

## A non-canonical phenomenon in Italian verb morphology: double forms realizing the same cell

Anna M. Thornton  
Università dell'Aquila

Poster presented at OxMorph1 – Oxford, August 2008

### Introduction

In the canonical approach to morphological description (Corbett 2005, 2007a, 2007b), a canonical paradigm can be expected to exhibit **UNIQUENESS OF REALIZATION**, i.e., “for any given lexeme, every cell of its paradigm will be filled in a unique way”. Deviations from this canonical situation are represented by doublets, i.e., cases in which a cell is filled by two synonymous forms which realize the same set of morphosyntactic properties.

Such deviations are claimed to be abundant in Italian, due to the peculiar history of the language used almost only in written form for centuries. De Mauro (1976: 29) states that

Allorché una stessa forma flessionale si presentava in due varianti, queste venivano entrambe conservate per secoli nella tradizione italiana, nella quale non potevano intervenire le tendenze selettive, economiche, operanti in lingue sottoposte a un uso più largo. E se talora il pur esiguo uso scritto ha differenziato almeno stilisticamente gli elementi polimorfi, facendo d’una forma quella canonica nella versificazione e d’un’altra quella propria della prosa (è il caso, per fare qualche esempio, di *esto* e *questo*, *gito* e *andato*, *alma* e *anima*, *augello* e *uccello*), più spesso tale differenziazione, ancora a metà Ottocento, mancava: così in *fo* e *faccio*, *vo* e *vado*, *devo* e *debbo* («poetico» era *deggio*), *dovuto* e *debito*, *dette diede* e *dié*, *visto* e *veduto*, *apparisco* e *appaio* ecc.

De Mauro’s examples constitute an almost complete catalogue of the areas in which doublets are attested in Italian verbal inflection.

A possible tentative classification of the kinds of doublets attested in Italian is offered in Table I, organized according to Corbett’s criteria for canonicity. The very presence of doublets is non-canonical, as it violates **UNIQUENESS OF REALIZATION**. Table I helps to situate different kinds of non-uniqueness.

Table I. Deviations from uniqueness of realization of cells in Italian verb morphology

Dimension in which the deviation takes place	Deviations instantiated by double forms in the same cell of the same lexeme	Examples
COMPOSITION/STRUCTURE (of the inflected word)	forms following two different conjugations in the same cell	<i>eseguo</i> / <i>eseguisco</i> ‘perform:1.PRS.IND’
LEXICAL MATERIAL (≈ shape of stem)	forms built on two different stems in the same cell	<i>devo</i> / <i>debbo</i> ‘must:1.PRS.IND’ <i>vado</i> / <i>vo</i> , <i>faccio</i> / <i>fo</i> ‘go:1.PRS.IND’, ‘do:1.PRS.IND’ <i>doliamo</i> / <i>dogliamo</i> ‘suffer:1PL.PRS.IND’ <i>giacciamo</i> / <i>giaciamo</i> ‘lie_down:1PL.PRS.IND’

		<i>parvi /parsi</i> ‘seem:1.PRF.IND’ <i>sederò / siederò</i> ‘sit:1.FUT’ <i>morrò / morirò</i> ‘die:1.FUT’ <i>udirò / udrò</i> ‘hear:1.FUT’ <i>visto / veduto</i> ‘see:PST.PTCP’
INFLECTIONAL MATERIAL (≈ shape of inflection)	forms with two different inflectional endings in the same cell	<i>temei / temetti</i> ‘fear:1.PRF.IND’ <i>va’ / vai, fa’ / fai, sta’ / stai</i> ‘go:2.IMP’, ‘do:2.IMP’, ‘stay:2.IMP’ <i>udente / udiente</i> ‘hear:PRS.PTCP’

The poster assesses the incidence of these different kinds of doublets in contemporary written Italian, based on investigation of the corpus containing 16 years (1985-2000) of issues of the daily newspaper *la Repubblica*, available online at the following address:

<http://sslmitdev-online.sslmit.unibo.it/corpora/corpus.php?path=&name=Repubblica>.

Some questions that the data might help to answer are the following:

- what doublets are actually attested in contemporary written Italian?
- is there any kind of functional differentiation between the two members of a pair of doublets?
- what properties characterize those pairs of doublets which appear not to be functionally differentiated?
- in cases where a tendency to the reduction / elimination of doublets can be observed, what is the direction of the change? i.e., what member of a doublet pair is eliminated?

## Data

### 1. Forms following two different conjugations in the same cell

#### 1.1. *-isc-* vs. *-isc-*less forms

An investigation was conducted of 18 verbs for which grammars signal the possibility of doublets with and without *-isc-* in the N-pattern cells. The set includes the four verbs for which *-isc-* forms were elicited in the off-line tests administered by Spina (2007) (*applaudire, bollire, cucire, mentire*) and all the verbs for which *-isc-* forms were elicited in unpublished off-line tests administered by Dressler and Thornton in 1995 (the four above, plus *eseguire, convertire, sdruccire, languire, inghiottire*).

The data are presented in Table II.

Table II *-isc-* vs. *-isc-*less forms in the *Repubblica* corpus  
(in order of decreasing amount of *-isc-* forms)

verb		1SG.PRS.IND	2SG.PRS.IND / 2SG.PRS.IMP	3SG.PRS.IND	3PL.PRS.IND	1/2/3SG.PRS.SBJV	3PL.PRS.SBJV	Total
<b>lambire</b>	with <i>-isc-</i>	0	0	206	79	7	1	293
<b>lambire</b>	without <i>-isc-</i>	0	0	0	0	0	0	0
<b>aborrire</b>	with <i>-isc-</i>	6	0	40	14	0	0	60
<b>aborrire</b>	without <i>-isc-</i>	21	0	63	25	1	2	112

<b>languire</b>	with <i>-isc-</i>	0	0	11	17	1	1	30
<b>languire</b>	without <i>-isc-</i>	<b>0</b>	<b>1</b>	<b>366</b>	<b>155</b>	<b>0</b>	<b>1</b>	<b>&gt;500</b>
<b>inghiottire</b>	with <i>-isc-</i>	0	1	8	5	1	0	15
<b>inghiottire</b>	without <i>-isc-</i>	6	5	287	64	12	2	376
<b>mentire</b>	with <i>-isc-</i>	0	0	5	4	2	0	11
<b>mentire</b>	without <i>-isc-</i>	ambiguous form* 1002 2/20	ambiguous form* 1299 0/20	ambiguous form* >10.000 1/20	206	ambiguous form* 230 1/20	12	>200 (> 800?)
<b>apparire</b>	with <i>-isc-</i>	0	0	2 in quotations (Carducci, D'Annunzio)	3 in quotations (Verdi, Bocchi)	1 in a quotation (Daniello Bartoli)	1 in a quotation (letter by Saba, 1923)	6 in quotations
<b>apparire</b>	without <i>-isc-</i>	<i>appaio</i> 46	<i>appari</i> 19	<i>appare</i> >10.000	<i>appaiono</i> 8289	<i>appaia</i> 918	<i>appaiano</i> 193	>19.000
<b>assorbire</b>	with <i>-isc-</i>	0	0	0	2	0	0	2
<b>assorbire</b>	without <i>-isc-</i>	9	2	1102	449	54	23	1639
<b>applaudire</b>	with <i>-isc-</i>	0	0	1	0	0	0	1
<b>applaudire</b>	without <i>-isc-</i>	30	8	1942	1045	19	10	>3000
<b>comparire</b>	with <i>-isc-</i>	0	1 dialectal	0	0	0	0	1 dialectal
<b>comparire</b>	without <i>-isc-</i>	<i>compaio</i> 21	<i>compari</i> 355 ambiguous form* 0/20	<i>compare</i> 4679 ambiguous form* 17/20	<i>compaiono</i> 2478	<i>compaia</i> 197	<i>compaiano</i> 84	>7500 with ambiguous forms
<b>assalire</b>	with <i>-isc-</i>	0	0	0	0	0	0	0
<b>assalire</b>	without <i>-isc-</i>	<i>assalgo</i> 0	<i>assali</i> 52	<i>assale</i> 355	<i>assalgono</i> 114	<i>assalga</i> 7	<i>assalgano</i> 1	529
<b>avvertire</b>	with <i>-isc-</i>	0	0	0	0	0	0	0
<b>avvertire</b>	without <i>-isc-</i>	451	66	>10.000	2891	206	70	>13.500
<b>bollire</b>	with <i>-isc-</i>	0	0	0	0	0	0	0
<b>bollire</b>	without <i>-isc-</i>	ambiguous form	ambiguous form	ambiguous form	72	ambiguous form	ambiguous form	≥72
<b>convertire</b>	with <i>-isc-</i>	0	0	0	0	0	0	0
<b>convertire</b>	without	2	2	240	93	53	24	>400

	-isc-							
<b>cucire</b>	with -isc-	0	0	0	0	0	0	0
<b>cucire</b>	without -isc-	6	32	210	67	3	1	>300
<b>divertire</b>	with -isc-	0	0	0	0	0	0	0
<b>divertire</b>	without -isc-	653	41	2832	835	192	72	>4500
<b>eseguire</b>	with -isc-	0	0	0	0	0	0	0
<b>eseguire</b>	without -isc-	43	7	1004	415	42	20	>1500
<b>nutrire</b>	with -isc-	0	0	0	0	0	0	0
<b>nutrire</b>	without -isc-	272	11	2108	956	98	47	>3400
<b>sdrucire</b>	with -isc-	0	0	0	0	0	0	0
<b>sdrucire</b>	without -isc-	0	0	0	0	0	0	0

\* It was not possible for practical reasons to examine all the tokens of this ambiguous form. The figure following the total refers to how many tokens are instances of the verb form being investigated among the first 20 random tokens of the word form in the corpus.

The corpus investigation shows that only eight of the 18 verbs display doublets with and without *-isc-*. *Lambire* has *-isc-* categorically, and the remaining nine verbs (including some which elicited *-isc-* responses in off-line tests, such as *bollire*, *convertire*, *cucire*, *eseguire*) have no *-isc-* forms in the corpus. *Sdrucire* has no forms belonging to the N-pattern cells in the corpus.

The *-isc-* forms are always less numerous than the *-isc-*-less forms, with ratios ranging from 1:1,9 (*aborrire*) to 1: >3000 (*apparire*, *applaudire*).

Most interesting are the data concerning the *-isc-* forms of *apparire*. All these forms appear in the corpus in quotations from non-contemporary authors, from Francesco Bocchi (1548-1618) and Daniello Bartoli (1608-1685) to Umberto Saba (1883-1957; the *-isc-* form appears in a letter written in 1923), through Verdi, Carducci and D'Annunzio. It can be concluded that *-isc-* forms of *apparire* are **literary** or **archaic** in contemporary written Italian.

The only *-isc-* form of *comparire* appears in a quotation of a Neapolitan saying, and can be classified as **dialectal**.

il massimo effetto col minimo sforzo, "sparagna e **comparisci**" gli avranno insegnato a Napoli.

In conclusion, very few verbs in the corpus have *-isc-* forms that are not diachronically, diaphasically or diatopically marked: *aborrire*, *languire*, *inghiottire*, *mentire*, *assorbire* and *applaudire*.

## 2. Forms built on two different stems in the same cell

This is the category richest in doublets. Several cases are examined in the following.

### 2.1. *vado* / *vo*

vado 6725 tokens    vo 140 tokens

All 140 tokens of *vo* in the *Repubblica* corpus were examined. 36 turned out to be false hits, constituted of misspellings such as <vo gliano>, <vo lontà> and the like; two tokens were in a language different from Italian. The remaining 102 tokens represent the forms shown below:

Forms realized by <vo> in the *Repubblica* corpus

a.	<i>vo</i>	‘andare:1SG.PRS.IND’	38
b.	<i>vo</i>	‘volere:3SG.PRS.IND’	46
c.	<i>vo</i>	‘volere:1SG.PRS.IND’	11
d.	<i>vo</i>	‘volere:2SG.PRS.IND’	7

*Vo* representing ‘andare:1SG.PRS.IND’ is not even the most frequent kind of *vo* in the corpus. All its 38 tokens occur in contexts which are non-standard or peculiar in some respect, as illustrated in Table III.

Table III Contexts of occurrence of *vo* ‘andare:1SG.PRS.IND’

13 tokens	<b>Quotations from songs and opera pieces</b> “Sola me ne vo per la città” ( <i>In cerca di te</i> , 1945) 7 tokens “Quando men vo soletta” ( <i>La Bohème</i> ) 4 tokens
12 tokens	<b>Tuscan speakers.</b> In all cases, the speakers are clearly identifiable in the text, and their being Tuscan is stated in the text or clearly understandable from other features of their speech besides the usage of <i>vo</i> : cf. the examples below. Speakers include well-known Tuscan celebrities such as Gino Bartali, Leonardo Pieraccioni and Pietro Pacciani.
8 tokens	<b>Quotations from Dante or analogies on Dante’s passages</b>
5 tokens	other cases

Some examples of *vo* in reported speech by Tuscan speakers:

Oggi volevo fare un appello al signor ministro e mi son detto: **vo** in televisione e **fo** l'appello, ma poi ho pensato che **il mio figliolo** mi avrebbe visto così conciato in televisione

'Che hai fatto?', le chiesi. **La mi rispose** male: 'Io so ' maggiorenne, **vo in do'** mi pare, **grullo** '. Avevo la **ramazza** in mano e gliene detti 7 o 8 di ramazzate.

"Che **vo** insieme a quella gente lì? **Un lo vede**, ha 50 anni e **un ha** una moglie, **gli è** un finocchio, quello non è una persona normale. [...]"

Original passage from Dante:

Or ti piaccia gradir la sua venuta:  
libertà **va** cercando, ch'è sì cara,  
come sa chi per lei vita rifiuta. (Purg. I,70-72)

Quotations from the *Repubblica* corpus:

- |    |   |
|----|---|
| a. | Perestrojka <b>vo</b> cercando ... mormora alla moglie un vecchio dentista russo davanti a un chiosco di giornali.  |
| b. | Qualità <b>vo</b> cercando, dice Romiti al pubblico, alla stampa e tra le mura non sempre discrete di casa propria.                                       |
| c. | Libertà <b>vo'</b> cercando, ma è sì cara ...   |
| d. | BOLOGNA -- L' articolo di fondo s'intitola "Il barbone del 2000" e cita (per contestarlo) Dante Alighieri: "Libertà <b>vo</b> cercando ch'è sì cara ...". |
| e. | "Civiltà <b>vo</b> cercando? ...".  |
| f. | "Continuità e stabilità, <b>vo</b> cercando", ha detto in sostanza lo stesso Putin, presiedendo l' ultima seduta del suo governo.                         |

It must be observed that (c) and (d) above, although presented as quotations from Dante, are in fact misquotations: first of all, Purg. I,71 has “libertà **va** cercando”, not “**vo** cercando”; besides, the quotation in (d) has “**ma** è sì cara” instead of “**ch**’è sì cara”.

It appears that in journalistic jargon a construction has established itself, that can be described as follows:

Construction : X                      vo              cercando

X = Quality noun, mainly ending in *-tà*

In this construction the verb form has changed from 3SG.PRS.IND to 1SG.PRS.IND. The variant *vo* has been selected to fill the construction for metric reasons (it is monosyllabic like *va*), and possibly also to echo Dante’s usage of this form in other parts of the *Comedy*, such as the one below:

Original passage from Dante:

E io a lui: "I' mi son un che, quando  
Amor mi spira, noto, e a quel modo  
ch'e' ditta dentro vo significando". ( Purg. XXIV,52-54)

This passage too has been (mis)quoted twice in the *Repubblica* corpus:

- |    |  |
|----|--|
| a. | questo da molto tempo noto, e al modo che ditta dentro <b>vo</b> significando senza mai andare oltre somma soddisfazione la più inutile masturbatio grillorum.         |
| b. | Io noto e al modo che ditta dentro come Fidel Confalonieri <b>vo</b> significando: penso infatti come lui che Sacchi avrebbe potuto vincere un paio di scudetti in più |

Even here, the quotations alter slightly the original text (“ch’e’ ditta” becomes the *facilior* “che ditta”) but the person of the verb isn’t altered.

The remaining five tokens of *vo* in the *Repubblica* corpus are illustrated below:

Other tokens of *vo* in the *Repubblica* corpus

- |    |  |
|----|--|
| a. | Il meccanico si rannicchiava nell' abitacolo guardando atterrito i suoi piedi. Tazio si avventava alla curva gridando: " <b>Vo</b> foeura, <b>vo</b> foeura!", ma per intima onestà si impegnava tanto da restare in istrada, e così non moriva per suo amaro dispetto.<br>(Tazio Nuvolari, 1892-1953) |
|----|--|

b.	"E' inutile: bisogna che io rinunci alle osservazioni, per oggi. Io <b>vo</b> dietro agli incantesimi della piuma". (Gabriele D'Annunzio)
c.	LA METAMORFOSI DI FAUSTO MENO DI LOTTA , PIÙ DI GOVERNO. ROMA -- No, no e poi no. Al governo non ci <b>vo'</b> . Sorridente come sempre, e come sempre gentile, Fausto Bertinotti ha rifiutato con parole piene di grazia la poltrona di ministro [...]
d.	" [...] io facevo tutti quei film, dopo le vittorie dello scorso anno, perché era un gioco, perché mi divertivo. Se però diventa un marchio di fabbrica, un obbligo, io smetto subito. Se vinco, <b>vo'</b> sul podio, saluto, e basta così " (Valentino Rossi, born in Tavullia (Urbino) in 1979)
e.	Io ringhio, digrigno, bestemmio, <b>vo</b> a casa.

Example (a) is a quotation from a Northern dialect; example (b) is literary and not contemporary; in example (c) the selection of *vo* is motivated by rhyming constraints; (d) will have to be explained through some speaker-related factor I haven't identified yet; (e) is from a contemporary text with a literary flavour.

It must also be observed that most tokens of *vo* in the *Repubblica* corpus do not correspond to 'andare:1SG.PRS.IND' but to one of the singular forms of *volere*, in texts that are heavily marked as non standard, or, often, in quotations from songs or proverbs in some regional variety of Italian or a dialect. Some examples are given below:

#### *Vo* realizing forms of the verb VOLERE in the *Repubblica* corpus

volere: 3SG.PRS.IND	La cassetta più venduta è una compilation di tutte queste canzoni, compresa "O ciuccio <b>vo</b> vulà".
	Insomma, come si dice a Roma, nun ce <b>vo'</b> sta'.
	PARLA ANNA, LA BABY DONNA DELLA NAPOLI CAMORRISTA. NAPOLI -- "Chi <b>vo'</b> male a chesta casa addà murì primma che trase". Così c'è scritto accanto alla porta d'ingresso dell'abitazione di Anna
	ORA IL POLITICO PARLA COSÌ. ROMA -- Una volta tanto bisogna ammetterlo, la miglior sintesi l'ha trovata Berlusconi: "Quanno ce <b>vo'</b> ce vo".
volere: 1SG.PRS.IND	Nel sogno (raccontato da Alceta a Melisso in un frammento dei Canti: " Odi, Melisso, io <b>vo'</b> contarti un sogno...), il poeta vede la luna staccarsi improvvisamente dal cielo [...]
	Se l'antica anima di Prato se n'è andata -- non quella malapartesca di Maledetti Toscani, genere "io son di Prato / <b>vo'</b> essere rispettato / e posa il sasso, sai".
	E la Carmen: "Danzar <b>vo'</b> la seguediglia, e l'alicante assaggiar!". E la Salomè: "Le tue labbra vo' baciari!".
volere: 2SG.PRS.IND	Nelle strade risuona il frastuono dei clacson, nei bus si ascolta Carosone che canta Tu <b>vo'</b> fa l'Americano.

## 2.2. *faccio* / *fo*

*faccio* > 10.000 tokens      *fo* 63 tokens

All 63 tokens of *fo* in the *Repubblica* corpus were examined. 7 turned out to be false hits, constituted by misspellings such as < fo rmata >, < fo tografie > and the like. The remaining 56

tokens all realize ‘fare:1SG.PRS.IND’ and distribute in a way parallel to the tokens of *vo* ‘andare:1SG.PRS.IND’, as illustrated in Table IV.

Table IV Contexts of occurrence of *fo* ‘fare:1SG.PRS.IND’

17 tokens	<b>Tuscan speakers</b> (14), or speakers of other local varieties using <i>fo</i> (3)
13 tokens	<b>Quotations from (non-contemporary) literary and poetic texts</b>
8 tokens	<b>Proverbs and sayings</b> “Non lo fo per piacer mio ma per far piacere a Dio”
8 tokens	“fo per dire”
7 tokens	<b>Quotations from songs and opera pieces</b> “Porto il mantello a ruota e fo il notaio” ( <i>Signorinella</i> , 1931) 3 tokens
3 tokens	other cases (G. Brera, unknown journalist, V. Rossi)

### 2.3. Some U-pattern verbs

Data from the *Repubblica* corpus were collected for five verbs described as capable of displaying stem alternations in 1SG.PRS.IND, 3PL.PRS.IND, 1SG.PRS.SBJV, 3PL.PRS.SBJV: *dovere*, *sedere*, *possedere*, *vedere* and *chiedere*.

#### **dovere**

devo >10.000 tokens	debbo = 1388 tokens
devono >10.000 tokens	debbono = 6274 tokens
deva 14	debba > 10.000
devano 1	debbano 4696

The pattern with *dovere* is interesting, because the /v/ forms are in the majority in the indicative while the /bb/ forms predominate almost without exception in the subjunctive. It is possible that the obligation to use /bb/ forms in the subjunctive is the factor sustaining high usage of the /bb/ by-forms in the indicative as well. The indicative /bb/ forms do not appear restricted to specific diatopic, diachronic or diaphasic varieties. Nor are most of the very few /v/ forms in the subjunctive; but four tokens of *deva* occur in a satirical piece that lists excerpts of spoken speech heard on radio and TV programs, with the intent of mocking the speakers for uttering anacolutha or other sequences felt to be unacceptable. Apparently, the very usage of *deva* as ‘dovere:1SG.PRS.SBJV’ is considered worthy of mockery.

" Ritengo che non si **deva** mai pensare a un dualismo " . ( Fabio Capello -- Il processo di Biscardi -- 25 aprile , alle 20,40 )  
 " Non credo si **deva** fare " . ( Antonio Matarrese -- Tg 1 -- 14 maggio , alle 20,36 ) .  
 " Non è che uno **deva** solo fare dei compiti che gli vengono assegnati " . ( Roberto Baggio -- Guida al campionato -- 4 settembre , alle 13,41 ) .  
 " Non capisco perché Alba **deva** difendersi " ( 19 marzo -- ore 23,40 -- Canale 5 -- " Maurizio Costanzo show " -- Maurizio Mosca )

#### **sedere**

siedo 118 tokens	seggo 6 tokens
siedono 2008 tokens	seggono 47 tokens
sieda 156 tokens	segga 8 tokens
siedano 113 tokens	seggano 12 tokens

The /gg/ forms are in the minority with respect to the /d/ forms, but they do not seem to be restricted to specific diatopic, diachronic or diaphasic varieties. Two forms realizing the same cell often occur in comparable contexts, such as the following:

Vuol dire Berlusconi a trattare con De Mita, nonostante Craxi **sieda** a Palazzo Chigi.

i socialisti sono ben lieti che al Senato (oltre che al Quirinale) **segga** un democristiano, così sarà poi più difficile pretendere anche Palazzo Chigi

### possedere

possiedo 140	posseggo 95
possiedono 1236	posseggono 755
possieda 174	possegga 132
possiedano 46	posseggano 65

The gap between the two series of forms is minimal for this verb, and in the 3PL.PRS.SBJV the /gg/ form has more tokens than the /d/ form. As in the case of *sedere*, the /gg/ series does not appear to be restricted to specific diatopic, diachronic or diaphasic varieties.

### vedere

vedo > 10.000	veggo 9
vedono > 10.000	veggono 4
veda 1944	vegga 2
vedano 685	veggano 0

The very few tokens of forms with /gg/ all occur in quotations from non-contemporary authors, from Boccaccio, through Machiavelli, Leonardo, Galileo, up to Giovanni Gentile (1875-1944) and Emilio Cecchi (1884-1966) (most of these authors were Tuscan or had extensive contact with Tuscan varieties during their lives). The /gg/ forms appear to be extinct in contemporary usage.

### chiedere

chiedo 7402	chieggo 1
chiedono	chieggono 0
chieda	chiegga 0
chieggano	chieggano 0

The /gg/ forms are completely extinct in contemporary usage. The only token of *chieggo* in the *Repubblica* corpus is explicitly mentioned as dialectal and / or antiquated:

didascalie scritte con un curioso impasto di dialetto e italiano ottocentesco tutto "chieggo" e "veggo"

## 2.4. 1PL.PRS.IND

The corpus was investigated for 1PL.PRS.IND forms of verbs for which possible alternations are recorded in this cell. The data are the following:

doliamo 1	dogliamo 0
giaciamo 0	giacciamo 0
piaciamo 3	piacciamo 17
taciamo 6	tacciamo 11

These few data point to a preference for the forms *piacciamo* and *tacciamo* over their doublets with a non-geminated affricate. For *tacciamo* this seems to be a relatively recent development: Fornaciari (1906) gives *taciamo* as the only possible form, while modern grammars such as Serianni (1988) list only *tacciamo*. Neither *taciamo* nor *piacciamo* occur in contexts marked from the diatopic, diachronic or diaphasic point of view.

## 2.5. parvi / parsi

The corpus was searched for all possible Passato Remoto 1/3SG and 3PL forms of *parere*. The only attested forms are the following:

parvi 0            parve 730            parvero 106

Neither sigmatic forms (*\*parsi*, *\*parse*, *\*parsero*) nor weak forms (*\*parei*/*\*paretti* type) are attested.

A somewhat different picture emerges with *apparire*, which has the following forms in the relevant cells:

apparii 0	apparvi 0	apparsi* 0/20
apparì 15	apparve 1341	apparse* 1/20
apparirono 10	apparvero 300	apparsero 0

*\*apparsi* and *apparse* are ambiguous forms that can also realize ‘apparire:PST.PTCP.M.PL’ and ‘apparire:PST.PTCP.F.PL’; both forms are very frequent (*apparsi* 2265, *apparse* 1805); inspection of the first 20 random tokens of each revealed only 1 token of Passato Remoto, which does not appear to be marked from the diatopic, diachronic or diaphasic point of view:

Noi ricordiamo il presidente Pertini quando **apparse** alla televisione per ricordare agli italiani che la campagna antisionista ed antisemita di allora era indegna di un paese civile.

## 2.6. Syncopated vs. unsyncopated Future and Conditional forms

The verbs described by Fornaciari (1906) as having syncopated forms in the Future, in the Present Conditional, or in both, were investigated. The results are as follows:

### andare

No syncopated forms attested in the *Repubblica* corpus for the Present Conditional.

Two forms attested in the future, *anderò* and *anderemo*, both in quotations from *Pinocchio*. A second token of *anderemo*, in a quotation attributed to Alberto Michellini, an Italian journalist and politician born in Rome in 1941, is probably a typo.

### godere

No syncopated forms attested in the *Repubblica* corpus for the Present Conditional.

A single token of *goderete* is in a quotation of a text written in 1614:

una citazione da una lettera che Giovanni Antonio Oldelli invia al fratello Giovanni Andrea a Colmar, 12 marzo 1614, tramite un messer Francesco spinazaro, cioè fabbricante di pettini da lino: mando para tre guanti e quelli che ho portato io di Roma, insieme una meza donzena di medaglie benedette, che sono di tanta indulgencia, et di tanto valore che non ponno esser di più, quali li **goderete** per amor mio et ne farete parte alli amici; qua non vi è altro di novo, solo che vi è stato tanta neve [...]

**udire**

The Present Conditional forms are very rare in the *Repubblica* corpus: there is 1 token of *udireste* and 1 token of *udrebbe*, in a translation of a poem by Rilke.

In the Future, the following forms are attested:

udirò 0	udirò 0
udirai 2	udirai 0
udirà 21	udirà 3
udremo 5	udiremo 1
udrete 1	udirete 3
udiranno 5	udiranno 0

The syncopated forms prevail in the future and do not appear to be specially marked from the diachronic or diatopic point of view. The entire verb UDIRE is somewhat diaphasically marked as literary with respect to SENTIRE ‘hear’ (among the first 2000 verbs in LIP, SENTIRE has rank 17 while UDIRE is absent).

**morire**

The following forms are attested:

## FUTURE

morirò 5	morirò 162
morrai 3	morirai 32
morrà 44	morirà 944
morremo 3	moriremo 111
morrete 0	morirete 27
morranno 3	moriranno 230

## PRESENT CONDITIONAL

morrei 3	morirei 23
morresti 0	moriresti 1
morrebbe 5	morirebbe 64
morremmo 0	moriremmo 9
morreste 0	morireste 0
morrebbero 1	morirebbero 35

The unsyncopated forms prevail in both the future and the conditional; the syncopated forms are sometimes found in non-contemporary texts, but this is not always the case. They cannot be considered extinct yet.

**2.7. Strong vs. weak Passato Remoto**

Several verbs are described as possessing alternative forms of Passato Remoto. Different sub-types are reviewed below.

**2.7.1. -ire verbs**

A few *-ire* verbs are described as having a “sigmatic” Passato Remoto alongside a regular one in 1/3SG and 3PL forms. The corpus data are summarized in Table V:

Table V “Sigmatic” vs. regular Passato Remoto of *-ire* verbs

Verb	“Sigmatic” Passato Remoto	Regular Passato Remoto
------	---------------------------	------------------------

	(tokens)	(tokens)
<i>aprire</i>	20	2023
<i>assalire</i>	--	64
<i>convertire</i>	--	128
<i>offrire</i>	31	1139
<i>scolpire</i>	--	40

Only *aprire* and *offrire* have sigmatic forms in contemporary usage. These forms are in the minority with respect to regular forms, but do not appear to be marked from the diatopic, diachronic or diaphasic point of view (at least, not more marked than the very usage of Passato Remoto itself, on which cf. Gambarara (1994)).

### 2.7.2. “Sigmatic” forms of *-ere* verbs

An opposite picture emerges from inspection of the *-ere* verbs described as possessing both “sigmatic” and regular Passato Remoto forms in 1/3SG and 3PL. The corpus data are summarized in Table VI:

Table VI “Sigmatic” vs. regular Passato Remoto of *-ere* verbs

Verb	“Sigmatic” Passato Remoto (tokens)	Regular Passato Remoto (tokens)
<i>annettere</i>	2/20 <i>annesse*</i> (199)	1 <i>annettei</i>
<i>appendere</i>	12 <i>appesero</i> + 1/20 <i>appese*</i> (558)	--
<i>assolvere</i>	149	--
<i>concedere</i>	61 <i>concessero</i> + 7/20 <i>concesse*</i> (1993)	1 <i>concedette</i>
<i>discutere</i>	52 <i>discussero</i> + 1/20 <i>discussi*</i> (1375) + 4/20 <i>discusse*</i> (1346)	--
<i>incutere</i>	1 <i>incusse</i>	1 <i>incuté</i>
<i>perdere</i>	621 <i>persero</i> + 1/20 <i>persi*</i> (2338) + 14/20 <i>perse*</i> (3459)	162 (cf. Table VII)
<i>rendere</i>	433 <i>resero</i> + 2/20 <i>resi*</i> (6849) + 5/20 <i>rese*</i> (5413)	--

\* These forms are ambiguous with PST.PTCP forms. The number of PRF tokens over the first 20 random tokens is given for each form; the total number of tokens of the word form is given in parentheses.

Clearly, sigmatic forms dominate for most verbs, and “regular” forms are at most hapaxes. The only exception is *perdere*, the most frequent verb in this group, and among the top 20 most frequent *-ere* verbs overall, which displays a fair amount of variation in the Passato Remoto, as shown in Table VII:

Table VII Passato Remoto forms of *perdere*

1SG			3SG			3PL		
<i>persi</i>	<i>perdei</i>	<i>perdetti</i>	<i>perse</i>	<i>perdé</i>	<i>perdette</i>	<i>persero</i>	<i>perderono</i>	<i>perdettero</i>
1/20* (2338)	1	7	14/20* (3459)	<i>perdé</i> 6 <i>perdè</i> 1	121	621	0	26

				total 7				
--	--	--	--	---------	--	--	--	--

\* These forms are ambiguous with PST.PTCP forms. The number of PRF tokens over the first 20 random tokens is given for each form; the total number of tokens of the word form is given in parentheses.

### 2.7.3. BERE

Table VIII Passato Remoto forms of *bere*

1SG			3SG			3PL		
<i>bevvi</i>	<i>bevei</i>	<i>bevetti</i>	<i>bevve</i>	<i>bev�</i>	<i>bevette</i>	<i>bevvero</i>	<i>beverono</i>	<i>bevettero</i>
7	--	--	61	1	2	11	--	--

Strong forms clearly prevail: the hapax *bev * is due to metric reason, as it occurs in the following quotation from the *Inno di Mameli*:

le il sangue Polacco – **bev ** col Cosacco -- ma il cor le bruci  .

### 2.7.4 DARE

Table IX Passato Remoto forms of *dare*

1SG		3SG		3PL	
<i>diedi</i>	<i>detti</i>	<i>diede</i>	<i>dette</i>	<i>diedero</i>	<i>dettero</i>
244	4/20* (3067)	3725	10/20* (4266)	971	381

\* These forms are ambiguous with PST.PTCP forms of *dire*. The number of ‘*dare*:PRF’ tokens over the first 20 random tokens is given for each form; the total number of tokens of the word form is given in parentheses.

Both variants are well-attested; preferences are hard to assess for 1SG and 3SG forms for practical reasons.

## 2.8. Strong vs. regular Past Participle

### 2.8.1. -ire verbs

Table X “Strong” vs. regular Past Participle of -ire verbs

Verb	“Strong” Past Participle	Regular Past Participle
<i>apparire</i>	√	1 <i>apparita</i> in quotation
<i>assorbire</i>	<i>assorto</i> is an ADJECTIVE	√
<i>costruire</i>	<i>costrutto</i> is a NOUN	√
<i>scolpire</i>	* <i>sculto</i> : 0 tokens	√
<i>seppellire</i>	√ <i>sepolto</i>	√ <i>seppellito</i>

The most interesting case is represented by *sepolto* and *seppellito*, true doublets that both serve as PST.PTCP forms. The data from the corpus are the following:

sepolto 2342	seppellito 640
sepolti 1110	seppellita 161
sepolti 1331	seppelliti 260
sepolte 513	seppellite 83 (this form is ambiguous with 2PL.PRS.IND)

In order to assess whether one of the two forms is specialized in adjectival vs. verbal usage, the corpus was searched for typically verbal vs. typically adjectival collocations:

	<i>sepolto</i>	<i>seppellito</i>
typically verbal usage	<i>ha sepolto</i> 135 (5,8%) <i>è stato sepolto</i> 144 (6,1%)	<i>ha seppellito</i> 100 (15,6%) <i>è stato seppellito</i> 47 (7,3%)
typically adjectival usage	<i>è sepolto</i> 292 (12,5%)	<i>è seppellito</i> 23 (3,6%)

Percentages are on the total tokens of the respective form

The data show that *sepolto* is preferred as an adjective and *seppellito* is preferred as a verb form, but both forms can be used in both capacities. This is a true case of doublets.

### 2.8.2. -ere verbs

Table XI “Strong” vs. regular Past Participle of -ere verbs

Verb	“Strong” Past Participle	Regular Past Participle
<i>concedere</i>	√	3 tokens (archaic)
<i>cuocere</i>	√	--
<i>parere</i>	√	--
<i>perdere</i>	√	√
<i>rendere</i>	√	2 tokens (archaic)
<i>spandere</i>	√	--
<i>vedere</i>	√	√

The general trend is very clear: the strong forms predominate, and regular forms are archaic, if attested at all. But there are two exceptions, which constitute true cases of doublets: *perso / perduto* and *visto / veduto*.

The corpus data are the following:

<i>perduto</i>	<i>perso</i>	<i>perduta</i>	<i>persa</i>	<i>perduti</i>	<i>persi*</i>	<i>perdute</i>	<i>perse*</i>
> 10.000	> 10.000	3918	3508	1962	19/20 (2338)	1572	6/20 (3459)

\* These forms are ambiguous with PRF forms. The number of PST.PTCP tokens over the first 20 random tokens is given in the first line of the cell; the total number of tokens of the word form is given in parentheses.

Data for *vedere* are more difficult to analyze because most forms are ambiguous with forms of the nouns *veduta* ‘view’, *vista* ‘eyesight’ and *visto* ‘visa’. Consider, however, the following data:

<i>veduto</i>	<i>visto*</i>	<i>veduti</i>	<i>visti*</i>
476 tokens	20/20 (> 10.000)	50 tokens	18/20 (> 10.000)

\*The number of PST.PTCP tokens over the first 20 random tokens is given in the first line of the cell; the total number of tokens of the ambiguous word form is given in parentheses.

Overall, it can be concluded that *perso/perduto* represent real doublets (of comparable frequency) in the Italian verb system; *visto/veduto* also constitute doublets, but the *-uto* forms are much less frequent.

## 2.9. “Dittongo mobile” in SEDERE

*Sedere* is a textbook case of N-pattern distribution of the two stems *sied-* and *sed-*. However, outside the N-pattern cells, the *sied-* stem spreads to cells where it wouldn't be expected, i.e., where it is unstressed, as the data below illustrate.

	Future		Present Conditional		Imperfect Indicative		Imperfect Subjunctive	
	-ie-	-e-	-ie-	-e-	-ie-	-e-	-ie-	-e-
1sg	26	2	4	1	--	37	--	4
2sg	1	1	--	--	4	--	--	
3sg	409	59	18	7	41	989	--	59
1pl	35	4	--	--	2	29	--	5
2pl	2	1	--	--	--	--	--	1
3pl	381	62	12	4	24	473	--	41
Total	854	129	34	12	71	1528	--	110

In the Future, there is a full pattern of doublets; very likely, the same holds for the Present Conditional and the Imperfect Indicative, even though forms for certain cells are not attested in the corpus; the Imperfect Subjunctive, on the contrary, seems to be totally immune to this kind of doublet.

## 3. Forms with two different inflectional endings in the same cell

### 3.1. The *temei / temetti* type

Conjugation tables appearing in Italian grammars and dictionaries suggest that alternation between the endings *-ei/-é/-erono* and *-etti/-ette/-ettero* in the cells for ‘1SG/3SG/3PL.PRF.IND’ of second conjugation verbs is free (“variazione libera” according to Salvi / Vanelli 2004: 94). Both sets of endings are listed as possible, usually arranged after the stem of *temere* (the only “regular” verb of second conjugation). Grammars (and dictionaries) differ in regard to which set is given precedence (implying preference?) or prominence, as already observed by Flamm (1987) and as shown below for a selection of recent authoritative works:

Lepschy / Lepschy 1981:	<i>cred-ei (-etti)</i>	
Serianni 1988	<i>io tem-ètti (tem-éi)</i>	
GRADIT (De Mauro 1999)	<i>io accedetti o accedei</i>	
Maiden / Robustelli 2000	<i>temei / temetti</i>	
Salvi / Vanelli 2004	BT-èi / -ètti	[BT = Base Tematica]

A set of 53 verbs listed in GRADIT as following the regular 2nd conjugation scheme was investigated. The data show that the two sets of endings are far from being in free variation in the corpus. Their distribution is **almost complementary**, and is regulated by a phonological factor: verbs whose stem ends in /t/ select the /t/-less set of endings, and vice versa (this tendency has been observed by several grammars, and was demonstrated by Flamm (1987) on corpus data).

The data for the two groups of verbs are listed in Tables XII and XIII:

Table XII Verbs with stem in /t/-

-ei	-etti	-é	-ette	-erono	ettero
51	0	1347	6	352	2

The 8 “ungrammatical” forms in Table XII are: *battettero*, *ripetette* (5), *ripetettero*, *fottette*. *Fottette* occurs in a context which appears to have undergone interference from Neapolitan: “Il generale Clark si **fottette** dalla paura”, said by Aldo Sersale, described as “l’anima incarnata di Positano” by the journalist quoting him. It is therefore not a standard Italian form. The other forms belong to top-frequency verbs (rank 1 and 3 among *-ere* verbs with a stem ending in /t/). The least restricted ending has spread to these top-frequency verbs, indicating a tendency to reduce the variation between the two sets of endings in favour of the series containing /tt/.

Table XIII Verbs with stem not ending in /t/

-ei	-etti	-é	-ette	-erono	ettero
1	278	18	4295	0	802

The 19 “ungrammatical” forms in Table XIII are: *credei*, *ricevè* (3) / *ricevé* (3), *vendé* (1) / *vendè* (2), *credé* (1), *cedè* (2), *temè* (1), *procedè* (1), *presiedé* (1) / *presiedè* (2), *tessè* (1).

*Credei* is the only non 3SG form attested, and this attestation occurs in a recently published text written by Giovanni Pascoli in the years between 1877 and 1879 – therefore, it is not attested in contemporary usage as represented by the *Repubblica* corpus.

Concerning the 18 3SG forms attested in the corpus, which are on the brink of ungrammaticality, it must be observed that 12 of them are also misspelled, displaying <è> instead of <é> in final position. This might be a signal that they were produced by writers not in complete command of standard written Italian.

Further data on the *-ei/-etti* alternation are represented by the occasional weak forms of *-ere* verbs with a strong Passato Remoto, shown in Tables VI, VII and VIII above. For the lower-frequency verbs *annettere*, *concedere*, *incutere* the complementary distribution of weak endings according to the phonological environment has no exceptions; for the high-frequency verb *perdere* the tendency to use the phonologically predictable endings is very strong, but a few tokens of the alternative endings are attested; the few weak tokens of *bere* distribute similarly, but weak forms seem to be avoided altogether for this verb.

### 3.2. *va’ / vai, fa’ / fai, sta’ / stai* ‘2SG.IMP’

Both the alternative imperative forms of *andare*, *fare* and *stare* are extremely frequent and homophonous with forms of the indicative. Therefore, for practical reasons it was not possible to examine all their tokens in the *Repubblica* corpus. Some data were collected analyzing the forms occurring in collocations that are good candidates for containing an imperative form of these verbs. The data are presented in Table XIV.

Table XIV Percentage of imperative tokens for short verb forms in four collocations

collocation	% imperative
va pure	1% (poetic)
vai pure	93,5%
sta attento	30 %*
stai attento	80%*
fa pure	0,6%
fai pure	77,3%
sta buono	33,3%
stai buono	66,6%

\* percentage on the first 20 random tokens of the collocation

Addio poesia , va pure / tutto fra noi è finito (Cardarelli)
--

The data show that forms in *-i* prevail as imperatives in all cases, even though forms in *-a* are never excluded. The two forms can be used by the same speaker, as shown in the following quotation:

" <b>Stai</b> calmo, <b>stai</b> zitto, <b>sta</b> attento a quello che dici"
---

Here the speaker is a non-native speaker, Ali Agca; but his speech seems to have been heavily edited by the journalist. All the quotations attributed to Agca in the text are perfectly grammatical, so there is no reason to doubt the grammaticality of this quotation, at least according to the journalist who included it in his article.

#### 4. Some conclusions

The data presented above give a picture of the doublets actually attested in contemporary written Italian (at least, in so far as it is represented by the *Repubblica* corpus).

**The number of verbs that have doublets in their paradigms is not very high**, in line with the expectation that uniqueness of realization should not be violated. Less than 30 verbs have doublets which have a frequency of 3 or more tokens each in the corpus. These verbs are shown in Table XV.

Table XV Verbs that have doublets with a frequency  $\geq 3$  in the *Repubblica* corpus, and their frequency rank in several databases

Verb	Rank in COLFIS (on all word-classes)	Rank among first 2000 verbs in LIF (on verbs only)*	Rank among first 2000 verbs in LIP (on verbs only)*
fare	21	2	2
dovere	42	8	8
andare	63	9	5
stare	64	13	11
vedere	65	10	9
dare	73	12	12
perdere	182	48	59
aprire	255	46	71
morire	331	29	68
offrire	332	85	156
piacere	339	50	60
apparire	380	120	347
ricevere	398	133	213
ripetere	560	119	83
vendere	645	186	77
sedere	1125	91	293
possedere	1278	238	--
bere	1413	111	338
tacere	2406	330	--
udire	2694	326	--
seppellire	2778	390	--
presiedere	3638	--	--
mentire	4784	338	--

inghiottire	5013	--	--
aborrere	19960	--	--
languire	23965	--	--

**COLFIS** = a 3.000.000 tokens frequency dictionary of written Italian

LIF = a 500.000 tokens frequency dictionary of written Italian

LIP = a 500.000 tokens frequency dictionary of spoken Italian

\*These data were computed from a database kindly made available by Miriam Voghera.

As the data in Table XV show, **in many cases doublets appear in the paradigms of top-frequency verbs** (as happens with other non-canonical phenomena, such as suppletion: cf. Corbett 2006). The main exception to the correlation between the existence of doublets in the paradigm and high frequency is represented by verbs that have *-isc-* and *-isc-less* forms, some of which are not frequent at all. A possible explanation is that uncertainty about inflectional class goes hand in hand with low frequency, while stem alternations must be supported by high frequency to be learned. Third conjugation verbs overall tend to have a lower ranking than second conjugation verbs: for example, among the first 100 top rank verbs in LIF, there are 30 *-ere* verbs but only 14 *-ire* verbs. And among ***-ire* verbs**, a tendency to reduce doublets in a way that yields a parallel **reduction of stem allomorphy** is observed:

- *-isc-* forms are less numerous than *-isc-less* forms (in spite of the fact that *-isc* forms would be the regular forms in the third conjugation – but they would introduce stem allomorphy! cf. Table II);
- regular Passato Remoto forms of *aprire* and *offrire* prevail by far over sigmatic ones (cf. Table V);
- unsyncoated Future and Conditional forms of *morire* prevail over syncoated ones (cf. § 2.6).

On the other hand, pairs of doublets in which the form that introduces stem allomorphy is more frequent than the “regular” form are represented also among *-ire* verbs, in the following cases:

- Passato Remoto of *apparire*, where strong forms prevail (§ 2.5);
- Future of *udire*, where syncoated forms prevail (§ 2.6)
- Past Participle of *seppellire*, where *sepolto* is more frequent than *seppellito* (§ 8.1).

Therefore, it is not easy to establish a tendency in the direction of changes aimed at eliminating doublets: among *-ire* verbs, in some cases there appears to be a tendency to reduce stem allomorphy, but in other cases there appears to be a tendency to preserve it.

A definite **tendency to preserve stem allomorphy** is observable in doublets occurring in the Passato Remoto and Past Participle of *-ere* verbs and in the Past Participle of *-ire* verbs: in all cases, the strong forms are more frequent than the “regular” ones (cf. Tables VI-XI).

Another characteristic feature of all kinds of doublets in Italian verb paradigms is that they **tend to occur in cells that belong to a partition class other than the default one**: there are no doublets, for instance, in 2PL.PRS.IND or in the Imperfect Subjunctive and the Infinitive. Doublets, as we have seen, typically occur in N-pattern or U-pattern cells, in 1SG/3SG/3PL.PRF, and in Past Participle. An exception to this pattern is the 2SG.IMP of *stare*, which has a doublet in a cell that does not constitute a partition class of its own in any verb (while doublets in the 2SG.IMP of *fare* and *andare* are less exceptional, as these verbs have some variant of the N-pattern). Like what has been found by Hippiisley et al. (2004) concerning exceptional cases of suppletion, the exceptionality of having doublets in a cell not normally containing them is accompanied by unexceptional behavior in other respects: *stare* is a top-frequency verb, i.e., a kind of verb for which having doublets in the paradigm is not exceptional. Truly exceptional is the case of doublets in the Future, Present Conditional and Imperfect Indicative of *sedere* (cf. § 2.9), where doublets appear in all cells of entire tenses, and of tenses that normally build their forms on the default stem. But here again, the exceptional presence of doublets in this case is accompanied by a non-exceptional feature, i.e., the

fact that *all* the cells have doublets. In some sense, the question comes down to an uncertainty as to which stem is the default one. Once a stem (*sied-* or *sed-*) is recognized as default, it is used to realize forms for all the cells in the tense.

In some cases, such as *vado / vo* and *faccio / fo*, the corpus data show that the two forms are not alternative forms that can be used by the same speaker in the same circumstances: one form is marked as belonging to a specific local variety, or to a non-contemporary variety (and therefore the two forms do not belong to the same I-language), and/or as belonging to specific text types, such as literary texts, poetry, songs, proverbs and sayings. This last case is interesting: forms marked as literary or poetic (cf. Serianni 2001 for a thorough presentation of poetic forms in Italian), or occurring only in songs and proverbs, could be said not to belong to the active repertoire of contemporary Italian speakers. On the other hand, the continual quotation of literary texts, poems, songs and proverbs containing these forms prevents them from exiting the passive repertoire of speakers (as has happened to forms such as *chieggo*, *veggo*, *sculto*), and occasionally allows these forms to resurface in contemporary usage (as in the *X-tà vo cercando* construction illustrated in §2.1.), preserving the existence of doublets in their cells.

## 5. Canonicity

The ideal pair of doublets is a pair of forms such that any speaker could use interchangeably one or the other to express the same meaning / realize the same cell of the paradigm of a given lexeme.

So a first criterion of canonicity for doublets is:

**Criterion 1:** no conditions > conditions

(where > = “more canonical than”)

(cf. criterion 13 in Corbett 2007: no outside conditions > conditions)

If one form in a pair of doublets is limited to certain text types (literary, poetic, used only in songs) or archaic or limited to some regional / local variety of the language, this pair is less canonical.

Any element that differentiates the conditions of usage of the two forms in a pair of doublets makes this pair a less canonical instance of doublets (the ultimate outcome of differentiation will be the establishment of two separate lexical items – lexical split, the opposite of incursion in suppletion).

If we assume this definition, some possible canonicity hierarchies, with examples from Italian verbs, are the following:

- diaphasically unrestricted > restricted  
*devo / debbo* > *vado / vo* (*vo* restricted to poetry, songs....)
- diachronically unrestricted > restricted  
*devo / debbo* > *chiedo / chieggo* (*chieggo* ancient)
- diatopically unrestricted > restricted  
*devo / debbo* > *vado / vo* (*vo* restricted to Tuscan variety)
- phonologically unconditioned > phonologically conditioned  
-isc/-isc-less doublets > -ei/-etti doublets
- semantically unconditioned > semantically conditioned  
*perso / perduto* > *sepolto / seppellito*  
(*sepolto* more frequent in adjectival usage, *seppellito* more frequent in verbal usage)

Is this a case of semantic conditioning or syntactic conditioning?

- morphologically unconditioned > morphologically conditioned  
Cf. Criteria 2 and 3 below
- syntactically unconditioned > syntactically conditioned

Sieda/segga > ??? [are there any cases of syntactically conditioned doublets? Maybe *sepolto / seppellito*?]

(*sieda/segga* is syntactically unconditioned, cf. corpus examples § 2.3)

A corollary of criterion 1:

- doublets of roughly equal frequency > doublets of very different frequency  
*sepolto / seppellito* > *visto / veduto*

Doublets interact with other non-canonical phenomena: first of all, with suppletion, as the existence of a pair of doublets often involves the possibility of having a regular vs. a suppletive form in a cell. Therefore, some variants of the criteria of canonicity established by Corbett (2007) for suppletion could apply to doublets.

**Criterion 2:** doublets in a single cell > doublets in several cells > doublets in all cells

cf. Criteria 9 and 12 in Corbett (2007):

- 9: more features > fewer features  
12: unique > non unique

Italian data:

*Stai / stai* ‘stare:2SG.IMP’ > *devo/debbo, devono/debbono* ‘dovere:1SG/3PL.PRS.IND’ > Future of *sedere*

If doublets occur in all cells, we could say it’s two lexical items (cf. the notion of “verbo sovrabbondante” in Italian grammars: *arrossare / arrossire*; cf. also the peculiar noun ORECCHIO which has both sg. and pl. forms in both genders: *orecchio / orecchia // orecchi / orecchie*)

The case of the Future of Italian *sedere* (§ 2.9 of the poster) is mid-way between ORECCHIO and some case in which only one cell in a paradigm contains doublets: in SEDERE all the PS/N forms are double in a given tense (maybe in three given tenses), but in other tenses of the same verb there are no doublets (there’s morphomic allomorphy or simple regularity).

*Stai / stai* is the only cell containing doublets in the paradigm of *stare*.

**Criterion 3:** doublets occur in unpredictable cells > doublets occur in cells that constitute a morphomic pattern > doublets occur in morphosyntactically defined cells

Cf. Corbett 2007 criterion 5: morphological distribution > morphosyntactic distribution

*stai/stai* 2SG.IMP of STARE >

*debbo/devo + debbono/devono* 1SG/3PL.PRS.IND of DOVERE (U-pattern distribution) >

Future of *sedere* (both *sied*-based and *sed*-based forms in all cells)

Examples of decreasing canonicity are the same for Criterion 2 and Criterion 3, which shows that the two criteria converge (as happens with several of the criteria established for suppletion).

**Criterion 4:** more variants > fewer variants

(identical to criterion 4 in Corbett 2007)

perse / *perdé* / *perdette* > *diede* / *dette*

- form built on default stem / form built on allomorphic stem vs. two forms built on two different allomorphic stems

apparì / *apparve* vs. *apparve* / *apparse*  
Which case is more canonical ?

**Criterion 5:** unparalleled doublets > parallel doublets in several paradigms

cf. Criterion 12 in Corbett (2007):

12: unique > non unique

*Devo/debbo, sepolto/seppellito* > *siedo/seggo* + *possiedo / posseggio*

Interaction with syncretism:

The 1/2/3SG.PRS.SBJ forms are always syncretic in Italian, even in the most irregular paradigms.

When there are doublets in these cells, is syncretism maintained?

Very likely yes; yes in my intuition, but the few forms in the corpus that would allow to check for this do not help. The cases in point are *deva* and the forms that could realize 1/2/3SG.PRS.SBJ in verbs displaying the *-isc-/-isc-less* alternation. Unfortunately, all tokens of *deva* realize 3SG.PRS.SBJ in the corpus, and no *-isc-/-isc-less* has enough tokens of both the forms that could in principle realize 1/2/3SG.PRS.SBJ to check for this factor.

It must be observed that forms realizing 1SG.PRS.SBJ and 2SG.PRS.SBJ are attested in the corpus for *debba* (*io debba* = 88 tokens, *tu debba* = 17 tokens), but are by far less frequent than tokens of *debba* realizing 3SG.PRS.SBJ; therefore, as *deva* is overall of extremely low frequency, its being attested only in forms realizing 3SG.PRS.SBJ could be just an effect of insufficient sample size, not a grammatical fact.

Interaction with heteroclisis:

Concerns the verbs with *-isc-/-isc-less* alternation

Maybe also concerns *sedere*?

Several criteria converge on establishing the following pairs as the most canonical cases of doublets in Italian verb inflection:

*stai / sta* ‘stare:2SG.IMP’

*perso / perduto*

*devo / debbo, devono / debbono*

*perse / *perdé* / *perdette**

*sepolto / seppellito*

## References

Corbett, Greville G., 2005, “The canonical approach in typology”, in *Linguistic Diversity and Language Theories*, ed. by Zygmunt Frajzyngier, Adam Hodges, David S. Rood. Amsterdam / Philadelphia, Benjamins, 25-49.

- Corbett, Greville G., 2007a, "Canonical typology, suppletion, and possible words", *Language* 83.1, 8-41.
- Corbett, Greville G., 2007b, "Deponency, Syncretism, and What Lies Between", *Proceedings of the British Academy* 145, 21-43.
- De Mauro, Tullio, 1976 [1963], *Storia linguistica dell'Italia unita*, Roma / Bari, Laterza.
- Flamm, Hans Werner, 1987, "Dovei o dovetti?", *Lingua nostra* XLVIII.1, 20-25.
- Fornaciari, Raffaello, 1906, *Grammatica italiana dell'uso moderno*, Firenze, Sansoni.
- Gambarara, Daniele, 1994, "Il passato remoto nell'italiano parlato", in De Mauro, T. (a cura di), *Come parlano gli italiani*, Scandicci (FI), La Nuova Italia, 183-194.
- GRADIT = *Grande dizionario italiano dell'uso*, ideato e diretto da Tullio De Mauro, Torino, UTET, 1999 [2007<sup>3</sup>]
- Hippisley, Andrew / Chumakina, Marina / Corbett, Greville G. / Brown, Dunstan, 2004, "Suppletion. Frequency, categories and distribution of stems", *Studies in Language* 28.2, 387-418.
- Lepschy, Anna Laura / Lepschy, Giulio, 1981, *La lingua italiana. Storia, varietà dell'uso, grammatica*, Milano, Bompiani.
- Maiden, Martin / Robustelli, Cecilia, 2000, *A Reference Grammar of Modern Italian*, London, Arnold.
- Salvi, Giampaolo / Vanelli, Laura, 2004, *Nuova grammatica italiana*, Bologna, Il Mulino.
- Serianni, Luca (con la collaborazione di Alberto Castelvechi), 1988, *Grammatica italiana*, Torino, UTET.
- Serianni, Luca, 2001, *Introduzione alla lingua poetica italiana*, Roma, Carocci.
- Spina, Rossella, 2007, *L'evoluzione della coniugazione italo-romanza. Uno studio di morfologia naturale*, Catania, ED.IT.