

Italian Verbs of Manner of Motion at the Syntax-Semantics Interface: a Distributional Analysis

Gabriella Lapesa Alessandro Lenci

University of Osnabrück University of Pisa

glapesa@uni-osnabrueck.de alessandro.lenci@ling.unipi.it

In this work, we carried out a quantitative analysis of the distributional correlates of Italian Manner of Motion (henceforth, MoM) verbs, to investigate the interaction between verb meaning and syntactic constructions in the linguistic encoding of motion events. Theoretically grounded in Levin’s take on argument realization¹ and methodologically based on the Distributional Semantics framework (Miller & Charles, 1991), our research relies on the assumption that the syntactic behavior of verbs is semantically determined, and, as a consequence, that verbs showing similar syntactic patterns must share some meaning components.

1 The Traditional Classification and its Problems

The typological classification of languages with respect to the lexicalization of Motion events is a widely investigated topic. The conceptual structure of motion events proposed by Talmy (1985, 1991) is at the basis of the classic opposition between Verb-framed and Satellite-framed languages. Verb-framed languages (e.g., Romance languages) encode the PATH element in the main verb root (example 1.a), Satellite-framed languages (e.g., Germanic languages) encode the PATH element in a directional complement (example 1.b):

- (1) a. *John ran into the room.*

[John]_{FIGURE} [ran]_{MANNER} [**into the room**]_{PATH}

- b. *Mario é entrato nella stanza correndo.*

[Mario]_{FIGURE} [**entered**]_{PATH} [the room]_{GOAL} [running]_{MANNER}

In the taxonomy proposed by Talmy, Italian is classified among the Verb-framed languages. However, data show that this classification is inappropriate with respect to (at least) three points. First of all, the size of the MoM lexicon is quite relevant. According to the dictionary-based classification carried out by Iacobini (2010),

¹ “The assumption that the syntactic behavior of verbs is semantically determined gives rise to a powerful technique for investigating verb meaning that can be exploited in the development of a theory of lexical knowledge” (Levin, 1993:14).

Italian MoM lexicon is richer than those of the other Verb-Framed languages. Iacobini's classification, shown in Appendix A, will constitute the starting point of the analysis presented in the poster. Second, the frequent use of MoM verbs in combination with directed motion adverbs (e.g. *saltare fuori*, “to jump out”; *scappare via*, “to run away”; *volare via* “to fly away”) can be considered as a cue of the emergence of verb-particle constructions (Iacobini and Masini, 2006), a strategy that is typical of Satellite-Framed languages (cfr. phrasal verbs). Third, some verbs show a tendency to merge MANNER and PATH (i.e. *eclissarsi*, “to disappear”; *gettarsi*, “to hurl oneself”; *irrompere*, “to burst into”; *sfuggire*, “to escape”).

1.1 New Proposals and the Need for Empirical Evidence

To account for the exceptions of Talmy's dichotomy, Slobin (2004) proposed to introduce a third group of languages, defined *equipollently framed*, in which manner and path verbs have equal status in the encoding of motion events. However, this tripartite classification still faces problems related to the identification of satellites and to the difficulties of defining what does it mean for a language to give “equal status” to manner and path verbs. More recently, Slobin (2008, 2009) suggested to shift the focus of the debate from typology to use, reformulating the typological opposition in terms of preference for Path-in-Verb (PIV) or Path-in-Non-Verb (PIN) constructions. Shifting the focus to use and preference to constructions turns the typologic dychotomy into a cline, in which “problematic” languages like Italian (or Serbo-Croatian) are more easily accommodated: this is the first main advantage of this proposal. In addition to that, the PIN/PIV hypothesis emphasizes the role of the structure of the motion event in determining the choice for lexicalization patterns. First results (Slobin, personal communication) suggest that the need of encoding a boundary-crossing event (e.g: *entrare in una stanza correndo*, “to run into a room”) triggers the use of PIN constructions also in languages that would have shown a preference for PIV constructions otherwise². Moreover, in languages with an overall preference for PIV constructions, boundary crossing events are “compatible only with manners of movement that are ballistic, high force dynamic” (Slobin, 2008).

2 Methodology

We analysed the way semantic classes of MoM distributionally behave in a large reference corpus. Distributional data about verb argument structure have been automatically extracted from the *La Repubblica* corpus (Baroni et al. 2004) through the LexIt framework (Lenci, forthcoming). We used stochastic association measures (Evert, 2008) traditionally applied to the study of word collocations to evaluate the strength of the correlation between verbs and syntactic frames, argument slots and lexical fillers or semantic classes. LexIt is a publicly available³ database which

2 For example, according to the PIN/PIV hypothesis, a speaker of Italian, when seeing John rapidly getting into the room should prefer the Path-in-Non-Verb construction (*John è corso nella stanza*, “John run into the room”) over the Path-in-Verb construction (*John è entrato nella stanza correndo*, “John entered the room by running”).

3 At the address <http://sesia.humnet.unipi.it/lexit/>

contains distributional information concerning 3933 Italian verbs: each verb is described in terms of a syntactic profile (subcategorization frames) and a semantic profile. The semantic profile is further articulated into lexical sets (Hanks and Pustejovsky, 1995) and selectional preferences (we implemented a variation of the algorithm described by Schulte-Im-Walde, 2006).

3 Results and Analysis

In this section we briefly present the distributional evidence found in *La Repubblica* for the emergence of verb particle constructions (section 4.1) and an analysis aiming at characterizing the behavior of MoM in interaction with boundary-crossing predicates (section 4.2). To characterize the distributional behavior of MoM, all the verbs belonging to a class were lumped together, into a unique label representing each MoM class. For example, in the tables presented in the following subsections, the data concerning the “Basic Level” class have been computed by summing the frequency of syntactic frames, slots and fillers of the three corresponding verbs *camminare* “walk”, *correre* “to run” and *incamminarsi* “to make one’s way”.

3.1 Emergence of Verb-particle Constructions

The topic of the emergence of verb-particle constructions in Italian has been tackled by several studies which investigated their locative, idiomatic and aspectual meaning nuances. In a dictionary-based study, Iacobini and Masini (2006) noticed that manner verbs constitute an optimal base for the creation of new verb-particle constructions, because they may require a specification of Path (e.g.: *saltare fuori*, “to jump out”) or an endpoint (e.g.: *grattare via*, “to scrape away”). This intuition found its distributional correlate in the strong association between the group of particles defined *telic* by Iacobini and Masini (*via*, “away”; *su*, “up”; *giù*, “down”; *fuori*, “out”)⁴ and some classes of MoM verbs, as shown in Table 1.

MoM Semantic Class	Fuori	Su	Giù	Via	Total
Rapid movement toward goal	937	39	7262	6427	14665
Punctual, repeteable movement	10320	127	253	0	10700
Rapid movement away from source	150	0	0	3325	3475
Rapid movement	22	9	342	1908	2281
Frictionless, silent movement	12	1	361	1395	1769
Labored progress	110	6	141	822	1079
Basic level	1	132	25	62	219
Relaxed walking	0	31	0	0	31
Impaired walking	0	15	0	0	15
Types of running	0	3	0	12	15
Varieties of walking	3	0	0	2	5

Table 1. MoM classes: Simple-ll association (Evert, 2008) to telic particles

⁴ These particles are defined *telic* because they systematically add an endpoint to the event encoded by their pivot verbs.

Table 1 reports the Simple-It (Evert 2008) correlation between MoM predicates and telic particles as fillers of the “adverbial modifier” slot. MoM classes are ranked according to the sum of the Simple-It values of the four telic particles. These data allow us to make a couple of considerations. First of all, the strong Simple-It association between telic particles and some of the MoM classes shows that these constructions are not idiosyncratic (in this case, Simple-It would have had a very low or negative value). This means that we are dealing with a compositional process, at least from a synchronic perspective. Second, the meaning component related to the directionality and speed of motion plays a crucial role in determining the association between Manner verbs and telic particles, which seems to be a feature of those verbs whose roots merge the MANNER and PATH meaning component.

3.2 Force Dynamics and Boundary Crossing

As outlined in section 2, the interaction between the manner component encoded in verb roots and boundary crossing events is central in Slobin's PIN/PIV hypothesis. As a first attempt to distributionally characterize this interaction, we considered the association strength with which these verbs play the syntactic role of modifiers in sentences whose main predicate encodes boundary crossing. Table 2 ranks the semantic classes of MoM as modifiers of a group of manually selected boundary crossing verbs⁵, with respect of the sum of the association strength between MoM modifier and boundary-crossing modified verb.

MoM Semantic Class	Boundary crossing
Impaired Walking	780
Labored Progress	277
Rapid Movement Toward Goal	271
Punctual, Repeteable Movement	270
Basic Level	202
Rapid Movement Away From Source	159
Rapid Movement	130
Varieties of Walking	105
Relaxed Walking	83
Types of Running	61
Frictionless, Silent Movement	58

Table 2. MoM classes: Simple-It association (Evert, 2008) to Boundary-crossing predicates

The correlation between MoM and boundary crossing verbs in the modifier-modified relation is clearly modulated by fine-grained semantic features. Crucially, the top-ranked classes are characterized by opposition of forces (*Impaired Walking*, *Labored Progress*), marked force-dynamics (*Rapid Movement toward Goal*) or ballistic features (*Punctual, Repeteable Movement*). The low compatibility of *Frictionless*,

5 Boundary crossing verbs: *attraversare* “cross”, *uscire* “go out”, *entrare* “enter”, *partire* “leave”, *arrivare* “arrive”, *rientrare* “go back in”, *sfociare* “flow (into)”, *traboccare* “spill over”, *fuoriuscire* “come out from”, *irrompere* “burst into”, *varcare* “cross”, *inciampare* “stumble”, *infilarsi* “slip into”, *intrufolarsi* “sneak in”, *superare* “go past”, *evadere* “evade”, *approdare* “land”, *sbarcare* “disembark”, *passare* “go past”, *sbucare* “come out”.

Silent Movement verbs confirms this tendency and supports Slobin's (2008) intuition: this is the MoM class that best fits Talmy's (1985) definition of *Force Dynamically Neutral* expressions, in which there are no forces opposing to each other.

4 Conclusion

This study represents an attempt to contribute to the debate concerning the typological classification of Italian with respect to the strategies of linguistic encoding of motion events. This work is characterized by a synchronic perspective (no claims are made concerning typological shifts), and by an empirical take on the problem: a reference corpus, a manual classification of MoM verbs and state of the art algorithms for the automatic analysis of verb argument structure constitute the main ingredients of our proposal. This approach is open to two lines of development: on the one hand, it would be interesting to track changes of frequency and distribution of MoM verbs in historical corpora, therefore taking a diachronic perspective; on the other hand, distributional data can be used to “challenge” the dictionary-based classification by carving a data-driven classification from the corpus. The comparison between the two classifications could give significant insights into the linguistic realization of meaning components of motion events.

References

- Baroni, M., Bernardini, S., Comastri, F., Piccioni, L., Volpi, A., Aston, G. & M. Mazzoleni (2004) Introducing the La Repubblica Corpus: A Large, Annotated, TEI(XML)-Compliant Corpus of Newspaper Italian. In *Proceedings of LREC 2004*, pages 1771-1774, Lisboa.
- Evert, S. (2008) Corpora and collocations. In Lüdeling, A. & Kyto, M (eds), *Corpus Linguistics. An International Handbook*, chapter 58. Mouton de Gruyter, Berlin.
- Hanks, P. & J. Pustejovsky (2005) A pattern dictionary for natural language processing. *Revue Française de linguistique appliquée*, (2005/2):63–82.
- Iacobini, C. (2010) The number and use of manner verbs as a cue for typological change in the strategies of motion events encoding. In Marotta, G., Lenci, A., Meini, L. & F. Rovai (eds). *Space in Language. Proceedings of the Pisa International Conference*, pages 495-514, Pisa.
- Iacobini, C. & F. Masini (2006) The emergence of verb-particle constructions in Italian: locative and actional meanings. *Morphology*, 16:155–188.
- Lenci, A. (forthcoming), Carving Verb Classes from Corpora" in Simone, R and F. Masini (eds) *Word Classes*, Amsterdam - Philadelphia: John Benjamins.
- Levin, Beth (1993) *English Verb Classes and Alternations. A Preliminary Investigation*. The University of Chicago Press, Chicago.
- Miller, G. & W. Charles (1991) Contextual Correlates of Semantic Similarity. *Language and Cognitive Processes*, 6(1):1–28.

- Schulte Im Walde, S. (2006) Experiments on the Automatic Induction of German Semantic Verb Classes. *Computational Linguistics*, 32(2):159–194.
- Slobin, D.I. (2008) Typology and Usage: beyond Verb-framed and Satellite-framed. Talk given at *Frames and Constructions: A Conference in Honor of Charles J. Fillmore*, Berkeley. University of California.
- Slobin, D.I. (2009) *From S-language and V-language to PIN and PIV*. Talk given at the Human Locomotion Across Languages Workshop, Nijmegen, Max Planck Institute for Psycholinguistics.
- Talmy, L. (1985) Lexicalization patterns: Semantic structure in lexical forms. In *Language Typology and Lexical Description: Vol. 3. Grammatical Categories and the Lexicon*, pages 36–149.
- Talmy, L. (1991) Path to realization: A typology of event conflation. *Proceedings of the Berkeley Linguistics Society*, 17:480–519.

Appendix A

Class	Verbs
Basic Level	camminare (to walk), correre (to run), deambulare (to walk), incamminarsi (to make one's way)
Varieties of Walking	ancheggiare (to waddle), ballare (to dance), calcare (to stamp, to tread), danzare (to dance), incedere (to sweep), marciare (to march), ondeggiare (to waddle), pavoneggiarsi (to strut about, to swagger), sculettare (to waddle), sfilare (to parade), sgambare (to stride), trotterellare (to toddle), zampezzare (to toddle)
Relaxed Walking	aggirarsi (to roam), bighellonare (saunter, lounge), dondolarsi (lounge), errare (wander), girandolare (ramble, saunter, stroll), girellare (ramble, saunter, stroll), gironzolare (ramble, saunter, stroll), girovagare (ramble, saunter, stroll), passeggiare (stroll), peregrinare (wander), vagabondare (rove), vagare (meander, wander)
Labored Progress	brancolare (to grope around), gattonare (to crawl), guadare (to wade), infilarsi (to squeeze oneself), insinuarsi (to insinuate oneself), penetrare (to penetrate), strascicare (to shamble), strascinarsi (to shuffle, to drag oneself), strisciare (to crawl, to creep, to worm), trascinarsi (to traipse, to shuffle, to drag oneself)
Impaired Walking	arrancare (to hobble, to limp), barcollare (to stagger, to totter), ciondolare (to totter, to stagger), claudicare (to hobble, to limp), incespicare (to stumble), inciampare (to stumble), pencolare (to totter, to stagger), tentennare (to stagger, to totter), traballare (to totter, to stagger), vacillare (to totter, to stagger), zoppicare (to hobble)
Types of Running	galoppare (to gallop), sprintare (to sprint), trottare (to trot)
Rapid Movement	filare (to race), frondarsi (to dash), guizzare (to scoot), precipitarsi (to dash, to streak), scattare (to shoot), scorrazzare (to scamper), sfrecciare (to dart), sgambettare (to scamper), volare (to fly)
Rapid Movement away from Source	battersela (to scam), dileguarsi (to flee), eclissarsi (to disappear), filarsela (to scam), fuggire (to abandon, to bolt), rinculare (to recoil), sbucare (to come out suddenly), scantonare (to steal away), scappare (to escape), sfilarsi (to slip away), sfuggire (to escape), sgattaiolare (to slip away), sgusciare (to slip away), squagliarsela (to steal away), svicolare (to slink away), svignarsela (to steal away), volatilizzarsi (to disappear)
Rapid Movement Toward Goal	accorrere (to run toward, to hasten), aggredire (to assault), assalire (to assault), attaccare (to pounce), avventarsi (to hurl oneself), braccare (to stalk), buttarsi (to hurl oneself), caricare (to charge), gettarsi (to hurl oneself), incalzare (to chase, to follow up closely), inseguire (to run after), investire (to assail), irrompere (to burst into), lanciarsi (to hurl oneself), piombare (to assail), riversarsi (to swarm), scagliarsi (to lunge), slanciarsi (to hurl oneself)
Frictionless, Silent Movement	intrufolarsi (to slip in), scivolare (to glide), sdrucciolare (to slip), serpeggiare (to sneak), slittare (to skid, to slide)
Punctual, Repeatable Movement	ballonzolare (to trip along), balzare (to jump), balzellare (to skip), piroettare (to pirouette), rimbalsare (to bound), saltare (to hop), saltellare (to hop), scavalcare (to jump, to step over, to across), sobbalzare (to jump), spuntare (to spring), sussultare (to leap), volteggiare (to vault), zompare (to hop)