GSR and Suppletion in Bolognese Clitics

Edward Rubin & Aaron Kaplan
University of Utah
e.rubin@utah.edu, a.kaplan@utah.edu

WCCFL 41
May 6, 2023

1 GSRs and Suppletion

- Gradient Symbolic Representations (GSRs; Smolensky & Goldrick 2016) allow phonological entities to be "partially present" in the input.
- This theory has been shown to account for certain kinds of phonologically conditioned morphological phenomena (Faust & Smolensky 2017, Zimmermann 2019).
- We apply this framework to clitic allomorphy in Bolognese (Romance; Italy) to assess its ability to account for suppletion.
- Bolognese makes a good test case: DEP penalizes both the appearance of a suppletive allomorph and epenthesis, which sometimes occurs as an alternative to suppletion.

2 Bolognese Clitics

- Bolognese has a fairly standard Romance clitic inventory:

(1) Clitic Pronouns in Bolognese

	NOM		DAT		ACC		PRT
	SING	PLUR	SING	PLUR	SING	PLUR	
1	a=/=ja	a=/=ja	m	s	m	s	
2	t	a=/=v	t	v	t	v	
3M	[(a)l]	i	i	i	[(a)l]	i	n
3f	l(a)	æl/æʎ	i	i	l(a)	i	
3RFLX			s	s	s	s	

- Our focus: interaction between allomorphy of 3MS.NOM and 3MS.ACC
- Both clitics display suppletion.

- Data in this work comes from Canepari & Vitali (1995), Vitali (2009), and from extensive work with native speakers.

3 Phonotactics

- Bolognese prohibits sonorant-final coda clusters:

(2)

tɛːvla	'table'	tɛːvel	'tables'
laŋteːrna	'lantern'	laŋteːreŋ	'lanterns'
liːvra	'hare'	liːver	'hares'

- Sonorant-initial onset clusters are also banned (except for a handful of root-internal [mC] clusters; e.g. [mdajɑŋ] 'medallion'). None exist underlyingly; epenthesis is visible with clitics:

(3) a. al= le= vad
3MS.NOM= 3MS.ACC= sees
'he sees him.'

b. al= le= tra
3MS.NOM= 3MS.ACC= throws
'he throws it.'

- Probably not a sonority sequencing fact (e.g. Clements 1991, Selkirk 1984): clusters that disobey sonority sequencing requirements are not rare (Rubin & Kaplan to appear):

(4)

zbdɛl	'hospital'
ftleŋna	'slice'
tskɲɔser	'to disavow'
vdand	'seeing'
forbz	'scissors'
poːrdg	'portico'

- We adopt the following constraint:

(5) *[+son]PERIPHERY: no sonorant-initial onset clusters or sonorant-final coda clusters.

4 Clitic Allomorphy: The Basics

4.1 3MS.NOM

- Prevocalic: [l] (6)
- Preconsonantal: [al] (7)

(6)

[l]=	arspɑnd
3MS.NOM=	responds

'he responds'

(7)

[al]=	vad
3MS.NOM=	sees

'he sees'

⇒ These are suppletive: no regular phonological process in Bolognese accounts for [a] epenthesis/deletion (Rubin & Kaplan 2022).

- [l] also appears post-verbally (e.g. in questions) with consonant-final verbs. Epenthesis is triggered by *[+son]PERIPHERY, which would not have been necessary with [al]:

(8) vad=el, *vad=al 'Does he see?'

- Our claim: [l] appears to avoid misalignment of [al] with respect to syllable boundaries:
  - *[a.l=arspɑnd] (cf. (6)): syllable boundary in the middle of the clitic
  - *[va.d=al] (8): clitic is not left-aligned with a syllable boundary
- The cover constraint ALIGN-[al]$_{NOM}$ penalizes both configurations.
- Ostensibly, a third allomorph [a] occurs before certain ACC and DAT clitics:

(9) a.

[a]=	m=	la=	da
3MS.NOM	1S.DAT	3FS.ACC	gives

'he gives it to me.'

b.

[a]=	t=	la=	da
3MS.NOM	2S.DAT	3FS.ACC	gives

'he gives it to you.'

c.

[a]=	s=	al=	da
3MS.NOM	1P.DAT	3MS.ACC	gives

'he gives it to us.'

d.

[a]=	v=	al=	da
3MS.NOM	2P.DAT	3MS.ACC	gives

'he gives it to you.'

- Rubin & Kaplan (2022): 3MS.NOM fuses with these (and other) clitics: [am], [as], etc., are single lexical items—"duplexes" that are the exponent of two sets of pronominal features.
- Revisions to (9) with the duplex analysis:

(10) a.

[am]=	la=	da
{3MS.NOM, 1S.DAT}	3FS.ACC	gives

'he gives it to me.'

b.

[at]=	la=	da
{3MS.NOM, 2S.DAT}	3FS.ACC	gives

'he gives it to you.'

c.

[as]=	al=	da
{3MS.NOM, 1P.DAT}	3MS.ACC	gives

'he gives it to us.'

d.

[av]=	al=	da
{3MS.NOM, 2P.DAT}	3MS.ACC	gives

'he gives it to you.'

- The duplex analysis explains why [al] occurs preconsonantly instead of the codaless [a], and why [a] appears only before certain clitics and in certain conditions.
- Again, suppletion: duplexes are not morphosyntactically identical to simplex clitics, so they must be separate lexical entries.

4.2 3MS.ACC

- Prevocalic: [l] (11)
- Preconsonantal: [al] (12)

(11) at= | l= | a dɛ
{3MS.NOM, 2S.DAT}= 3MS.ACC= has given
'he gave it to you.'

(12) at= | al= | da
{3MS.NOM, 2S.DAT}= 3MS.ACC= gives
'he gives it to you.'

- Suppletion, for the same reasons given for 3MS.NOM.
- No duplexes for this 3MS.ACC clitic.

4.3 Interaction of 3MS.NOM & 3MS.ACC

- Prevocalic interaction of 3MS.NOM and 3MS.ACC is as expected (13):
  - 3MS.ACC → [l] (prevocalic environment)
  - 3MS.NOM → [al] (preconsonantal environment)

(13) a. al= l= iŋdveŋna
3MS.NOM= 3MS.ACC= guesses
'he guesses it.'

b. al= l= a vest
3MS.NOM= 3MS.ACC= has seen
'he saw him.'

- Preconsonantal interaction is unexpected (14): [e] is epenthetic; [lC] onsets are disallowed—a situation that could have been avoided with 3MS.ACC [al].

(14) a. al= le= vad
3MS.NOM= 3MS.ACC= sees
'he sees him.'

b. al= le= tra
3MS.NOM= 3MS.ACC= throws
'he throws it.'

- A priori expectation: *[l= al= vad]
  - 3MS.ACC → [al] (preconsontal environment)
  - 3MS.NOM → [l] (prevocalic environment)

⇒ GSRs can account for this behavior.

5 Analysis

5.1 3MS.NOM & Duplexes

- All allomorphs appear in the input.
- Activity is assigned to whole allomorphs, not individual segments.

(15) /(0.1·l, 0.8·al)= vad/
3MS.NOM= sees
'he sees'

- Faithfulness favors allomorphs with greater underlying activity.
  - MAX rewards underlying activity preserved in a candidate (roots' activities are ignored in tableaux here).
  - DEP penalizes activity that must be added to bring an element's activity up to 1.

(16)

/(0.1·l, 0.8·al)=vad/	MAX 5	DEP 15	*H*
a. l=vad	0.1	−0.9	−13
☞ b. al=vad	0.8	−0.2	1
c. le=vad	0.1	−1.9	−28

- In this case, *[+son]PERIPH also favors [al=vad]:

(17)

/(0.1·l, 0.8·al)=vad/	*[+son]PERIPH 37	MAX 5	DEP 15	*H*
a. l=vad	−1	0.1	−0.9	−50
☞ b. al=vad		0.8	−0.2	1
c. le=vad		0.1	−1.9	−28

- Low activity is not fatal, in the right circumstances:

(18)

/(0.1·l, 0.8·al)=arspand/	ALIGN-[al]$_{NOM}$ 40	MAX 5	DEP 15	*H*
☞ a. l=arspand		0.1	−0.9	−13
b. a.l=arspand	−1	0.8	−0.2	−39

- High-weighted constraints can favor both a low-activity allomorph and epenthesis over a high-activity allomorph:

(19)

/vad=(0.1·l, 0.8·al)/	ALIGN-[al]$_{NOM}$ 40	*[+SON]PERIPH 37	MAX 5	DEP 15	*H*
a. vad=l		−1	0.1	−0.9	−50
b. va.d=al	−1		0.8	−0.2	−39
☞ c. va.d=el			0.1	−1.9	−28

- Duplexes are preferred over simplexes: generally, they're at least optional whenever the morphosyntactic conditions are met.

- Each duplex has an activity lower than the corresponding simplexes. E.g.:

(20) 3MS.NOM: /(0.1·l, 0.8·al, 0.45·am, 0.45·at, etc.)/

- Normally, they're suboptimal:

(21)

/(0.1·l, 0.8·al, 0.45·at)=vad/	MAX 5	DEP 15	*H*
a. l=vad	0.1	−0.9	−13
☞ b. al=vad	0.8	−0.2	1
c. at=vad	0.45	−0.55	−6

- But if 2S.DAT, e.g., is also in the input, it contributes another /0.45·at/, and candidates with that allomorph combine the activities of the 3MS.NOM /at/ and 2S.DAT /at/.

(22)

/(0.1·l, 0.8·al, 0.45·at)=(0.3·t, 0.45·at)=la=da/	*[+SON]PERIPH 37	MAX 5	DEP 15	*H*
a. l=t=la=da	−1	0.1 + 0.3	−1.6	−59
b. al=t=la=da		0.8 + 0.3	−0.9	−8
☞ c. at=la=da		0.45 + 0.45	−0.1	3

5.2 3MS.ACC & the Puzzling Interaction

- 3MS.ACC: [l] prevocalically (23), [al] preconsonantally (24):

(23)

/…= (0.95·l, 0.7·al)=a da/	*[+SON]PERIPH 37	MAX 5	DEP 15	*H*
☞ a. …=l=a dɛ		0.95	−0.05	4
b. …=al=a dɛ		0.7	−0.3	−1
c. …=le=a dɛ		0.95	−1.05	−11

(24)

/…= (0.95·l, 0.7·al)=da/	*[+SON]PERIPH 37	MAX 5	DEP 15	*H*
a. …=l=da	−1	0.95	−0.05	−33
☞ b. …=al=da		0.7	−0.3	−1
c. …=le=da		0.95	−1.05	−11

- Interaction between 3MS.NOM and 3MS.ACC: the combined preference for 3MS.NOM [al] and 3MS.ACC [l] is great enough to override other considerations:

(25)

/(0.1·l, 0.8·al)=(0.95·l, 0.7·al)=vad/	*[+SON]PERIPH 37	MAX 5	DEP 15	*H*
☞ a. al=le=vad		0.8 + 0.95	−1.25	−10
b. l=al=vad		0.1 + 0.7	−1.2	−14
c. al=l=vad	−1	0.8 + 0.95	−1.25	−47

- GSRs permit an account of Bolognese's suppletion, including the unexpected outcomes and the competition with epenthesis.

6 The Larger Context

- Embedding this analysis in a larger account of Bolognese clitics confirms the results from above.
- Optionality arises in some cases: we adopt Noisy Harmonic Grammar (NHG; Boersma & Pater 2016, Jesney 2007, Hayes 2017), implemented in R (R Core Team 2022).

6.1 Old Data

- 3MS.NOM with no other clitics ((6) & (7)):

(26) l= arspɑnd
3MS.NOM= respond.3S
'he responds'

(27) al= vad
3MS.NOM= see.3S
'he sees'

- Postverbal 3MS.NOM (8):

(28) vad=el 'Does he see?'

- 3MS.NOM duplexes (10); just (10b) included:

(29) at= la= da
{3MS.NOM, 2S.DAT} 3FS.ACC give.3S
'he gives it to you.'

- 3MS.ACC prevocalically and preconsonantally ((11) & (12)):

(30) at= [l]= a dɛ
{3MS.NOM, 2S.DAT}= 3MS.ACC= has given
'he gave it to you.'

(31) at= [al]= da
{3MS.NOM, 2S.DAT}= 3MS.ACC= gives
'he gives it to you.'

- 3MS.NOM with 3MS.ACC ((13)–(14)):

(32) a. al= l= iŋdveŋna
3MS.NOM= 3MS.ACC= guesses
'he guesses it.'

b. al= le= vad
3MS.NOM= 3MS.ACC= sees
'he sees him.'

6.2 New Data

- Duplexes are optional when just one of DAT and ACC clitics is present:

(33) a. [al]= [t]= di:z
3MS.NOM 2S.DAT says
'he says to you.'

b. [at]= di:z
{3MS.NOM, 2S.DAT} says
'he says to you.'

(34) a. [al]= [s]= di:z
3MS.NOM 1P.DAT says
'he says to us'

b. [as]= di:z
{3MS.NOM, 1P.DAT} says
'he says to us.'

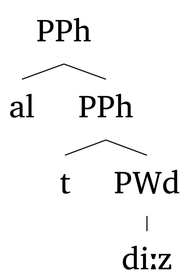
(35) a. [al]= [t]= tsa:ma
3MS.NOM 2S.ACC calls
'he calls you.'

b. [at]= tsa:ma
{3MS.NOM, 2S.ACC} calls
'he calls you.'

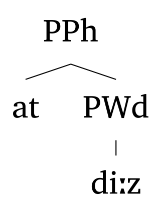
(36) a. [al]= [s]= tsa:ma
3MS.NOM 1P.ACC calls
'he calls us'

b. [as]= tsa:ma
{3MS.NOM, 1P.ACC} calls
'he calls us.'

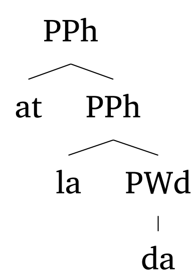
- Our account:
  - Cardinaletti & Repetti (2008): in Donceto (closely related to Bolognese), proclitics are outside the verb's PWd.
  - We implement this by assigning clitics to PPh.
  - Recursive PPhs (Ito & Mester 2007, 2009a,b, 2013): each clitic induces a new one.
  - *DUPLEX-PPh$_{min}$ discourages duplexes in the minimal (= lowest) PPh, competing with MAX and DEP, which favor duplexes (37a), (37b).
  - But when both DAT and ACC are present, the duplex is outside the minimal PPh, and *DUPLEX-PPh$_{min}$ doesn't penalize it (37c).

(37) a.


OK on *DUPLEX-PPh$_{min}$; worse on MAX/DEP

b.


*DUPLEX-PPh$_{min}$ violated; better on MAX/DEP

c.


OK on both *DUPLEX-PPh$_{min}$ and MAX/DEP

- With just one of DAT/ACC and a V-initial verb, duplexes are impossible:

(38) a. al= t= arspɑnd
3MS.NOM 2S.DAT responds
'he responds to you.'
b.* at=arspɑnd

(39) a. al= s= abraθa
3MS.NOM 1P.ACC hugs
'he hugs us.'
b.* as=abraθa

- Our account:
  - ONSET-PWd forces clitics to provide an onset for the verb.
  - CRISPEDGE-PWd (Ito & Mester 1999) prevents morphemes from straddling the PWd boundary.
  - Duplexes must violate one of these constraints; simplexes do not:

(40)

/3MS.NOM, 2S.DAT, arspɑnd/	ONSET-PWd	CRISPEDGE-PWd
☞ a. al=[t=arspɑnd]$_{PWd}$		
b. a[t=arspɑnd]$_{PWd}$		*!
c. [at=arspɑnd]$_{PWd}$	*!	
d. at=[arspɑnd]$_{PWd}$	*!	

- One more constraint: DEP-σ_1
  - Useful in ruling out extraneous alternations for 3MS.NOM (which is always word-initial, except in inversions).
  - DEP-σ_1 is identical to DEP, but it penalizes only initial-syllable epenthesis.

6.3 Noisy Harmonic Grammar

- Constraint weights are perturbed on each evaluation.
- Code written in R (R Core Team 2022), available at https://github.com/afkaplan/Bolognese
- Noise: Gaussian distribution with mean of 0 and standard deviation of 1
- Weights (41) and activities (42) given below:

(41)

Constraint	**Weight**
MAX	5
DEP	15
DEP-σ_1	28
*DUPLEX-PPh$_{min}$	34
*[+son]PERIPHERY	37
ONSET-PWd	55
CRISPEDGE-PWd	55

(42)

Clitic	Allomorph	Activity
3MS.NOM	[l]	0.1
	[al]	0.8
	duplexes	0.45
3MS.ACC	[l]	0.95
	[al]	0.7
2S.DAT	[t]	0.3
	[at]	0.45
2S.NOM	[t]	0.3

- Results (from 10,000 trials for each form):
  - Categorical data: all and only attested forms produced.
  - Optional duplexes (33):
    - al=t=diːz: 64.8%
    - at=diːz: 35.2%

7 Conclusion

- GSRs offer an account of suppletive allomorphy without requiring a suppletion-specific apparatus.
- Bolognese uses both [e]-epenthesis and suppletion to satisfy well-formedness constraints, both of which violate DEP. Nonetheless, each appears just where it should.
- NHG accounts for the system's optionality. A possible avenue for research: perturbed activity rather than perturbed weights.

References

Boersma, Paul & Joe Pater (2016) Convergence Properties of a Gradual Learning Algorithm for Harmonic Grammar. In *Harmonic Grammar and Harmonic Serialism*, John J. McCarthy & Joe Pater, eds., 389–434, Bristol, CT: Equinox.

Canepari, Luciano & Daniele Vitali (1995) Pronuncia e Grafia del Bolognese. *Rivista Italiana di Dialettologia* **19**: 119–164.

Cardinaletti, Anna & Lori Repetti (2008) The Phonology and Syntax of Preverbal and Postverbal Subject Clitics in Northern Italian Dialects. *LI* **39**(4): 523–563.

Clements, G.N. (1991) Place of Articulation in Consonants and Vowels: A Unified Theory. *Working Papers of the Cornell Phonetics Laboratory* **5**: 77–123.

Faust, Noam & Paul Smolensky (2017) Activity as an Alternative to Autosegmental Association. Paper presented at the 25th Manchestser Phononology Meeting.

Hayes, Bruce (2017) Varieties of Noisy HG. In *Proceedings of AMP 2016*, Karen Jesney, Charlie O'Hara, Caitlin Smith, & Rachel Walker, eds., Washington, DC: Linguistic Society of America.

Ito, Junko & Armin Mester (1999) Realignment. In *The Prosody-Morphology Interface*, René Kager, Harry van der Hulst, & Wim Zonneveld, eds., 188–217, Cambridge, U.K.: Cambridge University Press.

Ito, Junko & Armin Mester (2007) Prosodic Adjunction in Japanese Compounds. In *MIT Working Papers in Linguistics 55: Formal Approaches to Japanese Linguistics 4*, Yoichi Miyamoto & Masao Ochi, eds., 97–111, Cambridge, MA: Dept. of Linguistics and Philosophy, MIT.

Ito, Junko & Armin Mester (2009a) The Extended Prosodic Word. In *Phonological Domains: Universals and Derivations*, Baris Kabak & Jaent Grijzenhout, eds., 135–194, The Hague: Mouton de Gruyter.

Ito, Junko & Armin Mester (2009b) The Onset of the Prosodic Word. In *Phonological Argumentation*, Steve Parker, ed., 227–260, London: Equinox.

Ito, Junko & Armin Mester (2013) Prosodic Subcategories in Japanese. *Lingua* **124**: 20–40.

Jesney, Karen (2007) The Locus of Variation in Weighted Constraint Grammars. Poster presented at the Workshop on Variation, Gradience and Frequency in Phonology. Stanford, CA: Stanford U. July 2007.

R Core Team (2022) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, URL `https://www.R-project.org/`.

Rubin, Edward & Aaron Kaplan (2022) Lexical Selection in Bolognese Clitic Allomorphy. *Isogloss. Open Journal of Romance Linguistics* **8**(5): 1–17.

Rubin, Edward J. & Aaron Kaplan (to appear) Segmental and Prosodic Influences on Bolognese Epenthesis. In *Selected Papers from Stony Brook Epenthesis Workshop*, Lori Repetti, Ji Yea Kim, Veronica Miatto, & Andrija Petrovic, eds., LingSci Press.

Selkirk, Elisabeth (1984) On the Major Class Features and Syllable Theory. In *Language Sound Structure: Studies in Phonology*, Mark Aronoff & Richard T. Oehrle, eds., 107–136, Cambridge, MA: MIT Press.

Smolensky, Paul & Matthew Goldrick (2016) Gradient Symbolic Representations in Grammar: The Case of French Liaison. ROA-1286, Rutgers Optimality Archive, http://roa.rutgers.edu.

Vitali, Daniele (2009) *Dscårret in bulgnais?* Bologna: Perdisa, 2 ed.

Zimmermann, Eva (2019) Gradient Symbolic Representations and the Typology of Ghost Segments. In *Proceedings of the 2018 Annual Meeting on Phonology*, Katherine Hout, Anna Mai, Adam McCollum, Sharon Rose, & Matthew Zaslansky, eds., Linguistic Society of America.