ellis would be walking around [+E] in the dark’

(41)

- a. ***andare + a + V_{INFINITIVE}***
- b. *nessuno ha voglia di andarsene a dormire*
‘nobody would want to go to sleep [+E]’

The occurrence of *sene* in these aspectual constructions (where it is associated with all sorts of verbs) may contribute / have contributed spreading its use beyond the domain of motion. However, the “intransitivity” constraint still appears to hold.

5.5 A broader picture

What is the function of the *sene* construction within the Italian system? In this last section I would like to explore this question and put forward a first hypothesis.

The semantic job done by *sene* is quite similar to what *si* does in another “emphatic” pronominal verb construction in Italian, namely the so-called Transitive Middle Construction⁹ (Fried & Masini 2011; Masini 2008, 2012), where *si* marks the affectedness of the subject. See the following contrasts between *mangiare* and *mangiarsi* on the one hand (42) and *bere* and *bersi* on the other (43):

(42)

- a. *Sara mangia un gelato*
‘Sara eats an icecream’
- b. *Sara **si** mangia un gelato*
‘Sara eats “herself” an icecream’ → eat[+EMPHASIS]

(43)

- a. *Bevo una birra*
‘I drink a beer’
- b. ***Mi** bevo una birra*
‘I drink “myself” a beer’ → drink[+EMPHASIS]

The Middle Transitive *si* is also found in combination (unified) with inherent clitics:

(44) *ber-se-la*

drink.INF-SI-CL.3.SG.F.OBJ.ACC

‘to believe naively that something is true, to “buy” it’

As discussed in Masini (2012), the Middle Transitive Construction admits only prototypically transitive verbs, i.e. transitive verbs used with an Agent subject and an overtly expressed object (the null object alternation is not allowed). Within the limits of this restriction, it is very productive in contemporary (colloquial) Italian.

The *sene* construction shares the emphatic function with the Middle Transitive Construction, which results in an impact on the subject or another participant. However, as we pointed out, the emphatic *sene* only adds up to intransitive verbs, and a limited (albeit possibly increasing) portion of them. Therefore, the *sene* construction could be seen as a sort of Middle *Intransitive* Construction, with a more restricted use than the corresponding Middle Transitive Construction at present.

If we accept this hypothesis, then the *sene* construction is quite different from most other IVCCs:

- i) unlike the *sene* construction, generally IVCCs are lexically specified and basically unproductive;
- ii) unlike the emphatic *sene*, most inherent clitics have a “lexical” impact on the V and are obligatorily realized;
- iii) clitics with more “grammatical” meanings (like *la* in *finir-la* ‘to give it a rest’) are more similar to *sene*, but still different because basically unproductive.

⁹ The term *middle* here is used in the sense of Kemmer (1993).

6. Concluding remarks

The case study discussed in this paper illustrates two general points: the crucial interplay between representation and usage in the building of grammar, and the importance of using flexible theoretical tools to capture hybrid situations.

As for the first point, we can see that all the effects of usage on grammar mentioned by Bybee (2006: 719) are illustrated by IVCCs:

- i) low levels of repetition lead to conventionalization only (as in prefabs and idioms);
- ii) higher levels of repetition can lead to the establishment of a new construction with its own categories;
- iii) extreme high frequency leads to the grammaticization of the new construction, the creation of grammatical morphemes, and changes in constituency.

As for the second point, a constructionist view of the language architecture allows to capture and integrate various synchronic and diachronic aspects of IVCCs, most importantly the mismatch between their morphosyntax and their semantics, their emergence from concrete exemplars, and the different degrees of idiosyncrasy/regularity that characterizes the behavior of the various IVCCs. Indeed, at the present stage of development, we have a hybrid situation that includes: IVCCs with a regular behavior but unpredictable semantics that are represented as lexically specified constructions¹⁰; and a lexically semi-specified construction (called here Middle Intransitive Construction) that displays some degree of productivity and is set apart from other IVCCs. Although it seems flimsy to hypothesize the existence of a maximally abstract IVCC that acts as a template for the formation of new IVCCs, such a construction may still be present in the constructicon as a “closed” construction, namely as a generalization over a closed set of instances.

References

- Booij, G. (2010) *Construction morphology*. Oxford: Oxford University Press.
- Bybee, J. (2006) From usage to grammar: the mind’s response to repetition. *Language* 82, 711-733.
- Cordin, P. & A. Calabrese (1988) I pronomi personali. In L. Renzi et al. (Eds.), *Grande grammatica italiana di consultazione*, vol. 1. Bologna: Il Mulino, 535-592.
- De Mauro, T. (1999) *Introduzione*. In GRADIT 1999-2007. Turin: UTET.
- Espinal, M. T. (2009) Clitic incorporation and abstract semantic objects in idiomatic constructions. *Linguistics* 47(6), 1221-1271.
- Fried, M. & F. Masini (2011) *Ethical datives and related constructions*. Paper presented at the 44th Annual Meeting of Societas Linguistica Europaea (SLE), Logroño, 8-11 September 2011.
- GRADIT = Il grande dizionario italiano dell’uso, ed. by Tullio De Mauro, 1999. Turin: UTET.
- Hoffmann, T. & G. Trousdale (Eds.) (2013) *The Oxford handbook of Construction Grammar*. Oxford: Oxford University Press
- Jezek, E. (2003) *Classi di verbi tra semantica e sintassi*. Pisa: ETS.
- Jezek, E. (2005) Interazioni tra aspetto e diatesi nei verbi pronominali italiani. *Studi di grammatica italiana* 23, 239-281.
- Jezek, E. (2011) Verbi pronominali. In *Enciclopedia dell’Italiano*, vol. 2. Roma: Istituto della Enciclopedia Italiana.
- Kemmer, S. (2001) *The middle voice*. Amsterdam/Philadelphia: John Benjamins.
- Makkai, A. (1972) *Idiom structure in English*. The Hague: Mouton.
- Masini, F. (2008) *Costruzioni verbo-pronominali “intensive” in italiano*. Paper presented at the “XLII Congresso della Società di Linguistica Italiana” (SLI42), Pisa, 25-27 September 2008.
- Masini, F. (2012) *Costruzioni verbo-pronominali “intensive” in italiano*. In P. M. Bertinetto et al. (Eds.), *Language and the brain – Semantics* (Proceedings SLI42), Vol. 2. Roma: Bulzoni.

¹⁰ Possibly, other concrete IVCCs are in the making but not yet constructionalized. I leave this question for future research.

- Monachesi, P. (1996) *A grammar of Italian clitics*. PhD dissertation, University of Tilburg.
- Russi, C. (2008) *Italian clitics. An empirical study*. Berlin/New York: Mouton de Gruyter.
- Russi, C. (2011) Clitics of Italian *verbi procomplementari*: what are they? In M. Maiden et al. (Eds.), *Morphological autonomy: perspectives from Romance inflectional morphology*. Oxford: Oxford University Press, 382-400.
- Schwarze, Ch. & R. Cimaglia (2010) Clitici. In *Enciclopedia dell'Italiano*, vol. 1. Roma: Istituto della Enciclopedia Italiana.
- Viviani, A. (2006) I verbi procomplementari tra grammatica e lessicografia. *Studi di grammatica italiana* 25, 255-321.