Modelling Italian construction flexibility with distributional semantics: Are constructions enough?

Lucia Busso **Ludovica Pannitto** **Alessandro Lenci**
CoLing Lab, University of Pisa
{lucia.busso90,ellepannitto}@gmail.com, alessandro.lenci@unipi.it

Abstract

English. The present study combines psycholinguistic evidence on Italian valency coercion and a distributional analysis. The paper suggests that distributional properties can provide useful insights on how general abstract constructions influence the resolution of coercion effects. However, complete understanding of the processing and recognition of coercion requires to take into consideration the complex intertwining of lexical verb and abstract constructions.

Italiano. *Il lavoro unisce uno studio psicolinguistico sul fenomeno della coercion valenziale in Italiano con un'analisi distribuzionale.L'articolo suggerisce che le proprietà distribuzionali forniscano un'utile passaggio per capire l'influenza delle costruzioni alla risoluzione di effetti di coercion. Tuttavia, una piena comprensione del fenomeno richiede di prendere in considerazione la complessa relazione tra verbo e costruzione argomentale.*

1 Introduction

In Construction Grammar (Goldberg, 2006), the basic units of linguistic analysis are called *constructions* (Cxns), form-meaning pairings associated with autonomous, non-compositional abstract meanings, independently from the lexical items occurring in them. Examples of Cxns range from morphemes (e.g., *pre-, -ing*), to filled or partially-filled complex words (e.g., *daredevil*) to idioms (e.g., *give the devil his dues*) to more abstract patterns like the Ditransitive [Subj V Obj1 Obj2] (e.g., *he gave her a book*) (Goldberg, 2006).

Cxns appear at any level of linguistic analysis, but the level at which the notion of constructional meaning represents a radical departure from other theories of grammar is *argument structure*. These Cxns, such as the English Ditransitive, are claimed to be associated with an abstract semantic content. In this case, constructional meaning can be paraphrased as *X CAUSES Y TO RECEIVE Z.* One of the main supporting arguments in favour of constructions as independent and primitive objects of grammar is the flexibility with which argument Cxns and verbs interact with each other, as in example (1) in which the original intransitive sense of "to sneeze" is overridden by the Caused Motion Cxn, and thus takes a transitive sense of "making something move by sneezing".

(1) *John sneezed the napkin off the table*

This flexibility in combining Cxns and verbs is known as *valency coercion* (Michaelis, 2004; Boas, 2011; Lauwers and Willems, 2011; Perek and Hilpert, 2014).

This phenomenon, although vastly addressed for English, has not yet received a systematic investigation in other languages. For notable exceptions, see Boas and Gonzálvez-García (2014). In particular – to the best of our knowledge – no previous attempt to carry out an empirical investigation of valency coercion exists for Italian. However, even a simple corpus query reveals that the phenomenon is present in Italian, though it is not as pervasive as in English:

(2) *Tossì una risata leggera tra i suoi capelli* (He coughed a light laugh in her hair) [ItWac]

This paper presents an analysis of Italian constructional flexibility that combines psycholinguistic and computational evidence: first, we present the results of a behavioral experiment on valency coercion. Then, we model Cxns with distributional semantics to investigate whether the semantic shape of Italian argument Cxns can affect the interpretation and processing of coerced sentences.

2 Studying valency coercion: an acceptability rating task

MATERIALS AND SUBJECTS: The offline psycholinguistic experiment targets 9 Italian Cxns (see Table 1) that were selected using existing resources: *LexIt* (Lenci et al., 2012) and *ValPal* (Cennamo and Fabrizio, 2013). The resultant Cxns are of varying abstractness and schematicity levels (Barðdal, 2008).

Cxn	**frames**
CAUSED MOTION (CM)	NPj-V-NP -PPlocation
CAUSED MOTION + via (CMvia)	NPs-V-NPobj
DATIVE (DT)	NPs-V-NPj-PPrecipient
INTRANSITIVE MOTION (IM)	NPs-V–PPlocation
PASSIVE (PASS)	NPs-V-PP
PREDICATIVE (PRED)	NPs-V–AdjPpredicate
VERBA DICENDI explicit (sentential) (VDE)	NPs-V-cheVP
VERBA DICENDI implicit (sentential) (VDI)	NP-V-diVP

Table 1: Constructions used in the test.

For each Cxn, we built 21 sentences, which were subdivided into 3 experimental conditions: GRAMMATICAL (3a), COERCION (3b), IMPOSSIBLE (3c) (7 sentences per condition). The total number of stimuli amounts to 189 sentences. The structure of the test was inspired by Perek and Hilpert (2014). Between conditions, sentences differ only for their main verb, to have as little variation as possible.

(3) a. *Gianni ha detto che verrà domani* (Gianni said that he will come tomorrow)

b. *Gianni ha fischiettato che verrà domani* (Gianni whistled that he will come tomorrow)

c. *Gianni ha cucinato che verrà domani* (Gianni cooked that he will come tomorrow)

The coercion condition consists of verbs that display a partial semantic incompatibility with the constructional environment they are embedded in. They were selected by means of both native intuition and corpus query, selecting and refining cases that were either hapax or rare occurrences in the Italian corpus *ItWac* (Baroni et al., 2009).

120 Italian native speakers were tested: 39 adolescents (12-14 years old), 40 young adults (18-35 years old), and 41 adults (over 40). We tested subjects of different ages following extensive sociolinguistic literature that has shown that language use changes with age (Eckert, 2017; Labov, 2001; Wagner, 2012). Thus, it could be the case that grammaticality judgments on creative, non-standard sentences are also affected by age. Including different age groups in our analysis allows us to investigate a more representative sample of the population. To control for the possible influencing factor of education level, we only tested adult speakers either in possess of (at least) a bachelor degree or enrolled in a University course. Table 2 summarizes number, age groups and distribution of tested subjects.

Age group	Age range	distribution	Gender	Tot
Adolescents	12-14	mean: 12.9 sd:0.63	24 m (61,5%) 15 f (38,4%)	39
Young Adults	18-39	mean:27.3 sd:2.94	15 m (37,5%) 25 f (62,5%)	41
Adults	Over 40	mean: 56.7 sd:9.48	18 m (43,9%) 23 f (56,1%)	40

Table 2: Data about tested subjects.

A within-subject design was used, in which each subject sees all stimuli. Participants were asked to judge the acceptability of the (randomized) stimuli on a Likert scale from 1 - "completely unnatural" - to 7 - "perfectly natural". Presentation of the data varied across age groups: adolescents were given the test directly in their class. Young adults' judgments were collected through the online platform Figure Eight. Older adults, instead, were presented with a simple Microsoft Word document, in order to include participants who did not have familiarity with online data gathering.

RESULTS: We assessed statistical significance via linear mixed effect modelling, with by-subject and by-item intercepts.[1] Results show that coercion sentences (purple boxplot in Figure 1) are recognized as an intermediate condition between complete grammaticality and total ungrammaticality.[2] We consider this result to support the claim that coercion effects include a degree of semantic incompatibility that is nonetheless resolved in the interpretation process. Consistently

[1] model selection performed automatically via LRT with the R package *afex*. Models were performed with the R package *lmerTest* and R2 values were calculated with the MuMIn package (Singmann et al., 2016; Kuznetsova et al., 2017; Bartoń, 2013)

[2] $p < 0.0001$, R2c 0.61

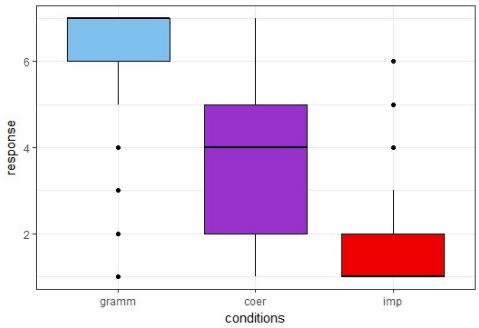

Figure 1: distribution of judgments in the 3 conditions

with the main tenets of Construction Grammar, we argue that the resolution of such incompatibility is driven by a dynamic interaction between the main verb and the constructional context (Kemmer, 2008; Kemmer and Yoon, 2013; Yoon, 2016). In a second analysis, we wanted to assess the effect of Cxn types on acceptability ratings. We used linear mixed effect modelling, adding an interaction between Cxn type and experimental condition.[3] Results indicate high variability in Cxn 'coercibility' (see Figure 2 and table 3). That is, some Cxns in our dataset were consistently judged as more natural by speakers in the coercion condition.

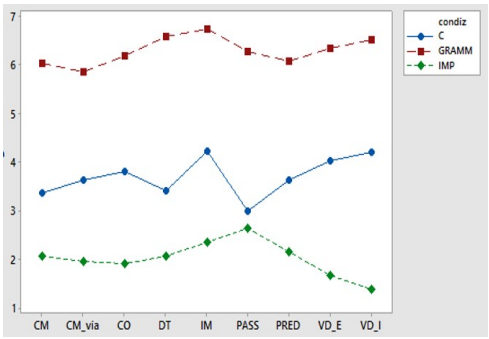

Figure 2: line plot of judgments

In particular, it appears that IM, VDE and VDI Cxns result to be more natural, while DT, PASS and (marginally) CO are the least naturally perceived ones in coercion sentences. Since coercion effects are said to be resolved by the general Cxn semantics overriding the lexical meaning of the verb, we hypothesize that the different flexibility degrees of the Cxns in the first experiment could be at least partially explained by distributional properties, such as type and token frequency, and semantic density of the Cxns in our

[3] $p < 0.0001$, R2c 0.76

	Estimate	Std. Error	t value	p value
coer	**3,64*** **	0,1	37,45	<0.0001
gramm	**2,66*** **	0,02	110,87	<0.0001
imp	**-1,79*** **	0,02	-74,84	<0.0001
CM	-0,14	0,16	-0,91	0,36
CMvia	-0,24	0,16	-1,53	0,13
CO	**-0,26.**	0,13	-1,95	0,05
DT	**-1,34*** **	0,17	-7,98	<0.0001
IM	**1,02*** **	0,16	6,40	<0.0001
PASS	**-0,73** **	0,26	-2,75	0,009
PRED	-0,07	0,26	-0,27	0,79
VDE	**1.06*** **	0,16	6,67	<0.0001
VDI	**0,70*** **	0,15	4,57	<0.0001

Table 3: fixed effects estimates of the coercion condition

dataset, the latter again estimated with distributional semantics.

Different degrees of flexibility could derive either from cognitive processes that reflect on language use, or emerge from repeated exposure and thus entrench in speakers' grammar. Both possible directions of this causal circle, however, ultimately allow us to fruitfully investigate construction flexibility using distributional semantics models. In other words, the higher 'coercibility' of novel instances of some Cxns could be due to speakers' sensitivity to distributional semantic features of the constructions (Barddal, 2006; Bybee, 2013; Zeschel, 2012; Perek and Goldberg, 2017).

3 A Distributional Semantic Model for argument constructions

PROCEDURE: Perek (2016) has shown that distributional semantics (Lenci, 2018) can be fruitfully used to model the semantic space covered by a Cxn. It has been argued in the literature that constructional meanings for argument Cxns arise from the meaning of high frequency verbs that co-occur with them (Goldberg, 1999; Casenhiser and Goldberg, 2005; Barak and Goldberg, 2017). Therefore, we modelled the semantic content of Cxns with the semantics of their most typical verb, each represented as a distributional vector.

We used the UDLex Pipeline[4] (Rambelli et al., 2017) to obtain a mapping between the Cxns of our dataset and the most frequent verbs that occur in them (these were selected considering verbs that appear at least 5 times in the relevant subcatego-

[4] The UDLex Italian dataset consist of 409,127 tokens.

rization frames). Table 4 summarizes the number of verbs considered for each of the eight Cxns.[5] Then, we built a Distributional Semantic Model (DSM) from the italian corpus *itWac* (Baroni et al., 2009) in order to represent verb meaning of the verbs obtained with UDLex. The 300-dimensional vectors (i.e., the embeddings) were created with the SGNS algorithm (Mikolov et al., 2013), using the most frequent 30,000 words as context, with a minimum frequency of 100.

Cxn	type freq (different verbs)	token freq (number of items)
CM	103	1538
CO	5	43
DT	90	1659
IM	51	1097
PASS	8	49
PRED	19	359
VD_E	12	116
VD_I	15	199

Table 4: Number of selected verbs per Cxn.

Following Lebani and Lenci (2017), we represented each Cxn as the weighted centroid vector of its typical verbs, as follows:

$$\overrightarrow{CXN} = \frac{1}{|V|} \sum v \in V frel(v, Cxn) \cdot \vec{v} \quad (1)$$

where V the set of the top-associated verbs v with Cxn and $frel(v, Cxn)$ is the co-occurrence frequency of a verb in a Cxn.

We measured the pairwise cosine similarity among the weighted Cxn vectors: as shown in Figure 3, the distributional behaviour of the Cxn vectors suggests that some Cxns in our dataset show similar distributional behaviour.

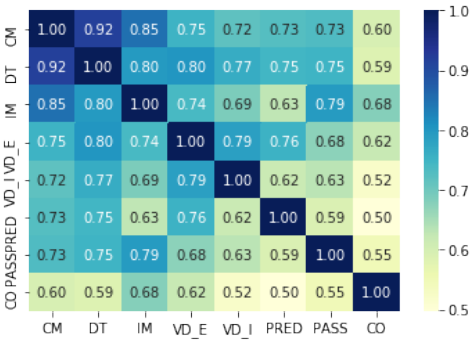

Figure 3: Construction semantic similarity.

[5]the Cxn CMvia was excluded due to the absence of corresponding subcategorization frames

Following Perek (2016), the semantic density of a Cxn is computed as the mean value of pairwise cosines between the verbs occurring in Cxn. Figure 4 plots the semantic densities of our Cxns.

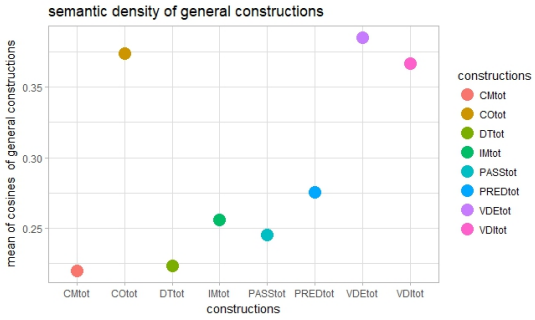

Figure 4: Construction semantic density.

Finally, to assess the effect of distributional properties on Cxns flexibility, we used semantic density, type frequency and token frequency (cf. Table 4) as predictors in linear mixed effect modelling. As dependent variable, we used the difference $gramm - coer$ and $coer - imp$. We performed two separate analyses for type and token frequencies without interactions to avoid multicollinearity effects. Predictors values were centered.

RESULTS: The estimates are reported in Tables 5 and 6 below. In the first two models frequency does not yield any effect. In the second models, instead, frequency appears to have an effect on the data. Hence, it appears that type and token frequency help discerning impossible from coercion instances of a Cxn, whereas only semantic density affects the higher naturalness of coercion phenomena. The more a Cxn is observed with semantically similar verbs (i.e., verbs that belong to the same classes or subclasses, which therefore increase the Cxn semantic density), the more the constructional meaning is easily coerced into novel instances.

4 Discussion

These findings support our claim that coercion effects are resolved by a dynamic interrelation between verb and Cxn (Kemmer, 2008; Kemmer and Yoon, 2013). Even though frequency effects are shown to affect Cxns extensibility to new items (Bybee, 2006), our results suggest that type and token frequency only facilitate the distinc-

(Gramm - coer) ∼sem. dens + type freq.				
	estimate	st. error	t value	p value
(Intercept)	**2.71*** **	0.11	25.02	<0.0001
Sem. density	**-0.34.**	0.16	-2.217	0.007
Type freq.	-0.13	0.16	-0.848	0.44
(Gramm - coer) ∼sem. dens + tok freq.				
	estimate	st. error	t value	p value
(Intercept)	**2.71*** **	0.11	25.02	<0.0001
Sem. density	**-0.35.**	0.16	-2.23	<0.1
Token freq.	-0.13	0.16	-0.89	0.42

Table 5: Fixed effects table for the first two models.

(Coer - imp) ∼sem. dens + type freq.				
	estimate	st. error	t value	p value
(Intercept)	**1.69*** **	0.15	10.87	<0.0001
Sem. density	**0.86***	0.22	3.38	<0.01
Type freq.	**0.47.**	0.22	2.1	<0.1
(Coer – imp) ∼sem. dens. + tok. freq.				
	estimate	st. error	t value	p value
(Intercept)	**1.69*** **	0.14	33.33	<0.0001
Sem. density	**0.91***	0.2	4.59	<0.001
Token freq.	**0.54***	0.2	2.71	<0.01

Table 6: Fixed effects table for the second two models.

tion between semantically incompatible and partially compatible formulations, whereas higher coercibility is only affected by semantic density.

We interpret this finding in light of the *upward strengthening hypothesis* (Hilpert, 2015), according to which a novel occurrence of a linguistic unit strengthens a superior node (i.e., the abstract Cxn) only if the former is categorized 'as an instance of a more abstract Cxn. If this categorization is not performed, or only superficially so, no upward strengthening will take place' (Hilpert, 2015, p.38). Higher coercibility is hence not affected by frequency of the Cxn because of the 'intermediate' grammaticality level of coercion, which does not allow unambiguous categorization. Coercion sentences result more natural if the target Cxn is observed with verbs belonging to similar semantic classes or subclasses, which increases Cxn semantic density. We could therefore assume that coercion effects in Italian elicit a *partial categorization*. The effect of semantic density, however, only explains part of the data. In fact, visual inspection of the relation between semantic density and the estimates of table 3 reveals that this effect does not explain the high coercibility of IM, or the low values of CO Cxns (see Figure 5).

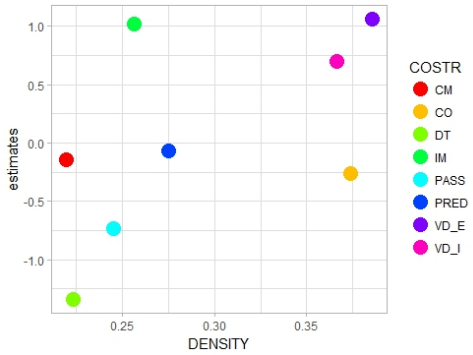

Figure 5: relation semantic density- estimates

All things considered, semantic properties (modelled with distributional vectors) of Cxns (e.g., its density) are only one of the factors influencing speakers processing and recognition of coercion effects. In fact, it has been argued that Romance languages are more *valency driven* than English (and Germanic languages in general) (Perek and Hilpert, 2014). The results of both experiments provide substantial evidence for an integrated account of Italian coercion effects, which should consider not only the properties of the general abstract Cxn, but rather the interaction of the mismatching verb with Cxn meaning.

These result also have interesting implications to understand the cognitive mechanisms underlying Cxn flexibility and productivity. In fact, these findings support the idea that Cxn meaning is abstracted from the semantics of prototypically occurring verbs. As we saw, several studies have argued in favour of this hypothesis for English, but the fact that we were able to adapt it to Italian suggests that the factors driving the acquisition of Cxns are - at least partially - not language-specific but rather general cognitive processes.

Acknowledgments:

The authors thank Lucia Passaro and Florent Perek for their help and valuable suggestions.

References

Libby Barak and Adele E. Goldberg. 2017. Modeling the Partial Productivity of Constructions.

Jóhanna Barðdal. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. 12.

Jóhanna Barddal. 2006. Predicting the Productivity of Argument Structure Constructions. *Annual Meeting of the Berkeley Linguistics Society*, 32(1):467, October.

Marco Baroni, Silvia Bernardini, Adriano Ferraresi, and Eros Zanchetta. 2009. The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation*, 43(3):209–226.

Kamil Bartoń, 2013. *MuMIn: multi-model inference, R package version 1.9.13*. CRAN http://CRAN.R-project.org/package=MuMIn.

Hans Christian Boas and Francisco Gonzálvez-García, editors. 2014. *Romance perspectives on construction grammar*. Number volume 15 in Constructional approaches to language. John Benjamins Publishing Company, Amsterdam ; Philadelphia.

Hans C. Boas. 2011. Coercion and leaking argument structures in Construction Grammar. *Linguistics*, 49(6), January.

J. Bybee. 2006. *Frequency of Use and the Organization of Language*. Oxford University Press.

Joan L Bybee. 2013. Usage-based theory and exemplar representations of constructions. In *The Oxford handbook of construction grammar*.

Devin Casenhiser and Adele E Goldberg. 2005. Fast mapping between a phrasal form and meaning. *Developmental Science*, 8(6):500–508.

Michela Cennamo and Claudia Fabrizio, 2013. *Italian Valency Patterns*. Max Planck Institute for Evolutionary Anthropology, Leipzig.

Penelope Eckert. 2017. Age as a Sociolinguistic Variable. In *The Handbook of Sociolinguistics*, pages 151–167. Wiley-Blackwell.

Adele E. Goldberg. 1999. The emergence of the semantics of argument structure constructions. In *The emergence of language*, pages 215–230. Psychology Press.

Adele E. Goldberg. 2006. *Constructions at work: the nature of generalization in language*. Oxford linguistics. Oxford University Press, Oxford ; New York.

Martin Hilpert. 2015. From hand-carved to computer-based: Noun-participle compounding and the upward strengthening hypothesis. *Cognitive Linguistics*, 26(1), January.

Suzanne Kemmer and Soyeon Yoon. 2013. Rethinking coercion as a cognitive phenomenon: Data from processing, frequency, and acceptability judgments.

Suzanne Kemmer. 2008. September. new dimensions of dimensions: Frequency, productivity, domains and coercion. In *meeting of Cognitive Linguistics Between Universality and Variation, Dubrovnik, Croatia*.

Alexandra Kuznetsova, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26.

W. Labov. 2001. *Principles of Linguistic Change, Social Factors*. Principles of Linguistic Change. Wiley.

Peter Lauwers and Dominique Willems. 2011. Coercion: Definition and challenges, current approaches, and new trends. *Linguistics*, 49(6), January.

Gianluca E. Lebani and Alessandro Lenci. 2017. Modelling the Meaning of Argument Constructions with Distributional Semantics. In *Proceedings of the AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding*, pages 197–204.

Alessandro Lenci, Gabriella Lapesa, and Giulia Bonansinga. 2012. Lexit: A computational resource on italian argument structure. In *LREC*.

Alessandro Lenci. 2018. Distributional Models of Word Meaning. *Annual Review of Linguistics*, 4:151–171.

Laura A. Michaelis. 2004. Type shifting in construction grammar: An integrated approach to aspectual coercion. *Cognitive Linguistics*, 15(1):1–67, January.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.

Florent Perek and Adele E. Goldberg. 2017. Linguistic generalization on the basis of function and constraints on the basis of statistical preemption. *Cognition*, 168:276–293.

Florent Perek and Martin Hilpert. 2014. Constructional tolerance: Cross-linguistic differences in the acceptability of non-conventional uses of constructions. *Constructions and Frames*, 6(2):266–304.

Florent Perek. 2016. Using distributional semantics to study syntactic productivity in diachrony: A case study. *Linguistics*, 54(1):149–188.

Giulia Rambelli, Alessandro Lenci, and Thierry Poibeau. 2017. UDLex: Towards Cross-language Subcategorization Lexicons. In *Proceedings of*

the Fourth International Conference on Dependency Linguistics (Depling 2017), September 18-20, 2017, Università di Pisa, Italy, pages 207–217. Linköping University Electronic Press.

Henrik Singmann, Ben Bolker, Jake Westfall, and Frederik Aust, 2016. *afex: Analysis of Factorial Experiments*. R package version 0.16-1.

Suzanne Evans Wagner. 2012. Age Grading in Sociolinguistic Theory: Age Grading in Sociolinguistic Theory. *Language and Linguistics Compass*, 6(6):371–382, June.

Soyeon Yoon. 2016. Gradable nature of semantic compatibility and coercion: A usage-based approach. *Linguistic Research*, 33(1):95–134, March.

Arne Zeschel. 2012. *Incipient productivity: a construction-based approach to linguistic creativity*. Number 49 in Cognitive linguistics research. De Gruyter Mouton, Berlin ; Boston.