I *si* di Univers-*ITA* e il tratto di persona

Elisa Di Domenico
Università per Stranieri di Perugia
elisa.didomenico@unistrapg.it

I. Introduzione

In questo studio presentiamo un'indagine sulla distribuzione dei clitici in enclisi o proclisi nelle frasi 'a ristrutturazione' (Voglio far*lo*/*Lo* voglio fare). L'indagine, effettuata sul *corpus* Univers-ITA **[6]**, offre lo spunto per una riflessione sulla natura dei dati- in particolare di quelli provenienti da *corpora*- nelle scienze del linguaggio. Considereremo in particolare il rapporto tra dati e teoria, la natura dei dati stessi (cosa si trova e cosa non si trova), alcune caratteristiche di un *corpus* che contribuiscono a rendere i dati attendibili o a mettere in luce spunti interpretativi.

II. Il *corpus*

Univers-ITA **[6]** è una collezione di 2160 testi scritti [250-500 parole] prodotti da studenti universitari su un unico tema e con un destinatario specifico. Nella consegna del compito viene raccomandato di usare uno stile il più possibile corretto e formale. Le sedi universitarie di raccolta ammontano a 30, diffuse su tutto il territorio nazionale*.

III. Fasi preliminari: La ricerca automatica e l'estrazione manuale dei contesti rilevanti

Fase 1: estrazione automatica delle frasi contenenti un pronome clitico (mi, ti, vi, lo, [la, li], le, gli, ne, ci, si), effettuata per tipo di clitico
[*N*= 20.330 => **19.276**]
Fase 2: Selezione manuale delle frasi 'a ristrutturazione'
[*N*= **2.618**] (Figura 1, Tavola 1)
Fase 3: Identificazione manuale dei casi di enclisi e proclisi per ciascun tipo di clitico (Tavola 1, Figura 2)
Risultati:

Figura 1: Consistenza (%) di ciascun clitico

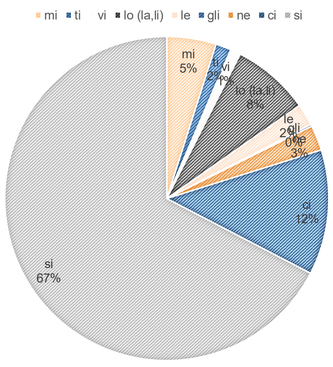
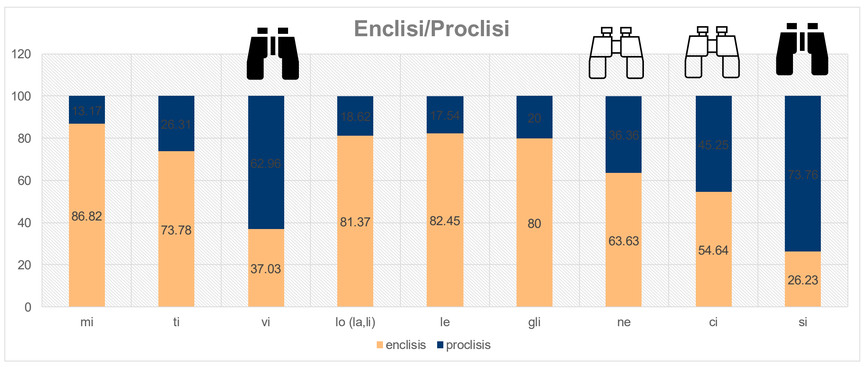


Tavola 1: Occorrenze

	Enclisi	Proclisi	Totale
mi	112	17	129
ti	28	10	38
vi	10	17	27
lo (la,li)	166	48	204
le	47	38	57
gli	4	1	5
ne	42	24	66
ci	179	148	327
si	463	1302	1765
Totale	1051	1567	**2618**

Figura 2. % enclisi e proclisi per ciascun clitico



- L'enclisi prevale sulla proclisi per tutti i clitici tranne *vi* e *si*
- La percentuale di enclisi è significativamente più bassa (p=.005) per *ci* e *ne***

IV. Considerazioni e ulteriori analisi

Per quanto riguarda *vi,* la preferenza per la proclisi si osserva solo quando *vi* non ha un valore personale:

Tavola 2. P e non-P *vi*

	Enclisi	Proclisi	Totale
Non-P	5 (27.77%)	13 (77.22%)	18
P	5 (55.55%]	4 (44.44%)	9
Totale	10	17	27

Per quanto riguarda *ci*, la preferenza per l'enclisi è significativamente maggiore (p=.001)quando *ci* ha un valore personale

Tavola 3. P e non-P *ci* ***

	Enclisi	Proclisi	Totale
Non-P	46 (60.52%)	30 (39.47%)	76
P	133 (77.32%)	39 (22.67%)	172
Totale	179	69	248

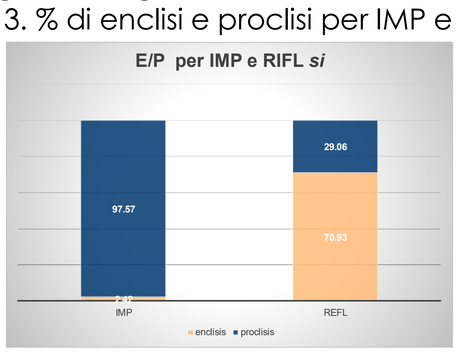
Uniti ai dati su *ne,* non dotato di un tratto di persona **[3]**, questi dati suggeriscono la **<u>Generalizzazione I</u>**:
I clitici dotati di un tratto di persona tendono maggiormente a distribuirsi in enclisi

-Può la Generalizzazione I spiegare i dati riguardanti *si*?

Nella letteratura si distinguono vari tipi di *si:*
(1)

a.	Durante le feste si mangia troppo	impersonale	IMP
b.	Questo concetto si definisce facilmente	medio	
c.	Le patate si mangeranno domani	passivo	
d.	Giulia si lava	riflessivo	RIFL
e.	Giulia si beve un caffè	benefattivo	
f.	Giulia si è arrabbiata	inerente	

Qui distingueremo:
IMP*si*: variabile libera, int. arbitraria **[5,7,8]**
REFL*si*: anaforico **[7]**
Applicando questa distinzione, il quadro che emerge è il seguente:
Figura 3. % di enclisi e proclisi per IMP e RIFL *si*



- I due tipi di *si* hanno una distribuzione E/P molto diversa
- RIFL*si* ha la distribuzione dei clitici P
- Nonostante il suo nome, IMP*si* non ha la distribuzione dei clitici non-P: solo nel 2,48% dei casi (28/1.156) è in enclisi. La proclisi appare un requisito.

=> La Generalizzazione I non è applicabile a IMP*si*

V. Conclusioni specifiche: due generalizzazioni indipendenti

L'analisi dei casi di IMP*si* in enclisi rivela un altro fatto interessante: quando IMPsi è in enclisi, la frase ha sempre un'interpretazione passiva:

(2) In questa didattica a distanza potrebbero anche veder*si* dei vantaggi [1259 Centro]

La generalizzazione relativa a IMP*si* risulta dunque la seguente:
<u>Generalizzazione II</u>
a) IMP*si* occorre quasi sempre in proclisi
b) Nei rari casi in cui ciò non avviene, la frase ha un'interpretazione passiva

In conclusione, l'analisi dei dati rivela due generalizzazioni indipendenti:
G.I: è relativa a tutti i clitici tranne IMP*si,* ripartendoli in due gruppi:

P-clitici	**non-P clitici**
mi, *ti*, p-*vi*, p-*ci*,	Loc/ex *vi*
lo (la, *li)*, *le*, *gli*,	Loc/ex/com etc. *ci*
RIFL*si*	*ne*
---	---
individui (+/- animati)	parti, luoghi, eventi

G.II: è relativa solo a IMP*si,* non coinvolge il tratto di persona, ma piuttosto un tratto[+/- soggetto], come suggerisce G.II b). G.II a) conferma [2] su una scala più ampia
Le due generalizzazioni naturalmente andranno spiegate: questo sarà oggetto di lavori futuri.

VI. Conclusioni generali

1) I dati non contengono in sé un'etichetta utile ad interpretarli, ed anche solo per arrivare a formulare delle generalizzazioni, il ricorso alla teoria è imprescindibile. **2)** Lavorando su un *corpus,* non sappiamo se ciò che non troviamo potrebbe comunque esserci. L'ampiezza del *corpus* è comunque importante [nel nostro caso i dati su *si*] per dare validità a ciò che si trova **3)** Le caratteristiche di interrogabilità del *corpus* Univers-ITA hanno permesso di inserire una nuova variabile –tipo di clitico- nel dibattito su ciò che influenza la distribuzione dei clitici in enclisi o proclisi nelle frasi 'a ristrutturazione' **[1, 2, 4, 9].**

Riferimenti bibliografici essenziali

[1] Benincà, P. (1986) "Punti di sintassi comparata dei dialetti italiani settentrionali". Holtus, G. and Ringger, K. (eds.), *Raetia Antiqua et Moderna: W. Theodor Elwert zum 80. Geburztag*. Tübingen: Niemeyer, 457-79. **[2] Berretta, M. (1986)** "Struttura informativa e sintassi dei pronomi atoni: Condizioni che favoriscono la 'risalita'". Stammerjohan, H. (ed.), *Tema/Rema in italiano*. Tübingen: Narr, 71-83. **[3] Cardinaletti, A.,Giusti, G. (1992)** "Partitive *ne* and the QP- Hypothesis". Fava, E. (ed.), *Proceedings of the XVII Meeting of Generative Grammar*. Torino: Rosemberg & Sellier, 121-41. **[4] Cardinaletti, A., Giusti, G., Lebani, G. (2025)** "Clitic Climbing across Italy. Variation, optionality, and the role of bilectalism". *Linguistic Variation*, 26(1), 76-96. **[5] Cinque G. (1988)** "On Si constructions and the theory of Arb". *Linguistic Inquiry* 19, 521–81. **[6] Grandi N., Ballarè. S., Chiusaroli F., Gallina F., Pascoli M. e Pistolesi E. (2023).** *Corpus Univers-ITA 2023*, doi: https://doi.org/10.60760/unibo/univers-ita. **[7] Manzini, M. R. (1986)** "On Italian *Si*". Borer, H. (ed.), *Syntax and Semantics 19: The syntax of pronominal clitics*. New York: Academic Press, 241-62. **[8] Pescarini, D. (2017)** "Parametrising arbitrary constructions". *Probus* ⟨10.1515/probus-2017-0005⟩. **[9] Rohlfs, G. (1968)** *Grammatica storica dell'italiano e dei suoi dialetti. Morfologia*. Torino: Einaudi.