

Pellegrini, Matteo

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

Études romanes de Brno. 2024, vol. 45, iss. 3, pp. 70-101

ISSN 2336-4416 (online)

Stable URL (DOI): <https://doi.org/10.5817/ERB2024-3-5>

Stable URL (handle): <https://hdl.handle.net/11222.digilib/digilib.80985>

License: [CC BY-SA 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Access Date: 03. 01. 2025

Version: 20241231

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

From Stems to Forms: Paradigm Zones in Italian Verb Inflection

MATTEO PELLEGRINI [matteo.pellegrini@unicatt.it]

Università Cattolica del Sacro Cuore, Italia

SOMMARIO

Questo lavoro presenta un'analisi della prevedibilità reciproca tra le celle del paradigma nella flessione verbale dell'italiano. L'approccio è astrattivo: il punto di partenza sono le forme flesse, senza assumere alcuna segmentazione *a priori*. La metodologia si basa su nozioni di teoria dell'informazione: l'incertezza nel prevedere una forma a partire da un'altra è misurata tramite l'entropia condizionale. I risultati sono usati per tracciare una mappatura del paradigma verbale dell'italiano in zone, tali per cui celle che appartengono alla stessa zona sono sistematicamente prevedibili le une dalle altre. Questa mappatura è confrontata con quella di alcuni importanti studi precedenti che hanno invece adottato un approccio costruttivo, basato sulla prevedibilità tra temi. Infine, la quantificazione dell'incertezza tramite l'entropia condizionale è usata come metrica di valutazione empirica per stabilire quale approccio dia luogo a un'analisi più economica tra quello astrattivo del presente lavoro e quello costruttivo degli studi precedenti.

PAROLE CHIAVE

paradigmi; flessione verbale; italiano; prevedibilità; entropia

ABSTRACT

This work presents an analysis of interpredictability between paradigm cells in Italian verbal inflection. The approach is abstractive: the starting point is inflected forms, without assuming any segmentation *a priori*. The methodology is based on information-theoretic notions: the uncertainty in predicting a form from another one is measured by means of conditional entropy. The results are used to draw a mapping of the verbal paradigm of Italian in zones, such that cells that belong to the same zone are systematically predictable from one another. This mapping is compared to the one of important previous studies that have adopted a constructive approach, based on predictability between stems. Lastly, the quantification of uncertainty through conditional entropy is used as evaluation metrics to establish which approach yields a more economical analysis between the abstractive one of the present work and the constructive one of previous studies.

KEYWORDS

paradigms; verbal inflection; Italian; predictability; entropy

RICEVUTO 2023-11-30; ACCETTATO 2024-02-20

1. Introduzione

Nei paradigmi flessivi, i diversi valori delle proprietà morfosintattiche sono talvolta espressi in maniera non uniforme e non totalmente prevedibile. Tale imprevedibilità può manifestarsi a livello delle desinenze selezionate in diverse forme flesse – che possono variare a seconda del lessema in questione, dando luogo a diverse classi flessive¹ – oppure a livello dei temi che compaiono in tali forme – che possono variare a seconda della cella del paradigma interessata, per via di fenomeni di allomorfia tematica o suppletivismo. A titolo esemplificativo, nella Tabella 1 si riportano le forme flesse che occupano una selezione delle celle del paradigma di nomi latini che appartengono a diverse declinazioni.

	ROSA _F [*] 'rosa' (I decl.)	LUPUS _M 'lupo' (II)	PUPPIS _F 'poppa' (III, in -i)	HOMO _M 'uomo' (III, in cons.)	IECUR _N 'fegato' (III, in cons.)
NOM.SG	<i>ros-a</i>	<i>lup-us</i>	<i>pupp-is</i>	<i>homo-ø</i>	<i>iecur-ø</i>
GEN.SG	<i>ros-ae</i>	<i>lup-ī</i>	<i>pupp-is</i>	<i>homin-is</i>	<i>iecinor-is/iecor-is</i>
ACC.SG	<i>ros-am</i>	<i>lup-um</i>	<i>pupp-im</i>	<i>homin-em</i>	<i>iecur-ø</i>
ABL.SG	<i>ros-ā</i>	<i>lup-ō</i>	<i>pupp-ī</i>	<i>homin-e</i>	<i>iecinor-e/iecor-e</i>
GEN.PL	<i>ros-ārum</i>	<i>lup-ōrum</i>	<i>pupp-ium</i>	<i>homin-um</i>	<i>iecinor-um/iecor-um</i>

Tabella 1. Alcune celle del paradigma di nomi latini di diverse declinazioni

* In questo lavoro, si userà il maiuscolo per i lessemi e il corsivo per le forme, seguendo la convenzione comune nei lavori di morfologia flessiva.

Si può osservare che la I, la II e la III declinazione utilizzano esponenti diversi in molte celle del paradigma, ad es. rispettivamente *-ae*, *-ī* e *-is* come esponenti di GEN.SG². Anche lessemi che sono tradizionalmente considerati come appartenenti alla stessa declinazione presentano a volte diversi esponenti in alcune celle, dando luogo a sottoclassi all'interno della classe flessiva in questione³: ad esempio, i nomi della III declinazione con tema in *-i-* marcano ABL.SG con *-ī* e GEN.PL con *-ium*, mentre quelli con tema in consonante nelle stesse celle utilizzano le desinenze *-e* e *-um*, rispettivamente. Inoltre, nomi maschili e femminili si differenziano dai neutri della stessa declinazione nelle forme flesse utilizzate per alcuni casi diretti: ad es., i nomi maschili e femminili della III hanno ACC.SG in *-em* o *-im* a seconda della sottoclasse, mentre i neutri presentano sistematicamente sincretismo con NOM.SG in tale cella. Infine, talvolta in diverse celle del paradigma compaiono temi diversi e non prevedibili sistematicamente l'uno dall'altro: si veda l'alternanza tra *homo-* in NOM.SG e *homin-* nelle altre celle per il lessema HOMO e – ancor più chiaramente,

1 L'assegnazione dei lessemi alla propria classe flessiva non è di solito essa stessa del tutto prevedibile sulla base di altri fattori (ad esempio, fonologici, sintattici, semantici o pragmatici; cfr. Corbett 2009: 6–8), sebbene tali fattori spesso influiscano almeno sulla probabilità di assegnazione della classe flessiva (cfr. Guzmán Naranjo 2019).

2 Per la notazione dei valori morfosintattici delle celle del paradigma, saranno adottate le convenzioni delle *Leipzig Glossing Rules*.

3 A questo proposito, si rimanda alla distinzione tra “macro-classi” e “micro-classi” flessive, introdotta da Dressler (2002) e recentemente riproposta da Beniamine, Bonami e Sagot (2017).

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

vista la natura suppletiva dell'alternanza– quella tra *iecur-* in NOM.SG e ACC.SG e *iec(in)or-* nelle altre celle per il lessema *IECUR*.

Tuttavia, è stato osservato che l'imprevedibilità relativa al contenuto di una cella è notevolmente ridotta se le forme che occupano altre celle sono note. Sempre restando nell'ambito della flessione nominale del latino, Wurzel (1984) ha individuato due catene di quelle che ha definito "relazioni implicative" tra le desinenze di diverse sottoclassi della III declinazione. Tali catene di relazioni implicative sono riportate ai punti (1) e (2), e il modo in cui vengono usate per rendere conto della struttura delle sottoclassi della III declinazione in latino è illustrato nella Tabella 2.

(1) *-im*, ACC.SG → *-ī*, ABL.SG → *-īs*, ACC.PL → *-ium*, GEN.PL

(2) *-um*, GEN.PL → *-ēs*, ACC.PL → *-e*, ABL.SG → *-em*, ACC.SG

	PUPPIS 'poppa'	IGNIS 'fuoco'	AURIS 'orecchio'	CIVIS 'cittadino'	HOMO 'uomo'
	(in <i>-i</i>)	(mista 1)	(mista 2)	(mista 3)	(in cons.)
ACC.SG	<i>pupp-im</i>	<i>ign-em</i>	<i>aur-em</i>	<i>civ-em</i>	<i>homin-em</i>
	↓		↑	↑	↑
ABL.SG	<i>pupp-ī</i>	<i>ign-ī</i>	<i>aur-e</i>	<i>civ-e</i>	<i>homin-e</i>
	↓	↓		↑	↑
ACC.PL	<i>pupp-īs</i>	<i>ign-īs</i>	<i>aur-īs</i>	<i>civ-ēs</i>	<i>homin-ēs</i>
	↓	↓	↓		↑
GEN.PL	<i>pupp-ium</i>	<i>ign-ium</i>	<i>aur-ium</i>	<i>civ-ium</i>	<i>homin-um</i>

Tabella 2. Alcune celle del paradigma di nomi latini di diverse sottoclassi della III declinazione

Malgrado la presenza di molte possibili combinazioni di esponenti (le cinque possibilità esemplificate dai nomi riportati nella Tabella 2), le generalizzazioni espresse in (1) e in (2) permettono di ridurre notevolmente l'incertezza sulla flessione di un nome. Ad esempio, se un nome ha ACC.SG in *-im*, si può prevedere con certezza che avrà ABL.SG in *-ī*, ACC.PL in *-īs* e GEN.PL in *-ium* (si tratterà quindi di un nome con tema in *-i-*); nella direzione opposta, se un nome ha GEN.PL in *-um*, non potrà che avere ACC.PL in *-ēs*, ABL.SG in *-e* e ACC.SG in *-em* (si tratterà quindi di un tema in consonante). Per quanto riguarda le classi miste, la semplificazione introdotta dalle catene di relazioni implicative proposte è più ridotta, ma comunque presente, come illustrato nella Tabella 2.

Sulla base di generalizzazioni di questo tipo, sono stati proposti anche principi di struttura paradigmatica più generali, quali il "Paradigm Economy Principle" e il "No-blur Principle" (Carstairs 1987; Carstairs Mc-Carthy 1994). Tralasciando i dettagli della loro formulazione (per cui si rimanda alle definizioni fornite nei lavori citati) per via della loro complessità, quello che importa sottolineare in questa sede è che anche questi principi si fondano sull'osservazione che la presenza di determinate desinenze in alcune celle riduce il numero delle desinenze ammissibili in altre celle.

Fino a questo punto, si sono prese in considerazione solo relazioni implicative riguardanti le desinenze. Tuttavia, analoghe generalizzazioni possono essere formulate anche riguardo ai temi che compaiono in diverse celle del paradigma. Infatti, esistono insiemi di celle all'interno delle quali compaiono sistematicamente gli stessi allomorfi tematici, e il lavoro pionieristico di Aronoff (1994) ha messo in luce che tali insiemi di celle sono a volte puramente morfologici, in quanto non possono essere individuati sulla base di fattori semantici o fonologici: si parla in questo caso di "morfomi". Una serie di studi di Maiden –in questa sede citeremo solo il primo (Maiden 1992), e la recente monografia che riassume i risultati della quasi trentennale opera di ricerca in questa direzione (Maiden 2018)– ha portato all'individuazione di schemi paradigmatici ricorrenti a carattere morfomico nella flessione verbale di molte lingue romanze.

Ad esempio, nella Tabella 3 sono riportati quattro verbi italiani che presentano alternanze tematiche molto diverse dal punto di vista formale: quella tra forme dittongate e non dittongate nel verbo SEDERE, quella tra forme in *-sc-* e forme semplici nel verbo FINIRE, il suppletivismo debole tra forme in *e-* e forme in *u-* nel verbo USCIRE, il suppletivismo forte tra forme in *va-* e forme in *and-* nel verbo ANDARE. La distribuzione degli allomorfi è però la stessa per tutti e quattro i verbi, con una variante che compare nella prima e nella seconda persona plurale del presente indicativo⁴ e l'altra che occupa le altre celle, dando luogo a quello che Maiden chiama "N-pattern"⁵ (Maiden 2018: 167–246).

	PRS.IND					
	1SG	2SG	3SG	1PL	2PL	3PL
SEDERE	<i>siedo</i>	<i>siedi</i>	<i>siede</i>	<i>sediamo</i>	<i>sedete</i>	<i>siedono</i>
ANDARE	<i>vado</i>	<i>vai</i>	<i>va</i>	<i>andiamo</i>	<i>andate</i>	<i>vanno</i>
USCIRE	<i>esco</i>	<i>esci</i>	<i>esce</i>	<i>usciamo</i>	<i>uscite</i>	<i>escono</i>
FINIRE	<i>finisco</i>	<i>finisci</i>	<i>finisce</i>	<i>finiamo</i>	<i>finite</i>	<i>finiscono</i>

Tabella 3. Distribuzione degli allomorfi tematici nel presente indicativo di alcuni verbi italiani

Sulla base di osservazioni simili, a partire dall'inizio del nuovo millennio sono state proposte analisi più dettagliate relative a singole lingue romanze, fondate sull'individuazione di un sistema di zone paradigmatiche caratterizzate dal fatto di condividere sistematicamente uno stesso allomorfo tematico per un dato lessema. Tra i vari studi, citiamo in questa sede quello di Bonami e Boyé (2003) per il francese, quello di Boyé e Cabredo Hofherr (2006) per lo spagnolo, quello di Guerrero (2014) per il catalano. Per quanto riguarda l'italiano, i lavori fondamentali sono quelli di Pirrelli (2000) e Pirrelli e Battista (2000) –in cui tali zone paradigmatiche sono chiamate "classi di partizione"– e quelli di Montermini e Boyé (2012) e Montermini e Bonami (2013) – che usano il termine di "spazio tematico", rifacendosi alla tradizione francese in cui si inseriscono, cfr. il già citato lavoro di Bonami e Boyé (2003).

4 Un'analoga distribuzione si riscontra nelle celle del congiuntivo, omesse nella tabella per brevità.

5 Per via della somiglianza tra la forma disegnata nel paradigma dalle celle interessate e la sequenza per indicare la lettera <n> in codice Morse, costituita da una linea seguita da un punto.

È importante notare che le analisi di cui abbiamo trattato fino a questo punto sono fondate sulla presenza di relazioni implicative che riguardano temi e desinenze, e di conseguenza sulla possibilità di individuare temi e desinenze all'interno delle forme flesse: l'approccio adottato può quindi essere definito "costruttivo", nella terminologia proposta da Blevins (2006, 2016), in quanto il punto di partenza dell'analisi è costituito dai morfî, tramite cui vengono assemblate le forme flesse nella loro interezza. Tuttavia, la segmentazione delle forme flesse in temi e desinenze non è un compito facile, per cui si possa arrivare a una soluzione univoca. A tale scopo possono infatti essere utilizzate diverse strategie. Ci sono ad esempio elementi (come le vocali tematiche nelle lingue romanze) per cui è difficile decidere se vadano considerati parte del tema –seguendo un principio che Loporcaro (2012) definisce "Maximize Stem"– o della desinenza –seguendo invece il principio "Maximize Ending". Sulla base di un'intuizione simile, ma con una formulazione più generale, Bonami e Beniamine (2021) distinguono tra il principio dello "unique discontinuous stem", secondo cui vengono considerati come parte del tema solo i segmenti che compaiono in tutte le celle del paradigma, dando pertanto luogo, per ciascun lessema, a un unico tema, che però sarà spesso discontinuo (ad esempio, sulla base dei dati della Tabella 3, il tema del verbo SEDERE sarebbe *s_ed-*); e quello dei "sets of continuous stems", secondo cui il tema deve essere composto di segmenti adiacenti, e di conseguenza per alcuni lessemi sarà necessario individuare diversi temi (ad esempio, per il verbo SEDERE avremo un tema *sied-* e un tema *sed-*).

Per ovviare a simili problemi, molti lavori recenti di morfologia flessiva (cfr., tra gli altri, Bonami e Boyé 2014, Beniamine 2018, Pellegrini 2023), adottano un approccio "astrattivo", nella terminologia dei già citati lavori di Blevins, usando cioè come punto di partenza dell'analisi direttamente le forme flesse, senza assumere una segmentazione *a priori*, ma casomai inferendola *a posteriori* dai dati sulla base di un processo di astrazione.

L'obiettivo di questo lavoro è fornire un'analisi puramente astrattiva della prevedibilità tra forme flesse nei paradigmi verbali dell'italiano, e confrontarla in maniera sistematica con le analisi basate sui temi –e quindi almeno parzialmente costruttive– proposte nei già citati lavori passati. A questo scopo, verranno innanzitutto ripetuti i punti fondamentali delle precedenti analisi basate sui temi (Sezione 2). Si passerà quindi alla descrizione dell'analisi astrattiva proposta (Sezione 3), illustrando il metodo adottato (3.1), i dati e gli strumenti utilizzati (3.2), e i risultati ottenuti (3.3). Le diverse analisi verranno poi messe a confronto in maniera sistematica (Sezione 4), rendendo conto delle differenze tra di esse (4.1) –in maniera analoga a quanto fatto da Bonami e Boyé (2014) per il francese– e proponendo poi un metodo per quantificare la loro efficacia nel catturare le generalizzazioni presenti nei dati (4.2). Chiude il lavoro l'esposizione delle conclusioni che si possono trarre dai risultati ottenuti (Sezione 5).

2. Prevedibilità tra temi: lo spazio tematico del verbo italiano

Per poter confrontare in maniera sistematica l'analisi proposta in questo lavoro, basata solo sulle forme flesse, con quelle precedenti basate sui temi, è necessario passare brevemente in rassegna i punti fondamentali di queste ultime.

Come si è accennato nella Sezione 1, i primi lavori a offrire una descrizione dettagliata della distribuzione paradigmatica degli allomorfi tematici nella flessione verbale dell'italiano sono

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

stati quelli di Pirrelli (2000) e Pirrelli e Battista (2000). L'esito dell'analisi da loro condotta può essere riassunto in una mappatura dello schema distribuzionale complessivo della coniugazione italiana, in cui diverse celle del paradigma fanno parte della stessa "classe di partizione" se tali celle contengono sistematicamente la stessa base tematica del verbo in questione⁶. Pertanto, tale mappatura può essere interpretata come un'astrazione della massima complessità raggiungibile da un verbo in italiano, almeno limitatamente all'allomorfia tematica⁷. I successivi lavori di Montermini e Boyé (2012) e Montermini e Bonami (2013) delineano un quadro simile della struttura paradigmatica dei verbi italiani, in quello che chiamano lo "spazio tematico" del verbo italiano, illustrato nella Tabella 4 nella sua versione più recente, quella di Montermini e Bonami (2013), che verrà usata come riferimento nel resto di questo lavoro.

	1SG	2SG	3SG	1PL	2PL	3PL		M.SG	F.SG	M.PL	F.PL
FUT.IND	S6	S6	S6	S6	S6	S6	PST.PTCP	S7	S7	S7	S7
PRS.COND	S6	S6	S6	S6	S6	S6	PRS.PTCP	S1			
PRS.SBJV	S2	S2	S2	S4	S4	S2	GER	S1			
PRS.IND	S2	S3	S3	S4	S1	S2	INF	S8			
IPRF.IND	S1	S1	S1	S1	S1	S1					
IPRF.SBJV	S1	S1	S1	S1	S1	S1					
PRET.IND	S5	S1	S5	S1	S1	S5					
IMP		S3		S4	S1						

Tabella 4. Lo spazio tematico del verbo italiano nella versione di Montermini & Bonami (2013)⁸

Inoltre, in questi ultimi lavori vengono espresse ulteriori generalizzazioni (i) su come ottenere un tema a partire da un altro e (ii) su come ottenere le varie forme flesse a partire dai temi. Per quanto riguarda il primo punto, vengono specificate relazioni formali di default tra i diversi temi, chiamate "stem-to-stem relations" da Montermini e Bonami (2013: 181-186), tradotto in "relazioni tema-tema" in questa sede. Nei verbi regolari della I coniugazione, come LAVARE, tutti i temi possono essere ricavati sulla base di tali relazioni di default, come illustrato nella Figura 1 e nella Figura 2. Nei verbi che invece presentano un certo grado di irregolarità, ad es. CRESCERE, alcuni temi non possono essere ricavati in questo modo, e devono quindi essere memorizzati come tali nel lessico, restando isolati come illustrato nella Figura 3.

Per quanto riguarda il secondo punto, vengono specificate relazioni formali tra temi e forme, che forniscono istruzioni su come ottenere le seconde a partire dai primi: tali relazioni sono chiamate "stem-to-form relations" da Montermini e Bonami (2013: 180-186), tradotto in "relazioni tema-forma" in questa sede.

6 Eccezion fatta per otto verbi altamente irregolari: ANDARE, AVERE, DARE, DIRE, ESSERE, FARE, SAPERE, STARE. Su questo aspetto, cfr. anche sotto alla Sezione 4.1.

7 Come fa notare Pirrelli (2000: 75), non esiste in realtà alcun verbo italiano che concretamente manifesti un tale grado di complessità: quello che conta un maggior numero di basi tematiche distinte (6) è DOLERE.

8 La "S" che precede l'indice di ogni zona sta per la parola inglese "Stem" 'tema'.

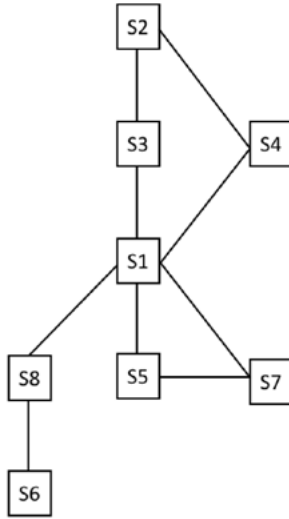


Figura 1. Grafo delle relazioni tema-tema dei verbi regolari della I coniugazione

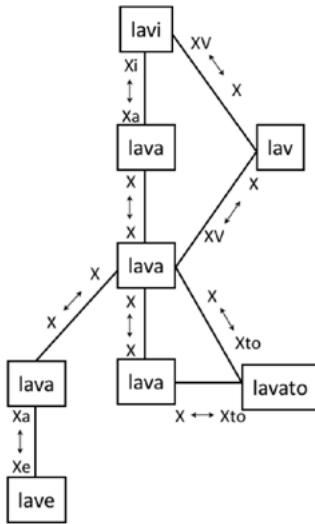


Figura 2. Grafo delle relazioni tema-tema del verbo LAVARE

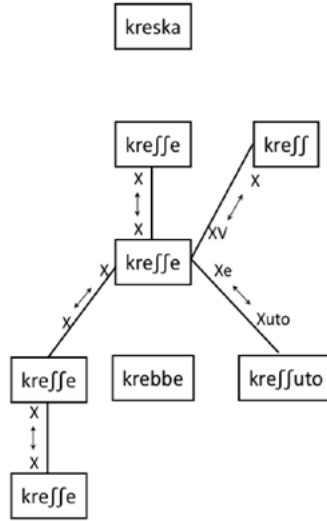


Figura 3. Grafo delle relazioni tema-tema del verbo CRESCERE

Le relazioni di default –sia quelle tema-tema, sia quelle tema-forma– possono essere valide per tutti i verbi, oppure avere una copertura più limitata e valere solo per verbi di determinate coniugazioni. Per chiarire il meccanismo che permette di generare le forme flesse in questo quadro, la Tabella 5 e la Tabella 6 mostrano le relazioni tema-tema e le relazioni tema-forma necessarie per rendere conto delle forme del presente indicativo di verbi appartenenti a diverse coniugazioni.

I diversi temi possono essere ottenuti l'uno dall'altro sulla base delle relazioni formali elencate nella Tabella 5. Ad esempio, il tema S4 può essere ricavato rimuovendo la vocale finale di S1, S2 o S3, indipendentemente dalla coniugazione, mentre il tema S2 può essere ricavato a partire da S1 e S3 sostituendo la vocale finale del tema con /i/ se il verbo è della I coniugazione, con /a/ se il verbo è della II o della III coniugazione, oppure a partire da S4 aggiungendo le rispettive vocali al tema.

temi	S1-S2	S1-S3	S1-S4	S2-S3	S2-S4	S3-S4
I	Xa-Xi	X-X	Xa-X	Xi-Xa	Xi-X	Xa-X
relazione II	Xe-Xa	X-X	Xe-X	Xa-Xe	Xa-X	Xe-X
III	Xi-Xa	Xi-Xe	Xi-X	Xa-Xe	Xa-X	Xe-X
LAVARE (I)	S1 /lava/ ↔ S2 /lavi/	S1 /lava/ ↔ S3 /lava/	S1 /lava/ ↔ S4 /lav/	S2 /lavi/ ↔ S3 /lava/	S2 /lavi/ ↔ S4 /lav/	S3 /lava/ ↔ S4 /lav/
TEMERE (II)	S1 /teme/ ↔ S2 /tema/	S1 /teme/ ↔ S3 /teme/	S1 /teme/ ↔ S4 /tem/	S2 /tema/ ↔ S3 /teme/	S2 /tema/ ↔ S4 /tem/	S3 /teme/ ↔ S4 /tem/
DORMIRE (III)	S1 /dormi/ ↔ S2 /dorma/	S1 /dormi/ ↔ S3 /dorme/	S1 /dormi/ ↔ S4 /dorm/	S2 /dorma/ ↔ S3 /dorme/	S2 /dorma/ ↔ S4 /dorm/	S3 /dorme/ ↔ S4 /dorm/

Tabella 5. Relazioni tema-tema del presente indicativo di verbi italiani di diverse coniugazioni

Le varie forme flesse possono essere costruite a partire dai rispettivi temi sulla base delle relazioni tema-forma elencate nella Tabella 6. Ad esempio, la prima persona plurale del presente indicativo si ottiene aggiungendo il segmento /jamo/ al tema S4 del verbo in questione (rispettivamente /lav/, /tem/, /dorm/ negli esempi forniti), indipendentemente dalla sua coniugazione. Per ottenere la terza plurale, invece, la vocale finale del tema S2 (rispettivamente /lavi/, /tema/, /dorma/) deve essere sostituita dal segmento /ano/ per i verbi della I coniugazione, /ono/ per i verbi della II e della III coniugazione.

	PRS.IND					
	1SG	2SG	3SG	1PL	2PL	3PL
tema	S2	S3	S3	S4	S1	S2
relazione	XV-Xo	XV-Xi	X-X	X-Xjamo	X-Xte	Xi-Xano Xa-Xono
LAVARE (I)	S2 /lavi/ → 1SG /lavo/	S3 /lava/ → 2SG /lavi/	S3 /lava/ → 3SG /lava/	S4 /lav/ → 1PL /lavjamo/	S1 /lava/ → 2PL /lavate/	S2 /lavi/ → 3PL /lavano/
TEMERE (II)	S2 /tema/ → 1SG /temo/	S3 /teme/ → 2SG /temi/	S3 /teme/ → 3SG /teme/	S4 /tem/ → 1PL /temjamo/	S1 /teme/ → 2PL /temete/	S2 /tema/ → 3PL /temono/
DORMIRE (III)	S2 /dorma/ → 1SG /dormo/	S3 /dorme/ → 2SG /dormi/	S3 /dorme/ → 3SG /dorme/	S4 /dorm/ → 1PL /dormjamo/	S1 /dormi/ → 2PL /dormite/	S2 /dorma/ → 3PL /dormono/

Tabella 6. Relazioni tema-forma del presente indicativo di verbi italiani di diverse coniugazioni

In tal modo, per un verbo regolare come LAVARE, basta conoscere un tema per essere in grado di riempire tutte le altre celle del paradigma, e anche per verbi parzialmente irregolari come CRESCERE un numero limitato di temi è sufficiente, permettendo quindi di rendere conto in maniera economica dei complessi *pattern* di allomorfa tematica presenti nel sistema verbale dell'italiano, oltre che delle diverse classi flessive.

3. Prevedibilità tra forme: un'analisi basata sull'entropia

3.1 Il metodo

Nel corso dell'ultimo decennio, sono stati proposti diversi modi per quantificare in maniera sistematica la prevedibilità delle relazioni implicative tra celle del paradigma.

Diversi lavori di Stump e Finkel (cfr., tra gli altri, Stump e Finkel 2013) descrivono un metodo fondato sulla teoria degli insiemi, e in particolare sulla nozione di "parti principali", vale a dire, un insieme di celle del paradigma a partire dalle quali è possibile prevedere tutte le altre forme di un lessema senza alcuna incertezza. Questa nozione è del resto ben attestata nella tradizione grammaticale delle lingue classiche. Nella pratica lessicografica del latino, ad esempio, le entrate lessicali dei verbi (come FERRE 'portare') elencano, oltre alla forma di citazione (PRS.ACT.IND.1SG, *ferō*), altre forme (nell'*Oxford Latin Dictionary* troviamo PRS.ACT.INF *ferre*, PRF.ACT.IND.1SG *tuli*,

SUP.ACC *latum*) che permettono di inferire il contenuto delle rimanenti celle del paradigma, anche quando vi siano fenomeni di allomorfia o suppletivismo (come nell'esempio appena portato). Gli strumenti sviluppati da Stump e Finkel (2013) permettono l'estrazione automatica di insiemi di parti principali e il calcolo di diverse misure di prevedibilità e predittività delle celle a partire dagli esponenti delle diverse celle del paradigma in lessemi di diverse classi flessive.

A partire da Ackerman, Blevins e Malouf (2009), per rendere conto di fatti simili è stata utilizzata un'altra strategia, basata sulla teoria dell'informazione e in particolare sulla nozione di entropia condizionale. In generale, l'entropia $-H(x)$ è una misura dell'incertezza sul valore assunto da una determinata variabile x ; più specificamente, l'entropia condizionale $-H(x|y)$ misura l'incertezza sul valore di una variabile x dato il valore di un'altra variabile y . Ackerman, Blevins e Malouf (2009) usano quest'ultima misura per quantificare l'incertezza nel prevedere la desinenza che compare in una cella conoscendo quella che compare in un'altra cella. In una serie di lavori recenti (cfr., tra gli altri, Bonami e Boyé 2014 e Beniamine 2018), questo metodo è stato modificato in modo da permettere di quantificare l'incertezza nel prevedere l'intera forma flessa che occupa una cella conoscendo quella che ne occupa un'altra, adottando quindi un approccio puramente astrattivo. È questo il metodo che verrà adottato anche in questo lavoro: pertanto, è necessario illustrarne brevemente i dettagli, in modo da chiarire i principi soggiacenti all'analisi che verrà svolta nelle prossime sezioni.

Si considerino i dati della Tabella 7.

verbo (coniugazione)	GER	PRS.IND.2PL	alternanza/contesto GER ↔ PRS.IND.2PL	alternanze applicabili	n. verbi
AMARE (I)	a'mando	a'ma:te	1 (_ando ↔ _a:te / C_#)	A (1)	2.085
VEDERE (II)	ve'dendo	ve'de:te	2 (_endo ↔ _e:te / C_#)	B (2,3)	355
SENTIRE (III)	sen'tendo	sen'ti:te	3 (_endo ↔ _i:te / C_#)	B (2,3)	287

Tabella 7. Classificazioni necessarie per il calcolo dell'entropia condizionale di GER dato PRS.IND.2PL

In un approccio basato esclusivamente sulle forme, il punto di partenza dell'analisi sono le forme che occupano le celle oggetto di indagine: in questo caso, il gerundio e la seconda persona plurale del presente indicativo, di cui vengono riportati nella seconda e nella terza colonna alcuni esempi relativi a verbi di diverse coniugazioni. Il primo passaggio della procedura consiste nel confrontare le diverse forme di uno stesso lessema e estrarre il pattern di alternanza formale tra le due forme, che può essere considerato come l'istruzione necessaria per ottenere una forma a partire dall'altra, e il contesto in cui tale alternanza si applica; in questo esempio, per semplicità consideriamo i diversi pattern come applicabili dopo consonante. Come si può osservare nella quarta colonna della Tabella 7, tali pattern permettono di distinguere tra le tre coniugazioni prese in considerazione in questo esempio: pertanto, possono essere considerati come in grado di riassumere il comportamento flessivo del verbo in questione, e costituiscono dunque la variabile x di cui si vuole valutare l'incertezza. Il secondo passaggio consiste nel classificare i verbi del campione sulla base dei pattern potenzialmente applicabili alla forma di cui si assume la conoscenza pregressa –GER– per prevedere quella che si vuole ottenere – PRS.IND.2PL– sulla

base della forma segmentale dei pattern estratti nella fase precedente e dei loro contesti di applicazione. Nell'esempio fornito, a un verbo come AMARE può essere applicato solo il pattern 1, dal momento che il gerundio termina in /ando/. Quanto a verbi come VEDERE e SENTIRE, il cui gerundio termina in /endo/, potenzialmente sia il pattern 2 sia il pattern 3 possono essere applicati: di conseguenza, tali verbi sono assegnati a una stessa classe (B), distinta da quella dei verbi come AMARE (A) –si veda la quinta colonna della Tabella 7. La classificazione ottenuta in questo secondo passaggio viene utilizzata come la variabile y di cui si dà per noto il valore nel calcolo dell'entropia condizionale. Sulla base di queste informazioni, dato un verbo il cui gerundio esce in /ando/, il PRS.IND.2PL non potrà che uscire in /a:te/, dal momento che soltanto un pattern (1) è disponibile per i verbi di questa classe (A). Dato invece un verbo il cui gerundio esce in /endo/, i pattern disponibili sono due (2 e 3): di conseguenza, ci sarà incertezza nel prevedere se il PRS.IND.2PL uscirà in /e:te/ o /i:te/.

Per quantificare il grado di difficoltà in maniera più accurata, è utile tenere in considerazione anche la frequenza dei diversi pattern e delle diverse classi nel lessico, come mostrato nella Tabella 8⁹.

x	$P(x)$	y	$P(y x)$
A	$\frac{2.085}{2.727}$	1	$\frac{2.085}{2.085}$
		2	$\frac{355}{642}$
B	$\frac{642}{2.727}$	3	$\frac{287}{642}$

Tabella 8. Variabili e loro probabilità per il calcolo dell'entropia condizionale di GER dato PRS.IND.2PL

I pattern 2 e 3 si applicano a un numero di verbi simile: di conseguenza, l'incertezza introdotta nel sistema sarà maggiore di quella che si avrebbe se uno dei due pattern fosse applicato solo a pochi verbi, nel qual caso l'applicazione dell'altro potrebbe essere considerata molto più probabile. Tuttavia, i verbi della II e della III coniugazione sono nel complesso molto meno frequenti di quelli della I coniugazione: di conseguenza, l'impatto dell'incertezza introdotta nel sistema dai verbi della II e della III sarà ridotto –dal momento che interessa una porzione più limitata del lessico verbale dell'italiano¹⁰.

I dettagli del calcolo dell'entropia condizionale per l'esempio in questione sono riassunti nella formula riportata al punto (3), il cui risultato finale costituisce una stima dell'incertezza complessiva nel prevedere la seconda persona plurale del presente indicativo di un verbo in italiano conoscendone il gerundio.

⁹ I dati sono tratti dal LeFFI, su cui cfr. sotto, Sezione 3.2.

¹⁰ Si fa dunque riferimento alla frequenza di *type*, piuttosto che a quella di *token*, in linea con l'osservazione di Bybee (1995: 433–435) che forme con alta frequenza di *token* spesso non contribuiscono all'estensione analogica dei pattern irregolari che eventualmente presentano, dal momento che sono più facilmente memorizzate che non prodotte tramite l'applicazione di tali pattern.

$$(3) \quad H(y|x) = - \left[\frac{2.085}{2.727} \times \left(\frac{2.085}{2.085} \times \log_2 \frac{2.085}{2.085} \right) + \frac{642}{2.727} \times \left(\frac{355}{642} \times \log_2 \frac{355}{642} + \frac{287}{642} \times \log_2 \frac{287}{642} \right) \right] = 0,234$$

3.2 Dati e strumenti

Nel seguito di questo lavoro, la metodologia appena descritta verrà applicata sistematicamente a tutte le possibili combinazioni di celle del paradigma verbale dell'italiano. Per farlo, è necessario innanzitutto disporre di un lessico che elenchi tutte le forme flesse di un campione sufficientemente ampio e rappresentativo di verbi. A questo scopo, verrà usato il LeFFI (Lessico di Forme Flesse dell'Italiano; cfr. Pellegrini e Cignarella 2020)¹¹. Questa risorsa è stata costruita a partire da un altro lessico, il GLAFF-IT (Calderone et al. 2017), da cui sono state estratte le forme flesse in trascrizione fonetica. È stata poi effettuata una selezione sulla base della frequenza dei lessemi, utilizzando i dati di un'altra risorsa, il COLFIS (Bertinetto et al. 2005), prima di procedere a una correzione manuale degli errori e delle incoerenze presenti nelle forme estratte dal GLAFF-IT. Rispetto alla versione presentata in Pellegrini e Cignarella (2020), che conteneva i 2.053 verbi con frequenza maggiore di 10 nel COLFIS, in quella utilizzata per questo lavoro (v1.1¹²) la soglia è stata innalzata a 5, arrivando dunque a un campione di 2.744 verbi. È questa la ragione delle differenze di dettaglio tra i risultati dell'analisi svolta in questo lavoro e quelli preliminari riportati da Pellegrini e Cignarella (2020: Tabelle 3, 4, *passim*), pur restando invariato il quadro generale e le conclusioni che se ne possono trarre.

Per quanto riguarda la modalità di costruzione di LeFFI, si rimanda a Pellegrini e Cignarella (2020: §2). In questa sede, ci limitiamo a riportare nella Tabella 9 una porzione della risorsa, nel formato in cui è rilasciata nella versione in questione, ovvero, con una colonna per ciascuna delle celle del paradigma verbale dell'italiano e una riga per ciascuno dei verbi del campione. Per ragioni tecniche di semplificazione dell'analisi, l'accento è marcato non all'inizio della sillaba su cui si colloca, ma subito prima della vocale che ne costituisce il nucleo. Si noti inoltre che nel LeFFI non è codificata la distinzione tra vocali medio-alte e medio-basse (/e/ vs. /ɛ/, /o/ vs. /ɔ/) propria dell'italiano standard, rispecchiando piuttosto le tante varietà di italiano in cui tale distinzione non ha carattere fonologico. Nello specifico, vengono utilizzate soltanto le vocali medio-alte (si veda sotto la Tabella 10).

11 <https://github.com/matteo-pellegrini/LeFFI>

12 <https://github.com/matteo-pellegrini/LeFFI/releases/tag/v1.1>

	INF	GER	IMP.1PL	IMP.2PL	PRS.IND.1SG	PRS.SBJV.3SG
ESSERE	'essere	ess'endo	sj'amo	sj'ate	s'ono	s'ia
FARE	f'are	fatʃ'endo	fattʃ'amo	f'ate	f'attʃo	f'attʃa
AVERE	av'ere	av'endo	abbj'amo	abbj'ate	'o	'abbja
MANGIARE	mandʒ'are	mandʒ'ando	mandʒ'amo	mandʒ'ate	m'andʒo	m'andʒi
DORMIRE	dorm'ire	dorm'endo	dormj'amo	dorm'ite	d'ormo	d'orma
SORRIDERE	sorr'idere	sorrid'endo	sorridj'amo	sorrid'ete	sorr'ido	sorr'ida
BACIARE	batʃ'are	batʃ'ando	batʃ'amo	batʃ'ate	b'atʃo	b'atʃi

Tabella 9. Una porzione del LeFFI

A partire da questi dati, è possibile calcolare automaticamente i valori di entropia condizionale che quantificano l'incertezza nel prevedere il contenuto di una cella a partire da quello di un'altra per tutte le possibili combinazioni di celle, utilizzando Qumin, un insieme di programmi in Python per l'analisi quantitativa dei sistemi flessivi (Quantitative modelling of inflection; cfr. Beniamine 2018). Questo strumento consente di estrarre automaticamente i pattern di alternanza formale e la classificazione sulla base delle alternanze potenzialmente applicabili, come illustrato nella Tabella 7, e di effettuare il calcolo dell'entropia, come esemplificato al punto (3). Di nuovo, si rimanda chi legge alla documentazione¹³ e alla pubblicazione originale (Beniamine 2018) per ulteriori dettagli sul funzionamento del programma. È però importante sottolineare che, per ricavare generalizzazioni sul contesto di applicazione dei pattern, è necessario che ciascuno dei segmenti che compaiono nei dati di input sia analizzato nei tratti fonologici che lo caratterizzano. L'analisi adottata per l'italiano ai fini di questo lavoro è riportata nella Tabella 10.

13 <https://qumin.readthedocs.io/>

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

Seg.	cons	son	approx	voice	cont	nas	lat	LAB	LAB DENT	COR	anter	DOR	high	low	back	rnd	acc
b	1	0	0	1	0	0	0	1	0	0	-1	0	0	0	0	0	-1
d	1	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	-1
g	1	0	0	1	0	0	0	0	0	0	-1	1	1	0	1	0	-1
v	1	0	0	1	1	0	0	1	1	0	-1	0	0	0	0	0	-1
m	1	1	0	1	0	1	0	1	0	0	-1	0	0	0	0	0	-1
ɱ	1	1	0	1	0	1	0	1	1	0	-1	0	0	0	0	0	-1
n	1	1	0	1	0	1	0	0	0	1	1	0	0	0	0	0	-1
ɲ	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	-1
ŋ	1	1	0	1	0	1	0	0	0	0	-1	1	1	0	1	0	-1
l	1	1	1	1	0	0	1	0	0	1	1	0	0	0	0	0	-1
ʎ	1	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0	-1
r	1	1	1	1	1	0	0	0	0	1	1	0	0	0	0	0	-1
ʒ	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	-1
z	1	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	-1
dʒ	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	-1
dz	1	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	-1
p	1	0	0	0	0	0	0	1	0	0	-1	0	0	0	0	0	-1
t	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	-1
k	1	0	0	0	0	0	0	0	0	0	-1	1	1	0	1	0	-1
f	1	0	0	0	1	0	0	1	1	0	-1	0	0	0	0	0	-1
ʃ	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	-1
s	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	-1
tʃ	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	-1
ts	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	-1
j	0	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	-1
w	0	1	1	1	1	0	0	1	0	0	-1	1	1	0	1	1	-1

Seg.	cons	son	approx	voice	cont	nas	lat	LAB	LAB DENT	COR	anter	DOR	high	low	back	rnd	acc
a	0	1	1	1	1	0	0	0	0	-1	-1	1	0	1	1	0	0
e	0	1	1	1	1	0	0	0	0	-1	-1	1	0	0	0	0	0
i	0	1	1	1	1	0	0	0	0	-1	-1	1	1	0	0	0	0
o	0	1	1	1	1	0	0	1	0	-1	-1	1	0	0	1	1	0
u	0	1	1	1	1	0	0	1	0	-1	-1	1	1	0	1	1	0
'a	0	1	1	1	1	0	0	0	0	-1	-1	1	0	1	1	0	1
'e	0	1	1	1	1	0	0	0	0	-1	-1	1	0	0	0	0	1
'i	0	1	1	1	1	0	0	0	0	-1	-1	1	1	0	0	0	1
'o	0	1	1	1	1	0	0	1	0	-1	-1	1	0	0	1	1	1
'u	0	1	1	1	1	0	0	1	0	-1	-1	1	1	0	1	1	1

Tabella 10. Tratti fonologici distintivi per i segmenti utilizzati in LeFFI

Legenda: cons = consonantico, son = sonorante, approx = approssimante, voice = sonoro, cont = continuo, nas = nasale, lat = laterale, LAB = labiale, LABDENT = labiodentale, COR = coronale, anter = anteriore, DOR = dorsale, high = alto, low = basso, back = posteriore, rnd = arrotondato, acc = accentato. 1 indica la presenza di un tratto, 0 la sua assenza, -1 segnala che il tratto non è rilevante per il segmento in questione.

3.3 L'analisi

Utilizzando i dati del LeFFI come input per i programmi di Qumin, si ottiene come output una matrice di valori di entropia condizionale che quantificano l'incertezza nel prevedere ogni cella del paradigma a partire da ciascuna delle altre, in entrambe le direzioni. Tale matrice è riportata nei materiali aggiuntivi online¹⁴. Nello specifico, i risultati sono strutturati in modo da riportare in ciascuna casella della tabella il valore di entropia condizionale relativo alla previsione della cella del paradigma nella colonna corrispondente a partire da quella nella riga corrispondente. Di conseguenza, relativamente a ogni cella del paradigma, i valori riportati nella rispettiva colonna forniscono indicazioni sulla prevedibilità, quelli riportati nella rispettiva riga sulla capacità predittiva.

A partire da questi risultati, è possibile ottenere una mappatura del paradigma come quella illustrata nella Tabella 11, in cui due celle A e B appartengono alla stessa zona (Z1-Z14) quando sono prevedibili l'una dall'altra senza alcun margine di incertezza, ovvero, in termini più formali, se e solo se $H(A|B) = H(B|A) = 0$.

¹⁴ <https://zenodo.org/records/13354026>.

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

	1SG	2SG	3SG	1PL	2PL	3PL		M.SG	F.SG	M.PL	F.PL
FUT.IND	Z6	Z6	Z6	Z6	Z6	Z6	PST.PTCP	Z7	Z7	Z7	Z7
PRS.COND	Z6	Z6	Z6	Z6	Z6	Z6	PRS.PTCP	Z11		Z11	
PRS.SBJV	Z2	Z2	Z2	Z4	Z4	Z2	GER	Z11			
PRS.IND	Z12	Z14	Z13	Z4	Z10	Z13	INF	Z8			
IPRF.IND	Z9	Z9	Z9	Z9	Z9	Z9					
IPRF.SBJV	Z1	Z1	Z1	Z1	Z1	Z1					
PRET.IND	Z5	Z1	Z5	Z1	Z1	Z5					
IMP		Z3		Z4	Z10						

Tabella 11. Zone paradigmatiche nella flessione verbale dell'italiano: $H(A|B) = H(B|A) = 0$

Tale mappatura è analoga nello spirito a quella proposta nei lavori passati in rassegna nella Sezione 2, dal momento che rende conto della struttura delle relazioni implicative tra le celle del paradigma verbale dell'italiano; ma se ne discosta in quanto è ottenuta tramite una procedura puramente astrattiva –quella descritta nella Sezione 3.1– che prende in considerazione direttamente le forme flesse, senza segmentarle preventivamente in temi e desinenze. È dunque interessante mettere a confronto in maniera sistematica l'analisi basata sulle forme di questo lavoro con quella basata su temi e desinenze dei lavori precedenti, per individuare eventuali differenze e indagarne le motivazioni: questo sarà l'argomento della prossima sezione.

Per il momento, invece, per semplificare la visualizzazione e l'analisi dei risultati complessivi dell'analisi proposta è utile concentrarsi su una versione ridotta del paradigma che contiene soltanto una cella per ognuna delle zone individuate: dal momento che tutte le altre possono essere ottenute a partire da questa senza alcuna incertezza, ciò permette una notevole riduzione delle dimensioni della matrice, senza perdita di informazione rilevante. La Tabella 12 mostra appunto i valori di entropia di tale versione ridotta del paradigma verbale dell'italiano. Per agevolare la lettura, sono utilizzate diverse sfumature di grigio a seconda del grado di incertezza nella previsione: più il colore è scuro, maggiore è il valore di entropia, e quindi l'incertezza.

Si può così notare che alcune celle risultano sistematicamente più facili o difficili da prevedere, oppure più o meno utili per prevedere altre celle. Ad esempio, la riga relativa alla cella IMP.1PL presenta valori di entropia notevolmente più alti delle altre: si tratta dunque di una cella con una limitata capacità predittiva. Ciò è dovuto ai fatti riassunti nella Tabella 13. Come si può notare, la forma contenuta in tale cella non permette di distinguere tra le diverse coniugazioni, dal momento che termina sempre in /j'amo/, per verbi sia della I (ad es. AMARE), sia della II (tanto quelli con infinito rizononico come SCENDERE, quanto quelli con infinito rizoatono come PERSUADERE) sia della III (ad es. VENIRE). Inoltre, tale terminazione è anche comune a verbi in cui il segmento /j/ è già presente nel tema (ad es. CAMBIARE) e a verbi in cui è introdotto dalla desinenza (come negli altri esempi appena citati). Di conseguenza, data una forma di IMP.1PL in /j'amo/, ci sono almeno cinque esiti possibili per la forma dell'infinito (in /'are/, /j'are/, /ere/, /'ere/ o /

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

	INF	GER	PST. PTCP. F.PL	IMP. 2SG	IMP. 1PL	IMP. 2PL	PRS. IND. 1SG	PRS. IND. 2SG	PRS. IND. 3PL	PRS. SBJV. 1SG	IPRF. IND. 1PL	IPRF. SBJV. 1PL	PRET. IND. 1SG	COND. 1PL
INF		0,012	0,022	0,45	0,49	0	0,44	0,43	0,4	0,42	0,012	0,02	0,03	0,032
GER	0,25		0,113	0,62	0,45	0,228	0,57	0,6	0,6	0,56	0,2314	0,24	0,1	0,235
PST.PTCP.F.PL	0,02	0,0007		0,46	0,45	0,0007	0,42	0,43	0,4	0,4	0,0007	0,01	0,01	0,014
IMP.2SG	0,15	0,0893	0,107		0,54	0,1409	0	0	0	0	0,1356	0,14	0,11	0,132
IMP.1PL	1,05	1,0091	0,863	1,14		1,1869	0,88	0,79	1	0,96	1,1893	1,2	0,84	0,763
IMP.2PL	0,05	0,012	0,024	0,45	0,46		0,43	0,43	0,4	0,42	0,012	0,02	0,04	0,035
PRS.IND.1SG	0,45	0,4389	0,205	0,47	0,54	0,3801		0,06	0,8	0,78	0,3946	0,39	0,22	0,154
PRS.IND.2SG	0,75	0,7452	0,425	1,01	0,48	0,7374	0,32		0,7	0,49	0,7527	0,76	0,41	0,38
PRS.IND.3PL	0,12	0,0852	0,103	0	0,5	0,1177	0	0		0	0,0975	0,12	0,11	0,117
PRS.SBJV.1SG	0,3	0,2623	0,279	0,23	0,36	0,2942	0,23	0	0,2		0,2948	0,29	0,28	0,296
IPRF.IND.1PL	0,05	0	0,011	0,55	0,45	0	0,44	0,43	0,5	0,42		0,01	0,03	0,033
IPRF.SBJV.1PL	0,05	0,0069	0,016	0,46	0,45	0,0069	0,45	0,45	0,4	0,43	0,0056		0,02	0,023
PRET.IND.1SG	0,03	0,0043	0,062	0,48	0,45	0,0043	0,46	0,46	0,5	0,44	0,0043	0		0,003
COND.1PL	0,41	0,4286	0,151	0,7	0,54	0,4427	0,48	0,48	0,6	0,57	0,4397	0,43	0,16	

Tabella 12. Valori di entropia condizionale per una versione ridotta del paradigma verbale dell'italiano

'ire/). Viceversa, la cella COND.1PL risulta essere più facilmente prevedibile rispetto a molte altre – i valori di entropia nella colonna corrispondente sono tendenzialmente più bassi che nel resto della matrice. Ciò è dovuto al fatto che in tale cella è neutralizzata la distinzione tra la I e la II coniugazione, oltre che quella tra le sottoclassi con infinito rizo-tonico e rizo-atono della seconda: in tutti questi casi, la forma sarà sempre in /er'emmo/, mentre soltanto nella III coniugazione la forma è distinta, in /ir'emmo/. Di conseguenza, anche a partire da una cella poco informativa – come appunto IMP.1PL – l'incertezza sarà ridotta: data una forma di IMP.1PL in /j'amo/, a fronte dei cinque esiti possibili elencati sopra per l'infinito, quelli disponibili per la cella COND.1PL sono soltanto tre (/er'emmo/, /jer'emmo/ o /ir'emmo/).

verbo	imp.1pl	inf	cond.1pl
AMARE (I)	/amj'amo/	/am'are/	/amer'emmo/
CAMBIARE (I)	/kambj'amo/	/kambj'are/	/kambjer'emmo/
SCENDERE (II)	/ʃendj'amo/	/ʃ'endere/	/scender'emmo/
PERSUADERE (II)	/perswadj'amo/	/perswad'ere/	/perswader'emmo/
VENIRE (III)	/sentj'amo/	/sent'ire/	/sentir'emmo/

Tabella 13. Le celle IMP.1PL, INF e COND.1PL di alcuni verbi italiani

4. Temi e forme a confronto

4.1 Differenze tra spazio tematico e zone paradigmatiche ricavate sulla base dell'entropia condizionale

In questa sezione, si cercherà di effettuare un confronto sistematico tra la mappatura dello spazio tematico del verbo italiano nella versione di Montermini e Bonami (2013)¹⁵, illustrata nella Tabella 4, e quella delle zone le cui celle possono essere ricavate le une dalle altre senza alcuna incertezza nella presente analisi basata sull'entropia condizionale, illustrata nella Tabella 11. Per comodità di lettura, le due tabelle sono riportate l'una vicino all'altra anche in questa sezione. Come si può osservare, la struttura del paradigma che emerge dalle due analisi presenta molte analogie, ma anche alcune significative differenze.

	1SG	2SG	3SG	1PL	2PL	3PL		M.SG	F.SG	M.PL	F.PL
FUT.IND	S6	S6	S6	S6	S6	S6	PST.PTCP	S7	S7	S7	S7
PRS.COND	S6	S6	S6	S6	S6	S6	PRS.PTCP	S1		S1	
PRS.SBJV	S2	S2	S2	S4	S4	S2	GER	S1			
PRS.IND	S2	S3	S3	S4	S1	S2	INF	S8			
IPRF.IND	S1	S1	S1	S1	S1	S1					
IPRF.SBJV	S1	S1	S1	S1	S1	S1					
PRET.IND	S5	S1	S5	S1	S1	S5					
IMP		S3		S4	S1						

Tabella 14. Lo spazio tematico del verbo italiano nella versione di Montermini & Bonami (2013)

	1SG	2SG	3SG	1PL	2PL	3PL		M.SG	F.SG	M.PL	F.PL
FUT.IND	Z6	Z6	Z6	Z6	Z6	Z6	PST.PTCP	Z7	Z7	Z7	Z7
PRS.COND	Z6	Z6	Z6	Z6	Z6	Z6	PRS.PTCP	Z11		Z11	
PRS.SBJV	Z2	Z2	Z2	Z4	Z4	Z2	GER	Z11			
PRS.IND	Z12	Z14	Z13	Z4	Z10	Z13	INF	Z8			
IPRF.IND	Z9	Z9	Z9	Z9	Z9	Z9					
IPRF.SBJV	Z1	Z1	Z1	Z1	Z1	Z1					
PRET.IND	Z5	Z1	Z5	Z1	Z1	Z5					
IMP		Z3		Z4	Z10						

Tabella 15. Zone paradigmatiche nella flessione verbale dell'italiano: $H(A|B) = H(B|A) = 0$

15 D'ora in avanti, per brevità ci riferiremo a questo lavoro come M&B.

Per quanto riguarda le analogie, si può notare che alcune zone corrispondono perfettamente a celle che presentano sempre lo stesso tema nell'analisi di M&B: è il caso di Z4, Z5, Z6, Z7 e Z8. Passando alle differenze, in generale l'analisi basata sull'entropia condizionale dà luogo a un numero maggiore di zone: 14, contro gli 8 temi individuati da M&B. Questo implica la presenza di celle che sono considerate da questi ultimi come costruite su uno stesso tema, ma risultano a volte non prevedibili senza alcun margine di incertezza nell'analisi basata sull'entropia. Ci sono infatti gruppi di celle che sono considerate come basate sul tema S1 da M&B, ma sono assegnate a una zona a sé nell'analisi proposta in questa sede, in particolare, le celle dell'imperfetto indicativo (Z9), la seconda persona plurale dell'indicativo e dell'imperativo (Z10), il participio presente e il gerundio (Z11). Quanto alle celle basate su S2, PRS.IND.1SG è assegnata a una zona a sé (Z12), mentre PRS.IND.3PL risulta nella stessa zona di PRS.IND.3SG (Z13), che viene invece considerata come basata sullo stesso tema di PRS.IND.2SG e IMP.2SG (S3) da M&B; queste ultime celle restano invece isolate nella nostra analisi, costituendo da sole rispettivamente le zone Z14 e Z3.

È utile indagare nel dettaglio i motivi delle differenze tra le diverse mappature. In primo luogo, la non completa sovrapponibilità delle due analisi è dovuta al fatto che, come si è accennato sopra alla Sezione 2, quella basata sui temi non rende conto di otto verbi molto frequenti e irregolari¹⁶, che esibiscono una distribuzione degli allomorfi tematici più complessa di quella prevista dallo spazio tematico proposto da M&B. Questi verbi sono invece inclusi nel nostro campione, generando incertezza nel prevedere le relative celle.

Ad esempio, la cella IPRF.SBJV.1PL può quasi sempre essere ricavata da IMP.2PL sostituendo la sequenza finale /te/ di quest'ultima con /ssimo/, ma questo non vale per alcuni verbi irregolari, tra cui DIRE, in cui le due forme sono basate su allomorfi tematici distinti, come si può osservare nella Tabella 14. Questa differenza genera un certo grado di incertezza nel prevedere le due celle l'una dall'altra, facendo sì che vengano assegnate a due zone distinte; ma dal momento che tali verbi irregolari non sono inclusi nell'analisi di M&B, le due celle vengono da questi considerate come basate sullo stesso tema.

verbo	IMP.2PL	IPRF.SBJV.1PL
AMARE	/am'ate/	/am'assimo/
SEDERE	/sed'ete/	/sed'essimo/
SENTIRE	/sent'ite/	/sent'issimo/
DIRE	/d'ite/	/ditf'essimo/

Tabella 16. Le celle IMP.2PL e IPRF.SBJV.1PL di alcuni verbi italiani

L'esclusione di tali verbi irregolari è motivata dall'esigenza di evitare di rendere la distribuzione notevolmente più complessa, per rendere conto del comportamento flessivo di un numero molto limitato di verbi, a maggior ragione in considerazione del fatto che è ragionevole ritenere che le forme di questi verbi siano memorizzate come tali nel lessico, vista la loro alta frequenza.

¹⁶ Si veda sopra alla nota 6 l'elenco completo.

Nell'approccio basato sull'entropia, invece, appare ragionevole includere tutti i verbi, dal momento che il carattere quantitativo della metodologia adottata (come abbiamo visto, si usa il numero di verbi in cui un pattern è attestato per valutarne la probabilità di applicazione) permette di tenere conto del grado minore di incertezza prodotto da fatti meno sistematici (in quanto attestati soltanto in pochissimi casi) come questi. Nell'esempio appena fornito, si può osservare nella Tabella 12 che i valori di entropia condizionale tra le celle IMP.2PL e IPRF.SBJV.1PL sono molto bassi – in entrambi i casi inferiori a 0,1. Di conseguenza, malgrado le due celle non possano essere assegnate a una stessa zona, le celle che appartengono alle zone in questione possono essere considerate più prevedibili le une dalle altre rispetto a quanto accade tra zone che includano forme basate su temi distinti nell'analisi di M&B. Nelle tabelle riportate in questo lavoro, si cerca di agevolare la visualizzazione di questo aspetto usando diverse sfumature di uno stesso colore per zone molto vicine in termini di prevedibilità, e colori del tutto diversi per zone più lontane.

Per rendere l'analisi proposta in questa sede più vicina a quella di M&B, una prima possibilità è dunque quella di usare una soglia diversa da zero per l'assegnazione di due celle a una stessa zona¹⁷. Se questa soglia è fissata a 0,01, come nella Tabella 17, le celle dell'imperfetto indicativo vengono assorbite nella zona Z1, coerentemente con il fatto che entrambe sono basate su S1 nell'analisi di M&B; se la soglia viene ulteriormente innalzata a 0,02, come nella Tabella 18, lo stesso accade anche per PRS.IND.2PL e IMP.2PL. Tuttavia, in entrambi i casi ci sono anche celle che confluiscono in una stessa zona malgrado siano considerate come basate su temi diversi da M&B: è il caso, ad esempio, del participio passato con soglia 0,01, e anche delle celle del preterito con soglia 0,02. Ciò è dovuto all'arbitrarietà della scelta della soglia, e alla difficoltà di individuarne una che corrisponda esattamente al grado di incertezza introdotto dai verbi irregolari esclusi.

17 È importante notare che, perché una cella venga assegnata ad una zona, è sufficiente che la prevedibilità sia inferiore alla soglia per una sola delle altre celle della zona, anche se rispetto alle altre la prevedibilità è invece superiore alla soglia. Ad esempio, PST.PTCP.F.SG è nella stessa zona di IPRF.SBJV.1PL nella Tabella 17, malgrado la prevedibilità reciproca di tali celle non risulti inferiore alla soglia di 0,01 (come si può vedere nella Tabella 12). Questo accade perché la prevedibilità di PST.PTCP.M.SG risulta inferiore alla soglia di 0,01 rispetto a IPRF.IND.3PL (celle non visibili nella Tabella 12 in quanto non tra quelle scelte come rappresentative; si vedano invece i risultati riportati nei materiali aggiuntivi online). Dal momento che, a loro volta, PST.PTCP.M.SG e PST.PTCP.F.SG appartengono alla stessa zona, così come IPRF.IND.3PL e IPRF.SBJV.1PL, anche PST.PTCP.F.SG e IPRF.SBJV.1PL vengono assegnate a tale zona. Ringrazio l'occhio attento di una delle persone che hanno revisionato questo lavoro per avermi reso consapevole della necessità di questo chiarimento.

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

	1SG	2SG	3SG	1PL	2PL	3PL		M.SG	F.SG	M.PL	F.PL
FUT.IND	Z6	Z6	Z6	Z6	Z6	Z6	PST.PTCP	Z1	Z1	Z1	Z1
PRS.COND	Z6	Z6	Z6	Z6	Z6	Z6	PRS.PTCP	Z9		Z9	
PRS.SBJV	Z2	Z2	Z2	Z4	Z4	Z2	GER	Z9			
PRS.IND	Z10	Z11	Z3	Z4	Z7	Z3	INF	Z8			
IPRF.IND	Z1	Z1	Z1	Z1	Z1	Z1					
IPRF.SBJV	Z1	Z1	Z1	Z1	Z1	Z1					
PRET.IND	Z5	Z1	Z5	Z1	Z1	Z5					
IMP		Z3		Z4	Z7						

Tabella 17. Zone paradigmatiche nella flessione verbale dell'italiano: $H(A|B) \wedge H(B|A) < 0,01$

	1SG	2SG	3SG	1PL	2PL	3PL		M.SG	F.SG	M.PL	F.PL
FUT.IND	Z6	Z6	Z6	Z6	Z6	Z6	PST.PTCP	Z1	Z1	Z1	Z1
PRS.COND	Z6	Z6	Z6	Z6	Z6	Z6	PRS.PTCP	Z7		Z7	
PRS.SBJV	Z2	Z2	Z2	Z4	Z4	Z2	GER	Z7			
PRS.IND	Z10	Z11	Z3	Z4	Z1	Z3	INF	Z8			
IPRF.IND	Z1	Z1	Z1	Z1	Z1	Z1					
IPRF.SBJV	Z1	Z1	Z1	Z1	Z1	Z1					
PRET.IND	Z1	Z1	Z1	Z1	Z1	Z1					
IMP		Z3		Z4	Z1						

Tabella 18. Zone paradigmatiche nella flessione verbale dell'italiano: $H(A|B) \wedge H(B|A) < 0,02$

Di conseguenza, a fini di confronto sembra più utile fare sì che le due analisi prendano in considerazione gli stessi verbi, escludendo dunque gli otto verbi irregolari in questione – insieme a tutti quelli che da essi sono derivati, e che dunque ne ereditano il comportamento flessivo, ad es. RIFARE da FARE – anche dal campione utilizzato per l’analisi basata sull’entropia. La mappatura che emerge da questo campione ridotto, costituito di 2.723 verbi, è riportata nella Tabella 19.

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

	1SG	2SG	3SG	1PL	2PL	3PL		M.SG	F.SG	M.PL	F.PL
FUT.IND	Z6	Z6	Z6	Z6	Z6	Z6	PST.PTCP	Z7	Z7	Z7	Z7
PRS.COND	Z6	Z6	Z6	Z6	Z6	Z6	PRS.PTCP	Z9		Z9	
PRS.SBJV	Z2	Z2	Z2	Z4	Z4	Z2	GER	Z9			
PRS.IND	Z10	Z12	Z11	Z4	Z1	Z11	INF	Z8			
IPRF.IND	Z1	Z1	Z1	Z1	Z1	Z1					
IPRF.SBJV	Z1	Z1	Z1	Z1	Z1	Z1					
PRET.IND	Z5	Z1	Z5	Z1	Z1	Z5					
IMP		Z3		Z4	Z1						

Tabella 19. Zone paradigmatiche nella flessione verbale dell'italiano: $H(A|B) = H(B|A) = 0$, esclusi verbi irregolari e derivati

Confrontando questi risultati con quelli ottenuti su tutto il campione (cfr. Tabella 11), come previsto, celle che erano assegnate a zone diverse in quest'ultimo caso confluiscono in una stessa zona con il campione ridotto, riducendo il numero di zone da 14 a 12. Ad esempio, le celle dell'imperfetto indicativo e la seconda persona plurale del presente indicativo e dell'imperativo vengono raggruppate nella zona Z1, coerentemente con quanto accade per queste celle nell'analisi basata sui temi. Rimangono tuttavia alcune differenze tra questa analisi e lo spazio tematico della Tabella 4: le zone Z9, Z10, Z11 e Z12 non trovano infatti corrispondenza in quest'ultima.

Un'altra ragione delle discrepanze tra le due analisi è legata al fatto che i temi di M&B sono formulati in maniera tale da consentire di prevedere tutte le forme all'interno della rispettiva zona, ma se si considerano le relazioni implicative tra le forme coinvolte, in alcuni casi c'è prevedibilità soltanto in una direzione. Si considerino ad esempio i dati della Tabella 20, in cui si riportano le forme di PRS.IND.2SG e PRS.IND.3PL insieme al tema S3 su cui tali forme si basano, come espresso da M&B. In virtù di tale formulazione, entrambe le forme possono essere ricavate a partire da S3, sulla base delle relazioni tema-forma in (4) e (5). Se però consideriamo le relazioni di prevedibilità tra forme, mentre la cella PRS.IND.2SG può in effetti essere ricavata senza alcuna incertezza da PRS.IND.3PL, lo stesso non vale per la previsione nella direzione opposta, come mostrato ai punti (6) e (7): data una seconda persona singolare in /i/, non è possibile identificare la coniugazione del verbo in questione, e di conseguenza ci sarà incertezza sulla terza plurale. Pertanto, le due celle vengono assegnate a due zone distinte.

lessema (coniugazione)	S3	PRS.IND.2SG	PRS.IND.3PL
AMARE (I)	/ 'ama/	/a' mi/	/ 'amano/
VEDERE (II)	/v' ede/	/v' edi/	/v' edono/
SENTIRE (III)	/s' ente/	/s' enti/	/s' entono/

Tabella 20. Le celle IMP.2PL e IPRF.SBJV.1PG e il tema S3 di alcuni verbi italiani

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

- (4) S3 /XV/ → PRS.IND.2SG /Xi/
- (5) S3 /Xa/ → PRS.IND.3PL /Xano/ ; S3 /Xe/ → PRS.IND.3PL /Xono/
- (6) PRS.IND.3PL /XVno/ → PRS.IND.2SG /Xi/
- (7) PRS.IND.2SG /Xi/ → PRS.IND.3PL /Xano/ ? ; /Xono/ ?

Per ovviare a questo problema, si può considerare sufficiente che tra due celle ci sia prevedibilità completa in una direzione perché tali celle siano assegnate a una stessa zona: in questo modo, PRS.IND.2SG e PRS.IND.3PL vengono assegnate alla stessa zona, coerentemente con quanto accade nell'analisi basata sui temi. La mappatura del paradigma verbale dell'italiano ottenuta in seguito a questa ulteriore modifica dei criteri di individuazione delle zone è mostrata nella Tabella 21.

	1SG	2SG	3SG	1PL	2PL	3PL		M.SG	F.SG	M.PL	F.PL
FUT.IND	Z5	Z5	Z5	Z5	Z5	Z5	PST.PTCP	Z1	Z1	Z1	Z1
PRS.COND	Z5	Z5	Z5	Z5	Z5	Z5	PRS.PTCP		Z1		Z1
PRS.SBJV	Z2	Z2	Z2	Z3	Z3	Z2	GER	Z1			
PRS.IND	Z2	Z2	Z2	Z3	Z1	Z2	INF	Z1			
IPRF.IND	Z1	Z1	Z1	Z1	Z1	Z1					
IPRF.SBJV	Z1	Z1	Z1	Z1	Z1	Z1					
PRET.IND	Z4	Z1	Z4	Z1	Z1	Z4					
IMP		Z2		Z3	Z1						

Tabella 21. Zone paradigmatiche nella flessione verbale dell'italiano: $H(A|B) \vee H(B|A) = 0$, esclusi verbi irregolari e derivati

Ora tutte le celle basate sullo stesso tema nell'analisi di M&B si trovano nella stessa zona. Ci sono però, viceversa, celle della stessa zona nell'analisi basata sull'entropia che sono invece considerate come basate su temi diversi da M&B. Il participio passato e l'infinito – basati rispettivamente su S7 e S8 – confluiscono in Z1, mentre PRS.IND.2SG, PRS.IND.3SG e IMP.2SG – basate su S3 – confluiscono in Z2, dando luogo a una struttura articolata in cinque zone, contro gli otto temi individuati da M&B.

Questa differenza è legata a casi in cui si riscontra la presenza di temi diversi dal punto di vista della loro costituzione segmentale, ma la metodologia adottata per l'estrazione dei pattern permette di individuare restrizioni fonologiche sull'applicabilità dei pattern di alternanza formale, rendendo le forme in questione prevedibili le une dalle altre senza alcuna incertezza.

Ad esempio, sulla base dei dati riportati nella Tabella 22, nell'analisi di M&B il tema su cui si basa PRS.IND.3SG (S3) è diverso da quello su cui si basa PRS.IND.3PL (S2). Ciononostante, le forme in questione risultano prevedibili l'una dall'altra nella misurazione basata sull'entropia, e questo non solo a partire dalla prima forma –dove ciò appare intuitivo, almeno sulla base dei

dati presentati – ma anche a partire dalla seconda, in virtù delle generalizzazioni ricavate dai dati del campione relativamente al contesto di applicazione dei diversi pattern di alternanza formale, come si può osservare nella quarta colonna. Sulla base di tali generalizzazioni, questi due verbi vengono assegnati a due classi distinte nella classificazione basata sulle alternanze applicabili a PRS.IND.3PL per ottenere PRS.IND.3SG, dato che a PUNGERE può essere applicato soltanto il pattern di alternanza formale con il contesto più ampio (2), mentre a PORRE può essere applicato anche quello ristretto a forme di terza singolare con le vocali /o/ oppure /a/ precedute dalle consonanti /b/, /m/ oppure /p/ (1).

verbo	PRS.IND. 3SG	PRS.IND. 3PL	alternanza/contesto (3PL → 3SG)	alternanze applicabili
PORRE	p'one	p'ongono	1 (_ngono ↔ _ne / [-lat]*[b-m-p][^h o-'a]_#)	A (1,2)
PUNGERE	p'undʒe	p'ungono	2 (_ngono ↔ _ndʒe / [-lat]*[+acc]_#)	B (2)

Tabella 22. Le celle PRS.IND.3SG e PRS.IND.3PL, con classificazione sulla base di alternanze applicate e applicabili, per alcuni verbi italiani

Dal punto di vista descrittivo, il fatto che generalizzazioni come queste siano valide sul campione considerato è sufficiente per prenderle in considerazione nel calcolo dell'entropia, e questo è il motivo alla base della differenza tra la nostra analisi e quella di M&B: PRS.IND.3SG e PRS.IND.3PL appartengono alla stessa zona nell'analisi basata sull'entropia, diversamente da quanto accade nello spazio tematico, in cui la cella PRS.IND.3SG è coindicizzata con PRS.IND.2SG e IMP.2SG, mentre PRS.IND.3PL con PRS.IND.1SG e alcune celle del congiuntivo presente.

È evidente però che tali restrizioni fonologiche sul contesto sono molto specifiche: questo solleva dubbi sul fatto che possano essere un artefatto delle caratteristiche dei dati del campione utilizzato, piuttosto che rispecchiare generalizzazioni effettivamente catturate dai parlanti; dubbio che può però essere risolto soltanto tramite un'accurata indagine psicolinguistica che richiederebbe un lavoro a parte, e non può essere svolta in questa sede: l'analisi presentata in questo lavoro – come del resto quella di M&B – ha soprattutto valenza descrittiva, e non ha pretese di replicare esattamente i meccanismi che i parlanti effettivamente operano nella produzione e nella comprensione delle forme.

4.2 L'entropia condizionale come metrica di valutazione delle diverse analisi

Nella sezione precedente, sono state discusse alcune differenze tra l'analisi della prevedibilità tra le celle del paradigma verbale dell'italiano proposta in questo lavoro e lo spazio tematico nella versione di M&B, individuando i fattori che le provocano e adottando alcuni aggiustamenti alla procedura adottata, in modo da rendere i due quadri più simili.

In primo luogo, l'analisi proposta in questo lavoro e quella di M&B differiscono, come si è visto, nell'approccio teorico: mentre in questo lavoro l'approccio è puramente astrattivo – il punto di partenza è costituito dalle forme flesse contenute nelle varie celle del paradigma – in quello di

M&B l'approccio è almeno parzialmente costruttivo –è infatti presupposta una segmentazione delle forme in temi e desinenze, e sono di conseguenza individuate relazioni formali tra temi da un lato, tra temi e forme dall'altro.

Ma i due lavori differiscono anche nella metodologia adottata ai fini dell'analisi: mentre in questo lavoro le zone paradigmatiche sono ottenute sulla base del calcolo dell'entropia effettuato a partire dalle alternanze estratte automaticamente tramite Qumin dal lessico LeFFI, in quello di M&B i temi e le relazioni sono identificati in seguito a un'accurata indagine manuale dei dati della flessione verbale dell'italiano. Come si è visto nell'ultimo esempio della Sezione 4.1, anche questo fattore genera significative differenze tra le due analisi.

Per appianare questa seconda differenza, la medesima metodologia basata sull'entropia adottata in questo lavoro può essere applicata anche alla prevedibilità tra temi come formulati da M&B, e delle forme a partire da tali temi. Usando come punto di partenza l'analisi di M&B, per prevedere una forma f_y basata su un tema S_y a partire da un'altra forma f_x basata su un tema S_x , è necessario innanzitutto (i) estrarre il tema S_x utilizzato nella forma di partenza f_x ; poi (ii) ottenere il tema S_y utilizzato nella forma di arrivo f_y a partire dal tema S_x ; infine (iii) ricavare la forma di arrivo f_y a partire dal tema S_y . I diversi passaggi sono schematizzati al punto (8) ed esemplificati al punto (9).

$$(8) \quad f_x \rightarrow S_x \rightarrow S_y \rightarrow f_y$$

$$(9) \quad \text{PRS.IND.1PL} /lavj'amo/ \rightarrow S4 /lav/ \rightarrow S1 /lav'a/ \rightarrow \text{PRS.IND.2PL} /lav'ate/$$

Di conseguenza, l'entropia che quantifica l'incertezza nel prevedere la forma f_y a partire dalla forma f_x nell'analisi di M&B può essere calcolata sommando i valori di entropia che quantificano l'incertezza di questi diversi passaggi, come mostrato al punto (10) e applicato alle celle usate come esempio al punto (11).

$$(10) \quad H(f_y | f_x) = H(S_x | f_x) + H(S_y | S_x) + H(f_y | S_y)$$

$$(11) \quad H(\text{PRS.IND.2PL} | \text{PRS.IND.1PL}) = 0,25 + 1 + 0 = 1,25$$

Si ottiene così un valore di entropia che quantifica l'incertezza nella previsione tra le celle in questione nell'analisi parzialmente costruttiva di M&B; valore che può essere confrontato a quello ottenuto nell'analisi puramente astrattiva svolta in questo lavoro tra le stesse celle, riferito alla previsione tra le forme flesse in questione, senza passaggi intermedi. In questo caso, il valore ottenuto con l'analisi astrattiva (1,19), è paragonabile a quello ottenuto con l'analisi parzialmente costruttiva, ma leggermente più basso. Questo indica che le generalizzazioni che si possono estrarre con gli strumenti utilizzati sulla relazione tra le forme flesse sono più utili di quelle che emergono in seguito alla scomposizione in temi e desinenze ai fini della prevedibilità tra queste due celle.

È dunque interessante svolgere tale confronto in maniera sistematica su tutte le possibili combinazioni di celle del paradigma. In questo modo, l'entropia condizionale può essere usata come una metrica di valutazione empirica delle diverse analisi: minore l'entropia, maggiore l'efficacia dell'analisi nel rendere conto degli stessi dati.

Dai temi alle forme: zone paradigmatiche nella flessione verbale dell'italiano

Per fare ciò, a partire dalle forme di LeFFI sono stati estratti i diversi temi come formulati da M&B. Dal momento che M&B non prendono esplicitamente in considerazione i pattern accen- tuali, per rendere le due analisi del tutto confrontabili è stata utilizzata una versione di LeFFI da cui è stata rimossa la notazione dell'accento, oltre a rimuovere i verbi irregolari esclusi dall'ana- lisi di M&B e quelli da essi derivati, come era stato fatto sopra nella Sezione 4.1.

I risultati del confronto sono sintetizzati nella Figura 4. Nella riga centrale si trova il numero di combinazioni di celle per cui l'entropia calcolata con l'approccio costruttivo che utilizza anche i temi $-H(S)$ è uguale a quella calcolata con l'approccio astrattivo che utilizza soltanto le forme flesse $-H(f)$. Nella parte in alto si trovano le combinazioni di celle per cui la prima è più alta della seconda - vale a dire, $H(S) > H(f)$. Nella parte in basso si trovano invece le combinazioni di celle in cui accade l'inverso - vale a dire, $H(f) > H(S)$. In entrambi i casi, si distinguono combinazioni di celle in cui la differenza è molto ampia (> 1), più ridotta (tra 0,1 e 1), oppure molto marginale ($< 0,1$).

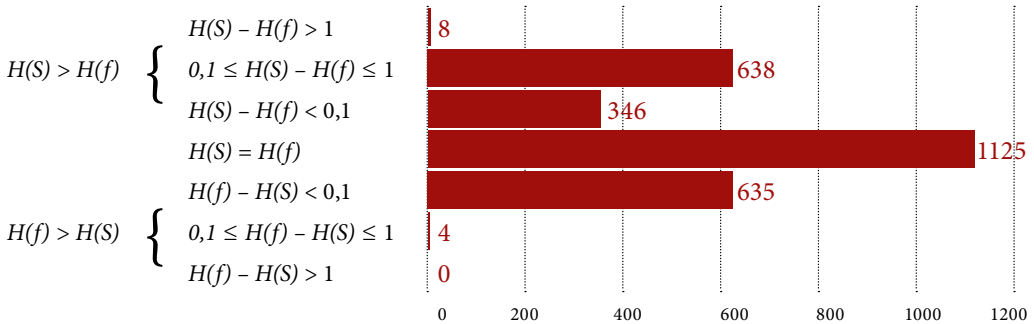


Figura 4. Confronto tra valori di entropia ottenuti a partire dai temi $-H(S)$ - e a partire dalle forme $-H(f)$

Da questi risultati, emerge che per molte combinazioni di celle (1.125 su un totale di 2.756), le due analisi sono perfettamente equivalenti, e laddove c'è una differenza questa è spesso trascurabile (minore di 0,1 per 981 delle 1.631 combinazioni in cui i risultati non sono identici).

Tuttavia, i casi in cui l'entropia calcolata solo sulla base delle forme flesse è maggiore di quella calcolata a partire da temi e forme (639) non solo sono meno numerosi di quelli in cui accade il contrario (992) in generale, ma sono anche in massima parte attestati relativamente a combi- nazioni di celle tra cui la differenza è minima, minore di 0,1 (635 casi), molto raramente quan- do si considera una differenza maggiore di tale soglia (4 casi). Al contrario, l'entropia calcolata a partire da temi e forme è spesso sensibilmente maggiore rispetto a quella calcolata solo sulla base delle forme: la differenza è maggiore di 0,1, in 638 casi, e addirittura maggiore di 1 in 8 casi. Questo indica che, in un'analisi basata sull'entropia, l'approccio puramente astrattivo adottato in questo lavoro dà luogo a un sistema più prevedibile di quello parzialmente costruttivo adottato da M&B, pur partendo da un numero minore di assunzioni di partenza, dal momento che non viene data per stabilita una segmentazione delle forme in temi e desinenze. È però utile indagare nel dettaglio i fattori che provocano le differenze più macroscopiche tra i valori di entropia otte- nuti con le due diverse modalità. Ad esempio, per quanto riguarda la previsione di PRS.IND.1PL

a partire da PRS.IND.1SG, l'entropia calcolata a partire da forme e temi è notevolmente più alta di quella calcolata solo sulla base delle forme flesse: la differenza è addirittura maggiore di 1, come si può ricavare dai risultati riportati ai punti (12) e (13).

$$(12) \quad H(\text{PRS.IND.1PL} \mid \text{PRS.IND.1SG}) = 0,52$$

$$(13) \quad H(S2 \mid \text{PRS.IND.1SG}) + H(S4 \mid S2) + H(\text{PRS.IND.1PL} \mid S4) = 0,78 + 0,24 + 0,6 = 1,62$$

Nella Tabella 23 sono mostrati i motivi di tale differenza. La cella PRS.IND.1SG non permette di distinguere tra verbi di diverse coniugazioni, dato che esce in /o/ in tutti i casi. Dal momento che il tema S2 su cui tale cella è basata nella formulazione di M&B, in cui corrisponde alla forma di PRS.SBJV.1/2/3SG, distingue invece tra verbi della I coniugazione (con S2 in /i/) e verbi delle altre coniugazioni (con S2 in /a/), c'è una notevole incertezza ($H = 0,78$) nel prevedere il tema a partire dalla forma. Tuttavia, questa fonte di incertezza non è invece presente se a partire dalla stessa cella si vuole prevedere direttamente la forma di PRS.IND.1PL, dal momento che anche in quest'ultima cella non si fa distinzione tra verbi di diverse coniugazioni – la terminazione è in ogni caso in /jamo/.

In maniera simile, l'incertezza nel prevedere S4 a partire da S2 ($H = 0,24$) è provocata dal fatto che un tema S2 in /i/ può essere relativo a verbi della I coniugazione con S4 in consonante (ad es. AMARE, S4 /am/) oppure con S4 in /j/ (ad es. SPECCHIARE, S4 /spekkj/); ma anche questa fonte di incertezza scompare se si prevede direttamente PRS.IND.1PL da PRS.IND.1SG, dal momento che in PRS.IND.1PL /j/ è presente in entrambi i casi.

L'unica fonte di incertezza nella previsione di una forma a partire dall'altra è la presenza di verbi in cui l'uscita di PRS.IND.1PL non presenta il segmento /j/ (ad es. VOLTEGGIARE); ma quest'incertezza è comunque presente anche nel calcolo dell'entropia relativa ai rispettivi temi, dal momento che a un S2 in /i/ può corrispondere un S4 in consonante (come /voldedd3/) oppure in /j/ (come /spekkj/).

	PRS.IND.1SG	S2	S4	PRS.IND.1PL
AMARE (I)	/amo/	/ami/	/am/	/amjamo/
VEDERE (II)	/vedo/	/veda/	/ved/	/vedjamo/
SENTIRE (III)	/sento/	/senta/	/sent/	/sentjamo/
VOLTEGGIARE (I)	/voldedd3o/	/voldedd3i/	/voldedd3/	/voldedd3amo/
SPECCHIARE (I)	/spekkjo/	/spekki/	/spekkj/	/spekkjamo/

Tabella 23. Le forme PRS.IND.1SG e PRS.IND.1PL e i temi S2 e S4 di alcuni verbi italiani

Passiamo ora a un caso in cui l'entropia calcolata solo sulla base delle forme flesse è sensibilmente più alta di quella calcolata a partire da forme e temi. Questo è quello che succede per la previsione di IMP.2SG a partire da PST.PTCP.M.SG, come si può vedere ai punti (14) e (15).

$$(14) \quad H(IMP.2SG | PST.PTCP.M.SG) = 0,4$$

$$(15) \quad H(S7 | PST.PTCP.M.SG) + H(S3 | S7) + H(IMP.2SG | S3) = 0 + 0,08 + 0 = 0,08$$

Dal momento che al punto (14) l'unico valore di entropia non nullo è quello relativo alla previsione di S3 a partire da S7, possiamo confrontare direttamente questa previsione con quella delle rispettive forme, tralasciando la previsione di S7 a partire dal participio e quella dell'imperativo a partire da S3.

Se si calcola l'entropia di prevedere IMP.2SG a partire da PST.PTCP.M.SG (cfr. Tabella 24), viene estratto uno stesso pattern di alternanza formale (1) per i verbi della I coniugazione e per quelli della III della sottoclasse che non presenta /ʃʃ/ all'imperativo. Il contesto di applicazione del pattern dovrà quindi essere abbastanza ampio da coprire entrambi i casi – in questo esempio, leggermente semplificato rispetto alla procedura come applicata automaticamente da Qumin, è considerato come genericamente applicabile dopo vocale. Analogamente ampio risulta il contesto di applicazione del pattern di alternanza formale (3) proprio dei verbi della III coniugazione della sottoclasse che presenta /ʃʃ/ all'imperativo, dato che lo stesso pattern è applicato anche dopo la vocale /a/ in una coppia di forme come *nato-nasci*. Dato che i contesti di applicazione di questi due pattern sono gli stessi, verbi che presentano tali pattern sono assegnati a una stessa classe nella colonna relativa alle alternanze potenzialmente applicabili: ciò implica che ci sarà incertezza per tutti i verbi che appartengono a tali classi, che sono molti, come si può osservare nell'ultima colonna.

Se invece si calcola l'entropia di prevedere S3 a partire da S7 (cfr. Tabella 25), il pattern di un verbo della I come AMARE (1) risulta diverso da quello di un verbo della III come SENTIRE, dal momento che S3, per come è formulato da M&B, distingue tra queste due coniugazioni (essendo identico a PRS.IND.3SG, esce in /a/ per verbi della I, in /e/ per verbi delle altre coniugazioni). Di conseguenza, il contesto del pattern (1) può essere formulato in modo che risulti applicabile soltanto dopo la vocale /a/. Pertanto, per tutti i verbi che presentano S7 in *-at* l'incertezza sulla forma di S3 sarà ridotta, dato che il pattern di alternanza con inserzione di /ʃʃ/ è applicato ad un solo verbo della classe; l'incertezza sarà invece più marcata per i verbi che presentano S7 in *-it*, che però sono molto meno numerosi. Dal momento che l'incertezza riguarda una porzione molto più limitata del lessico¹⁸, il valore di entropia calcolato risulta molto più basso.

18 Le differenze di dettaglio nelle frequenze delle diverse classi sono sempre dovute al fatto che S3 è ricavata a partire da una cella diversa (PRS.IND.3SG) rispetto a quella in questione (IMP.2SG); di conseguenza alcuni verbi presentano un pattern irregolare per la prima e non per la seconda, e viceversa.

verbo (coniugazione)	PST.PTCP. M.SG	IMP.2SG	alternanza/contesto	alternanze applicabili	n. verbi
AMARE (I)	amato	ama	1 (_to ↔ _ / V_#)	A (1,3)	2.080
VENDERE (II)	venduto	vendi	2 (_uto ↔ _i / C_#)	B (2)	57
USUFRUIRE (III)	uzufruito	uzufruiffi	3 (_to ↔ _ffi / V_#)	A (1,3)	218
SENTIRE (III)	sentito	senti	1 (_to ↔ _ / V_#)	A (1,3)	48
NASCERE (II)	nato	naffi	3 (_to ↔ _ffi / V_#)	A (1,3)	1

Tabella 24. Le celle PST.PTCP.M.SG e IMP.2SG, con classificazione sulla base di alternanze applicate e applicabili, per alcuni verbi italiani

verbo (coniugazione)	S ₇	S ₃	alternanza/contesto	alternanze applicabili	n. verbi
AMARE (I)	amat	ama	1 (_t ↔ _ / a_#)	A (1,3)	2.080
VENDERE (II)	vendut	vende	2 (_ut ↔ _e / C_#)	B (2)	59
USUFRUIRE (III)	uzufruit	uzufruiffè	3 (_t ↔ _ffe / V_#)	C (3,4)	218
SENTIRE (III)	sentit	sente	4 (_it ↔ e_ / C_#)	C (3,4)	55
NASCERE (II)	nat	naffe	3 (_t ↔ _ffe / V_#)	A (1,3)	1

Tabella 25. I temi S₇ e S₃, con classificazione sulla base di alternanze applicate e applicabili, per alcuni verbi italiani

Tirando le somme di quanto si può concludere sulla base dei due esempi appena discussi, le differenze nei risultati ottenuti con le due modalità risultano spesso dovute al fatto che i temi di M&B sono concretamente ricavati a partire da determinate forme (ad es. S₂ da PRS.SBJV.1/2/3SG nel primo esempio, S₃ da PRS.IND.3SG nel secondo). Questo fa sì che emergano differenze significative nell'applicazione della procedura per il calcolo dell'entropia nel momento in cui tali temi siano coinvolti in previsioni che riguardano celle diverse da quelle da cui sono ricavati (ad es. PRS.IND.1SG nel primo esempio, IMP.2SG nel secondo).

In un caso come quello del primo esempio, vi sono fattori che introducono incertezza nel sistema che compaiono soltanto nelle previsioni che coinvolgono i temi, e possono invece essere evitati se si opera esclusivamente sulle forme. Dal momento che il dato a cui i parlanti sono concretamente esposti, e quello che devono dunque prevedere al momento della produzione, sono appunto le forme, piuttosto che temi e desinenze in isolamento, questi elementi di incertezza vanno considerati come introdotti dall'analisi, che dunque ha l'effetto di rendere più difficili le previsioni, invece che più semplici, almeno relativamente a queste celle.

In un caso come quello del secondo esempio, invece, l'analisi genera una semplificazione del compito, ma tale semplificazione è in ultima istanza dovuta al fatto che viene presa in considerazione un'altra forma – quella a partire da cui il tema è concretamente ricavato – rispetto a quella

in questione. Di conseguenza, viene introdotta artificiosamente un'informazione aggiuntiva, che è ovviamente utile, ma rende le due previsioni non del tutto confrontabili¹⁹.

5. Conclusioni

In questo lavoro, dopo aver passato in rassegna le analisi della prevedibilità nella flessione verbale dell'italiano basate sulla distribuzione dei temi nelle celle del paradigma svolte in lavori precedenti (Sezione 2), ne è stata proposta una ottenuta tramite un approccio puramente astrattivo, basato esclusivamente sulle forme flesse, senza assumere una segmentazione in temi e desinenze. La metodologia adottata si fonda sulla nozione di entropia condizionale, propria della teoria dell'informazione (Sezione 3.1). Tale metodologia è stata applicata sistematicamente ai dati di un lessico di forme flesse dell'italiano –LeFFI– utilizzando come strumento un insieme di script in python per la modellizzazione quantitativa di fatti di morfologia flessiva, Qumin (Sezione 3.2). I risultati così ottenuti (Sezione 3.3) sono stati poi confrontati sistematicamente con quelli dei lavori precedenti basati sui temi.

Da tale confronto, emerge (Sezione 4.1) che il quadro a cui le diverse analisi danno luogo è nel complesso simile: molte delle zone paradigmatiche individuate sulla base dell'entropia corrispondono a celle che sono considerate come basate sullo stesso tema nei lavori precedenti. Sono però emerse anche significative differenze, dovute, oltre che al diverso approccio (parzialmente costruttivo nelle analisi basate sui temi, puramente astrattivo in questo lavoro), anche a scelte metodologiche diverse, ad esempio riguardo all'esclusione di verbi altamente frequenti e irregolari, oppure alla scelta se prendere in considerazione solo casi in cui ci sia prevedibilità reciproca tra le celle in questione o ritenere sufficiente che si possa prevedere una cella a partire dall'altra, e non viceversa, oppure infine relativamente alle modalità di ottenimento delle zone paradigmatiche (tramite un'accurata ispezione manuale dei dati nelle analisi basate sui temi, tramite un'estrazione automatica sulla base dei valori di entropia in questo lavoro).

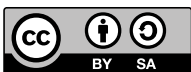
Per appianare le differenze metodologiche, la procedura seguita in questo lavoro è stata applicata anche per quantificare l'incertezza nella previsione dei temi l'uno dall'altro, e delle forme a partire dai temi e viceversa, usando come punto di partenza l'analisi svolta da Montermini e Bonami (2013). Si ottengono così (Sezione 4.2) dei risultati che sono facilmente confrontabili con quelli relativi alla previsione delle forme le une dalle altre, come calcolati nell'analisi proposta in questa sede. In questo modo, l'entropia condizionale può essere utilizzata come metrica di valutazione dell'economia descrittiva delle diverse analisi, mostrando che anche così facendo i risultati restano in generale abbastanza simili, ma laddove c'è una differenza l'approccio puramente astrattivo adottato in questa sede è in grado di dare luogo a un sistema che risulta più prevedibile, pur partendo da un numero minore di assunzioni di partenza.

¹⁹ Un risultato del genere andrebbe forse confrontato piuttosto con quella che Bonami e Beniamine (2016) chiamano “joint predictiveness”, prendendo in considerazione più di una forma nel calcolo dei valori di entropia, e valutando così l'incertezza nel prevedere il contenuto di una cella dando per noto il contenuto di più di una cella, invece che di una sola, come si è fatto in questo lavoro. Bonami e Beniamine (2016) mostrano che l'entropia diminuisce all'aumentare del numero di celle considerate: di conseguenza, è plausibile che un'analisi di questo tipo avvicini i valori di entropia a quelli calcolati sulla base dei temi in casi come questi; per ragioni di spazio, questa analisi non può però essere svolta in questa sede.

Riferimenti bibliografici

- Ackerman, F.; Blevins, J. P.; & Malouf, R. (2009). Parts and wholes: implicative patterns in inflectional paradigms. In J. P. Blevins, & J. Blevins (Eds.), *Analogy in grammar: Form and acquisition* (pp. 54–82). Oxford: Oxford University Press.
- Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes*. Cambridge, Mass.: MIT Press.
- Beniamine, S. (2018). *Classifications flexionnelles. Étude quantitative des structures de paradigmes*. PhD Thesis. Université Sorbonne Paris Cité-Université Paris Diderot (Paris 7).
- Beniamine, S.; Bonami, O.; & Sagot, B. (2017). Inferring inflection classes with description length. *Journal of Language Modelling*, 5(3), 465–525.
- Bertinetto, P. M.; Burani, C.; Laudanna, A.; Marconi, L.; Ratti, D.; Rolando, C.; & Thornton, A. M. (2005). CoLFIS (Corpus e Lessico di Frequenza dell'Italiano Scritto).
- Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics*, 42(3), 531–573.
- Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford: Oxford University Press.
- Bonami, O.; & Beniamine, S. (2021). Leaving the stem by itself. In M. Haag, S. Moradi, A. Petrovic, & J. Rees-Miller (Eds.), *All Things Morphology: Its independence and its interfaces* (pp. 81–98). Amsterdam: John Benjamins.
- Bonami, O.; & Boyé, G. (2003). Supplétion et classes flexionnelles. *Langages* 37(152), 102–126.
- . (2014). De formes en thèmes. In F. Villoing, S. Leroy, & S. David (Eds.), *Foisonnements morphologiques. Études en hommage à Françoise Kerleroux* (pp. 17–45). Paris: Presses Universitaires de Paris-Ouest.
- Boyé, G.; & Cabredo Hofherr, P. (2006). The structure of allomorphy in Spanish verbal inflection. *Cuadernos de Lingüística*, 13, 9–24.
- Bybee, J. (1995). Regular morphology and the lexicon. *Language and cognitive processes* 10(5), 425–455.
- Calderone, B.; Pascoli, M.; Hathout, N.; & Sajous, F. (2017). Hybrid method for stress prediction applied to GLAFF-IT, a large-scale Italian lexicon. In J. Gracia, F. Bond, J. P. McCrae, P. Buitelaar, C. Chiarcos, & S. Hellman (Eds.), *Language, Data and Knowledge. First International Conference, LDK 2017, Galway, Ireland, June 19–20, 2017, Proceedings* (pp. 26–41). Cham: Springer.
- Carstairs, A. (1987). *Allomorphy in inflexion*. London: Croom Helm.
- Carstairs-McCarthy, A. (1994). Inflection Classes, Gender, and the Principle of Contrast. *Language*, 70(4), 737–788.
- Corbett, G. G. (2009). Canonical inflectional classes. In F. Montermini, G. Boyé & J. Tseng (Eds.), *Selected proceedings of the 6th Décembrettes: Morphology in Bordeaux* (pp. 1–11). Somerville: Cascadilla Proceedings Project.
- Dressler, W. U. (2002). Latin inflection classes. In A. M. Bolkestein, C. H. M. Kroon, H. Pinkster, H. W. Rummelink, & R. Risselada (Eds.), *Theory and Description in Latin Linguistics: Selected Papers from the XIth International Colloquium on Latin Linguistics* (pp. 91–110). Amsterdam: Gieben.
- Guerrero, A. (2014). *Analyse thématique de la flexion en catalan central standard*. PhD Thesis. Université Toulouse le Mirail.
- Guzmán Naranjo, M. (2019). *Analogical classification in formal grammar*. Berlin: Language Science Press.
- Loporcaro, M. (2012). Stems, endings and inflectional classes in Logudorese verb morphology. *Lingue e linguaggio*, 11(1), 5–34.

- Maiden, M. (1992). Irregularity as a determinant of morphological change. *Journal of Linguistics*, 28, 285–312.
- . (2018). *The Romance verb. Morphomic structure and diachrony*. Oxford: Oxford University Press.
- Montermini, F.; & Bonami, O. (2013). Stem spaces and predictability in verbal inflection. *Lingue e linguaggio*, 12(2), 171–190.
- Montermini, F.; & Boyé, G. (2012). Stem relations and inflection class assignment in Italian. *Word Structure*, 5(1), 69–87.
- Pellegrini, M. (2023). *Paradigm Structure and Predictability in Latin Inflection: An Entropy-based Approach*. Cham: Springer.
- Pellegrini, M.; & Cignarella, A. T. (2020). (Stem and Word) Predictability in Italian verb paradigms: An Entropy-Based Study Exploiting the New Resource LeFFI. In F. Dell'Orletta, J. Monti, & F. Tamburini (Eds.), *Proceedings of the Seventh Italian Conference on Computational Linguistics (CLiC-it 2020)*. Torino: Accademia University Press.
- Pirrelli, V. (2000). *Paradigmi in morfologia: un approccio interdisciplinare alla flessione verbale dell'italiano*. Pisa & rRoma: Istituti editoriali e poligrafici internazionali.
- Pirrelli, V. & Battista, M. (2000). The paradigmatic dimension of stem allomorphy in Italian verb inflection. *Italian Journal of Linguistics*, 12(2), 307–380.
- Stump, G. T. & Finkel, R. A. (2013). *Morphological typology: From word to paradigm*. Cambridge: Cambridge University Press.
- Wurzel, W. U. (1984). *Flexionsmorphologie und Natürlichkeit: ein Beitrag zur morphologischen Theoriebildung*. Berlin: Akademie-Verlag.



This work can be used in accordance with the Creative Commons BY-SA 4.0 International license terms and conditions (<https://creativecommons.org/licenses/by-sa/4.0/legalcode>). This does not apply to works or elements (such as images or photographs) that are used in the work under a contractual license or exception or limitation to relevant rights.