

Explotació de corpus orals: l'experiència del COD (Corpus Oral Dialectal del català)

Maria-Rosa Lloret

Departament de Filologia Catalana

Universitat de Barcelona

mrosa.lloret@ub.edu



Resultats del projecte ECOD

Explotació d'un Corpus Oral Dialectal: anàlisi de la variació lingüística i desenvolupament d'aplicacions informàtiques

HUM2004-1504/FILO (1^a fase), 2004-07

HUM2007-65531/FILO (2^a fase), 2007-10

Pàg. web del projecte: <http://www.ub.edu/lincat>

EQUIP INVESTIGADOR

Universitat de Barcelona (9)

- **Maria-Rosa Lloret** (IP)
- **Joaquim Viaplana**
- **Maria-Pilar Perea**
- Montserrat Adam
- Sebastià Bonet
- Andreu Bosch
- Josefina Carrera
- Josep Moran
- Clàudia Pons

Universitat Pompeu Fabra (1)

- **Esteve Clua**

Universitat de València (1)

- Jesús Jiménez

Universitat Catòlica de València (1)

- Ricard Herrero (HUM2007)

Euskal Herriko Unibertsitatea (1)

- Gotzon Aurrekoetxea (HUM2004)

Universität Salzburg (2)

- Xavier Casassas (HUM2007)
- Hans Goebel

Universität Basel (1)

- Germà Colon

Meertens Instituut (Amsterdam) (1)

- Franz Hinskens

Becaris predoctorals (U Barcelona) (9)

- Elisenda Campmany (2003-06; dir.: Lloret)
- Laia Querol (2003-06; dir.: Lloret)
- Laia Rosàs (2004-07; dir.:Perea)
- Núria Busquet (2005-08; dir.:Perea)
- Joan-Pau Callejón (2005-08; dir.: Bonet)
- Roser Marí (2006-09; dir.: Perea)
- Maria Cabrera (2007-10; dir: Carrera & Pons)
- Esteve Valls (2007-10; dir.: Clua & Viaplana)
- Margalida Adrover (2008-11; dir.: Lloret)

OBJECTIUS

1. Completar i informatitzar les **dades** del *Corpus Oral Dialectal* (COD) del català contemporani, recopilat en projectes anteriors. (J. Viaplana, amb M-R. Lloret, M-P. Perea i E. Clua)
2. Anàlisi dels **fenòmens lingüístics** amb relació a la variació dialectal. (M-R. Lloret)
3. Anàlisi **dialectomètrica** i representació de la **distància lingüística**. (E. Clua)
4. Elaboració d'un **atles sonor** foneticomorfològic. (M-P. Perea)

D'on provenen les dades del COD?

- **Projectes anteriors**, dirigits per Joaquim Viaplana:
 - PB94-0951 (1994-1997)
 - PB97-0889 (1998-2001)
 - BFF2001-3798 (2002-2004)
- **Enquestes**, bàsicament entre 1994-96:
 - 82 localitats (caps de comarca o similars)
 - 303 entrevistes (mínim 3 per localitat)
- **Qüestionari foneticomorfològic** (aprox. 600 ítems)
Text lliure (10')
- **Informants** 30-45 anys, no escolaritzats en català;
nivell cultural mitjà



COD

82 localitats

303 entrevistes

Qüestionari

1. COMPLETAR

Ex.: VERBS REGULARS: conjugació 1 (*cantar*):

- Cada dia al matí jo una cançó. [canto]
- Mentre ens esperàvem, jo una cançó. [cantava]

2. IDENTIFICAR

Ex.: FONÈTICA

- [Imatge] (perla)

3. TRADUIR del castellà, francès o italià

Ex.: COMBINACIÓ DE PRONOMS PERSONALS

- Si tus hijos quieren ese coche, *se lo* compras. [els el / els hi compres]

Metodologia

- Enregistrament digital.

TRACTAMENT DADES QÜESTIONARI:

- Transcripció ortogràfica i fonètica de les dades segmentades morfològicament. Base de dades dissenyades amb Microsoft® Access, **específica** per a cada base de dades.
- Supervisió de les dades transcrites.
- Exportació de les dades a base de dades (© ORACLE) per poder **relacionar** cada registre amb registre sonor.
- Selecció automàtica de la forma més **freqüent** ($3 > 1$).

Exemple base de dades creada per a verbs

Fonètica	Arrel	Ep	Ext	Tema	Mode	MT	Ep2	NP
kəntəríə	kənt			ə		ríə		
pərðərém	pərð	ə				ré		m
bəyént	bə		ɣ	é	nt			
bəwɾə	bəw				r		ə	
sénten	sént						e	n
serβiʃko	serβ		íʃk			o		

Base de dades (fonètica)

The screenshot shows the Microsoft Access application window titled "Microsoft Access - [Consulta_completa]". The menu bar includes "Archivo", "Edición", "Ver", "Insertar", "Formato", "Registros", "Herramientas", and "Ventana". The toolbar contains various icons for file operations, editing, and navigation. The main window displays a form with the following fields and values:

INFORM	MDB
LOCALITAT	Reus
ID_FON	0001
MOT_ANT	
MOT_ANT_FON	
ARREL	káz
VOC_INTER	<input type="checkbox"/>
NASAL	<input type="checkbox"/>
DIMINUT	
VOC_FINAL	e
PLURAL	<input type="checkbox"/>
seqüència	
sencer	
FORMES ALT.	
OBSERVACIONS	
OBSERVACIO_Fon	

At the bottom of the form, there is a record navigation section showing "Registro: 1 de 55283" and a "Vista Formulario" button. The Windows taskbar at the bottom shows the system tray with the time "8:32" and several open applications including "Espai Person...", "Elisenda C", "Microsoft Wo...", "BASES DE D...", "Microsoft ...", and "Microsoft Ac...".

Base de dades (clítics pronominals)

Microsoft Access - [Total]

Archivo Edición Ver Insertar Formato Registros Herramientas Ventana ?

Clítics

INFOR	DFM	PLUR2		VO_FE1	
LOC	La Bisbal	V_FIPOST2		PL1	
ID_PF	0001	VOC_VERB		VO_FST1	
MOTS_ANT	Cada dia	CONS_VERB	rèntu	VO_I2	
MOTS_ANT	káðe ði	VERB_VOC		P_2	
V_I1	a	VERB_CONS		VO_FE2	
PR01	m	VERB_W		PL2	
V_FIPRE1		VERB_R		VO_ST2	
PLUR1		VERB_NT		MOTS_POST	la cara
V_FIPOST1		VO_I1		MOTS_POST	le káre
V_I2		P_1		SENCER	am
PR02				Observacions	
V_FIPRE2					

Registro: 1 de 24376

Vista Formulario

Inicio Espai Person... Elisenda C Microsoft Wo... BASES DE D... Microsoft Ac... Microsoft ... 8:37

Base de dades (articles)

Microsoft Access - [total]

Archivo Edición Ver Insertar Formato Registros Herramientas Ventana ?

Articles

ORT	el gos / cap	MOT_ANT	
SING	1	CONS_ANT	
MASC	1	VOC_ANT	
FEM	0	ART	l
PL	0	VOC_FINAL	u
INFORM	RIMA	PLURAL	
LOCAL	L'Alguer	MOT_POST	káp
ID_ART	0001	SENCER	lu

Registro: 1 de 4871

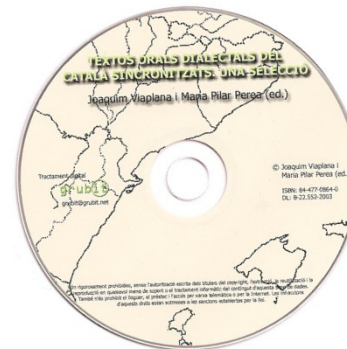
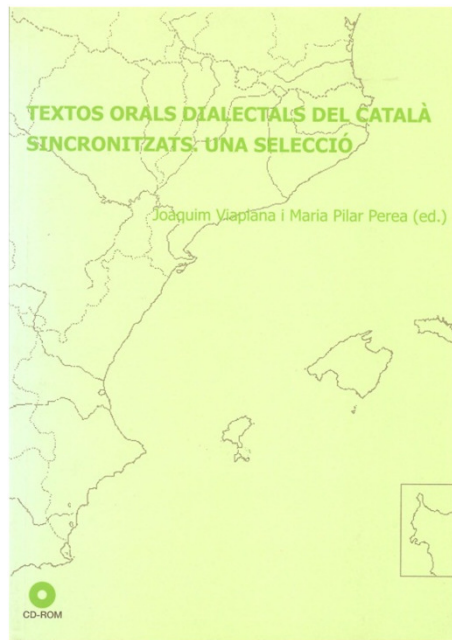
Vista Formulario

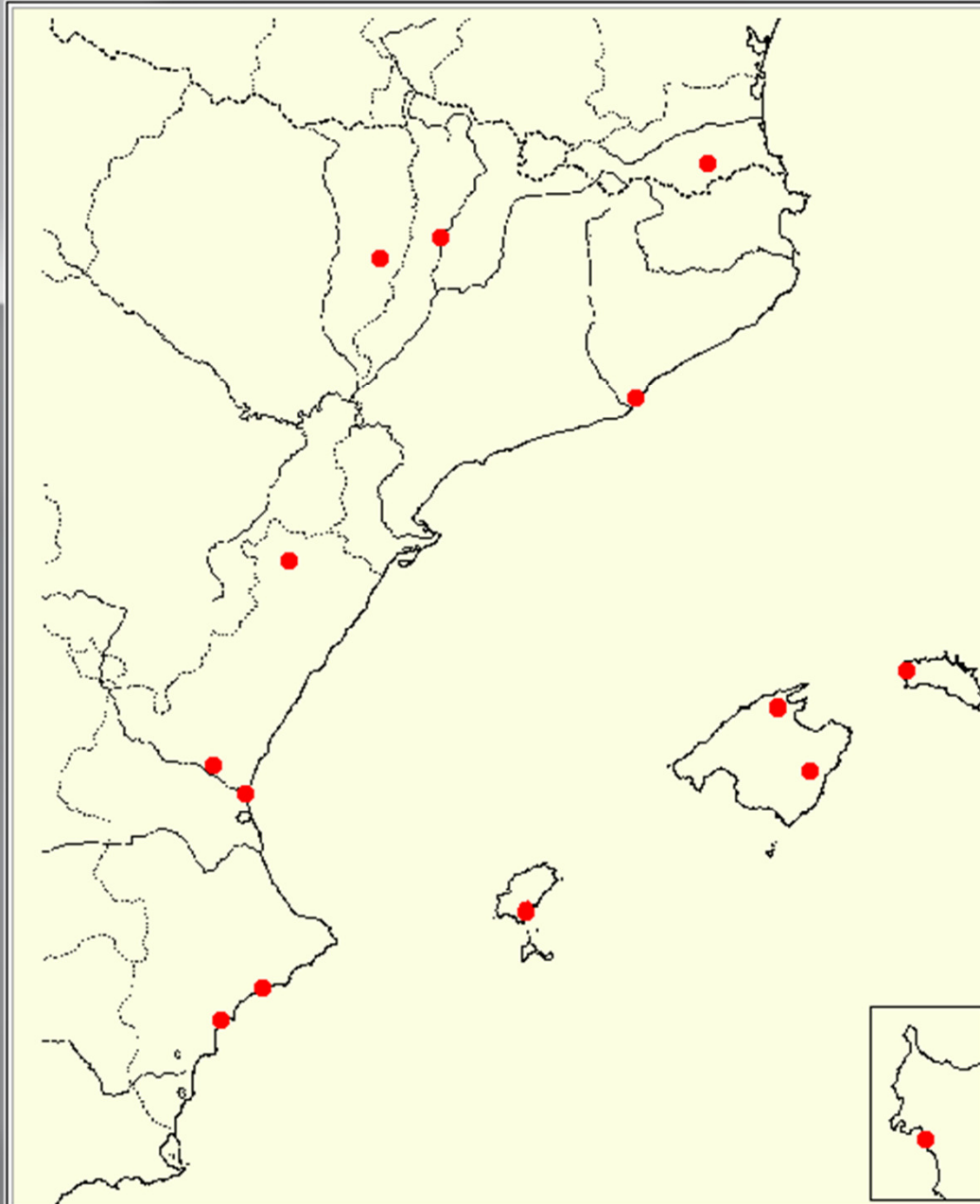
Inicio Espai Person... Elisenda C Microsoft Wo... BASES DE D... Microsoft ... 8:43

TRACTAMENT DADES TEXTOS LLIURES:

- En principi, per comprovar resultats enquestes.
- Aplicació docent: Transcripció ortogràfica i foneticoortogràfica de **13 textos de diferents dialectes** publicats en versió sincronitzada, en paper i en CD-ROM:

Viaplana & Perea (ed.) (2003): *Textos orals dialectals del català sincronitzats. Una selecció*. Barcelona: PPU.





Els punts corresponen a 14 localitats del domini lingüístic català, en les quals s'han associat els textos sincronitzats. Per començar, feu clic sobre els punts marcats en el mapa.

Presentació



Sincronitzador



Informació

Població: L'Alguer

Dialecte: Algerès
Subdialecte:
Codi informant: RIMA
Durada: 0:09:11

T. Fonètica T. Fonoortogràfica

 **Reproduir**

Els punts corresponen a 14 localitats del domini lingüístic català, en les quals s'han associat els textos sincronitzats. Per començar, feu clic sobre els punts marcats en el mapa.

Presentació 

**Transcripció fonètica: L'Alguer**

1 'ez um be'ħisiħ 'ħok (.) e: 'tens de 'na a 'vera 'tota las kus'teraz
 2 ðe r al'ye (.) de 'na (.) a de 'ka de 'kasa (.) a l aspa'ransa a'pòlina
 3 'sez 'ma e'nar a 'pòlina (.) 'tens i de 'vera: la: lu la lukali'tat i 'k ez
 4 a'pòlina ðe 'rokas (.) 'ez e: (.) uħ 'ħok a'saj pejza'džistik (.) i 'tu
 5 'tens i ðe 'nal 'vewra (.)

6 E: I tant.

7 e 'j a wna jyre'zjeta 'mentre 'tu 'vaz e ke s a'kira l aspa'ransa (.)
 8 'anti a lu 'vujt i ðe se'tembra sa 'fa la: (.) 'j ez e: la preyaru'ria
 9 ðe 'tu rus kampa'poruz (.) 'j ez la 'misa (.) i: a la 'siz de 'taħda la
 10 'diw tam'be wħ kapa'ħa (.) la 'diw aħ kata'ra (.)

11 E: En català.

12 ner 'se (.) e: (.) 'i e'z una 'koza (.) 'ez un ra'kħħt as'torio de 'tota

**Transcripció fonoortogràfica: L'Alguer**

1 És un bellíssim lloc. E... tens d'anar a vere totes les costeres
 2 de l'Alguer, d'anar a... de ca- de Caça a l'Esperança, Pòllina.
 3 Ets mai anada a Pòllina? Tens de vera la lo- la localitat que és
 4 a Pòllina de Roques, és e... un lloc assai paisatgístic i tu
 5 tens d'anà-l veure.

6 E: I tant.

7 E... hi ha una iglesieta mentre tu vas, e... que s'aquidra l'Esperança.
 8 *Anti* a lo vuit de setembre se fa la..., hi és... la pregadoria
 9 de tots los campagnolos, hi és la missa, i... a les sis de tarda la
 10 diu també un capellà, la diu en català.

11 E: En català.

12 Per se- e... i és una cosa... és un record històric de totes

Exercicis

Digueu a quin dialecte o subdialecte pertany cadascuna de les paraules o de les seqüències de paraules següents.
De vegades, més d'una resposta és correcta.


1. any nou [an 'nou]
 balear
 alguerès
 rossellonès

2. cantava [kan'tava]
 mallorquí
 menorquí
 alguerès

3. pocs [ˈpots]
 mallorquí
 alguerès
 menorquí

4. apagar [əpə'ja]
 mallorquí
 menorquí
 alguerès

5. tendra [ˈtendra]

 **Recompte** **Netejar solucions**
Solucions **Netejar marcats**

Exercicis

Digueu a quin dialecte o subdialecte pertany cadascuna de les paraules o de les seqüències de paraules següents.
De vegades, més d'una resposta és correcta.


1. any nou [an 'nou]
 balear
 alguerès
 rossellonès

2. cantava [kan'tava]
 mallorquí
 menorquí
 alguerès

3. pots [ˈpots]
 mallorquí
 alguerès
 menorquí

4. apagar [əpe'ja]
 mallorquí
 menorquí
 alguerès

5. tendra [ˈtendra]

 **Recompte**
Solucions

Netejar solucions
Netejar marcats

Encerts: 3
Errades: 1
En blanc: 90

Exercicis

Digueu a quin dialecte o subdialecte pertany cadascuna de les paraules o de les seqüències de paraules següents. De vegades, més d'una resposta és correcta.

1. any nou [an 'nou] 💡 ✗ alguerès

balear
 alguerès
 rossellonès

2. cantava [kan'tava] 💡 ✔

mallorquí
 menorquí
 alguerès


3. pocs ['pots] 💡 ✔

mallorquí
 alguerès
 menorquí

4. apagar [əpe'ja] 💡 ✔

mallorquí
 menorquí
 alguerès

5. tendra ['tɛnɾə]

 Recompte Solucions Netejar solucions Netejar marcats Encerts: 3 Errades: 1 En blanc: 90

EXPLOTACIÓ DEL COD

(HUM2004, HUM2007)

1. **Difusió** de les dades del COD (accés on-line)

- a) Dades del qüestionari
- b) Textos orals

2. **Anàlisi** de les dades:

- a) Anàlisi **dialectomètrica** i representació de la distància lingüística
- b) Anàlisi de **fenòmens lingüístics** amb relació a la variació dialectal

1. Difusió de les dades del COD

Coord.: Viaplana, amb Lloret, Perea i Clua

a) Bases de dades (qüestionari)

1ª fase: Publicació de les bases de dades en CD-Rom (2007)



El Corpus Oral Dialectal (COD) del català contemporani conté informació dels sis principals dialectes del català, obtinguda d'entrevistes efectuades als caps de comarca -o equivalents- del domini lingüístic català entre 1994 i 1996, amb informants d'entre 30 i 45 anys. Aquest CD-ROM aplega els resultats aconseguits a partir del qüestionari en vuit bases de dades (© Microsoft Access), cadascuna de les quals té una estructura adaptada al contingut: aspectes fonètics rellevants, morfologia verbal regular, clítics pronominals, articles, possessius, pronoms personals forts, demostratius i locatius. Apte per a usos docents, és també una bona eina per a la recerca dialectal i per a l'estudi de la situació actual i de les perspectives futures de les varietats dialectals del català.

The Corpus Oral Dialectal (COD) of contemporary Catalan contains information on the six main dialects of Catalan, obtained through a series of interviews carried out in the main county (comarca) capitals -or equivalent towns- of the Catalan linguistic domain between 1994 and 1996, with informants aged between 30-45 years old. This CD-ROM gathers the results drawn from the questionnaire in eight databases (© Microsoft Access), each of which has a structure adapted to its contents: relevant phonetic aspects, regular verbal morphology, pronominal clitics, articles, possessives, strong personal pronouns, demonstratives and locatives. It can be used for teaching, and it can also be a useful tool in dialect research as well as in the study of the current state and future perspectives of the Catalan dialects.

C O D
Corpus Oral Dialectal
Departament de Filologia Catalana
Universitat de Barcelona

Joaquim Viaplana (UB)
Maria-Rosa Lloret (UB)
Maria-Pilar Perea (UB)
Esteve Clua (UPF)

 PPU Departament de Filologia Catalana
Universitat de Barcelona

Joaquim Viaplana (UB)
Maria-Rosa Lloret (UB)
Maria-Pilar Perea (UB)
Esteve Clua (UPF)

Dipòsit Legal: B-47506-2007
ISBN: 978-84-477-0990-8

Tractament digital **grubit**

 PPU

Corpus Oral Dialectal

Arxiu Informació Quant a ...

Articles Clítics Demonstratius Fonètica Locatius Possessius Pronoms personals Verbs Informació

Condicions Resultat de la consulta

Cerca


Condicions per a articles

Articles	Localitats	Informants
al als del dels el els en la	Alacant Albocàsser Alcoi Alzira Amposta Andorra Balaguer Banyoles	AAC AAG AAS ACB ACF ACM ACT ADCA

Gènere: Tots Nombre: Tots Personal

CORPUS ORAL DIALECTAL (COD)

Departament de Filologia Catalana
Universitat de Barcelona



Joaquim Viaplana (UB)
Maria-Rosa Lloret (UB)
Maria-Pilar Perea (UB)
Esteve Clua (UPF)

Dipòsit legal: B-47506-2007
ISBN: 978-84-477-0990-8

PPU

Inicio Symposia Adobe Reader Correu electrònic :: S... Correu electrònic :: R... Corpus Oral Dialectal Documento1 - Microso... ES 14:40

Corpus Oral Dialectal

Arxiu Informació Quant a ...

Articles Clítics Demonstratius Fonètica Locatius Possessius Pronoms personals **Verbs** Informació

Condicions Resultat de la consulta

Cerca

Condicions per a verbs

Verbs	Localitats	Informants
<ul style="list-style-type: none"> cantar canviar conèixer coure dir dormir eixir moure 	<ul style="list-style-type: none"> Albocàsser Alcoi Alzira Andorra Balaguer Benavarri Berga Borriana 	<ul style="list-style-type: none"> AAC AAG AAS ACB ACF ACM ACT ADCA

beure


Dialectes	Temps	Núm. persona
<ul style="list-style-type: none"> Alguerès Balear Central Nord_occ 	<ul style="list-style-type: none"> Part Perfet PS 	<ul style="list-style-type: none"> 2 3 4 5

PI

1

CORPUS ORAL DIALECTAL (COD)

Departament de Filologia Catalana
Universitat de Barcelona



PPU

Joaquim Viaplana (UB)
 Maria-Rosa Lloret (UB)
 Maria-Pilar Perea (UB)
 Esteve Clua (UPF)

Dipòsit legal: B-47506-2007
 ISBN: 978-84-477-0990-8

Windows taskbar: Inicio, Symposia, Adobe Reader, Correu electrònic :..., Correu electrònic :..., Corpus Oral Dial..., Documento1 - Mic..., Microsoft Excel, ES, 14:47

Condicions Resultat de la consulta

Exporta Imprimeix

Que la consulta inclogui les paraules següents:

beure

Que la consulta inclogui les localitats següents:

Alacant

Amposta

Banyoles

Barcelona

Que la consulta inclogui els informants següents:

Localitat	Inf.	Conj.	Dialecte	Ort. 1	Ort. 2	Temps	Persona	Fon.	Arrel	Ep.	Ext.
Alacant	BPM	2b	Valencià	bec	beure	PI	1	bék	bé		k
Alacant	CBS	2b	Valencià	bec	beure	PI	1	bék	bé		k
Alacant	FJCP	2b	Valencià	bec	beure	PI	1	bék	bé		k
Amposta	JPM	2b	Nord_occ	bec	beure	PI	1	bék	bé		k
Amposta	CFR	2b	Nord_occ	bec	beure	PI	1	bék	bé		k
Amposta	ARG	2b	Nord_occ	bec	beure	PI	1	bék	bé		k
Banyoles	JST	2b	Central	bec	beure	PI	1	bék	bé		k
Banyoles	PGG	2b	Central	bec	beure	PI	1	bék	bé		k
Banyoles	CBF	2b	Central	bec	beure	PI	1	bék	bé		k
▶ Barcelona	CSJ	2b	Central	bec	beure	PI	1	bék	bé		k
Barcelona	CVB	2b	Central	bec	beure	PI	1	bék	bé		k
Barcelona	DAC	2b	Central	bec	beure	PI	1	bék	bé		k
Barcelona	ECC	2b	Central	bec	beure	PI	1	bék	bé		k
Barcelona	FFS	2b	Central	bec	beure	PI	1	bék	bé		k
Barcelona	MAHC	2b	Central	bec	beure	PI	1	bék	bé		k
Barcelona	MIBF	2b	Central	bec	beure	PI	1	bék	bé		k
Barcelona	MFI	2b	Central	bec	beure	PI	1	hék	hé		k

Verbs

Localitat	Inf.	Dialecte	Conj.	Temps	Verb	Pers.	Fonètica	Arrel	Ep	Ext	Tema	Mode	I
Barcelona	CSJ	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	CVB	Central	2b	PI	beure	1	bék	bé		k			
Banyoles	PGG	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	ECC	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	FFS	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	MAHC	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	MIBF	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	MFI	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	PAC	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	MPT	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	TEG	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	GPC	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	CFB	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	JBT	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	JFM	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	MAA	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	MABA	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	MOBA	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	MBM	Central	2b	PI	beure	1	bék	bé		k			
Banyoles	CBF	Central	2b	PI	beure	1	bék	bé		k			
Amposta	JPM	Nord_occ	2b	PI	beure	1	bék	bé		k			
Amposta	CFR	Nord_occ	2b	PI	beure	1	bék	bé		k			
Alacant	CBS	Valencià	2b	PI	beure	1	bék	bé		k			
Alacant	FJCP	Valencià	2b	PI	beure	1	bék	bé		k			
Barcelona	AMT	Central	2b	PI	beure	1	bék	bé		k			
Barcelona	DAC	Central	2b	PI	beure	1	bék	bé		k			
Banyoles	JST	Central	2b	PI	beure	1	bék	bé		k			
Amposta	ARG	Nord_occ	2b	PI	beure	1	bék	bé		k			
Alacant	BPM	Valencià	2b	PI	beure	1	bék	bé		k			

2ª fase: Mapes interactius on-line (coord. Perea, 2010)

Mapes de Variació Dialectal - Mozilla Firefox

http://www.grubit.net/mapadialectal/

Verbs: servir III B present indicatiu 1a p. sing.

Só	Verb	F. estàndard	Conjugació	Temps	Persona	F. fonètica
	servir	serveixo	III B	Present indicatiu	1a p. sing	serβisko
	servir	serveixo	III B	Present indicatiu	1a p. sing	sərβéfu
	servir	serveixo	III B	Present indicatiu	1a p. sing	sərβéfu
	servir	serveixo	III B	Present indicatiu	1a p. sing	sərβéfu
	servir	serveixo	III B	Present indicatiu	1a p. sing	serβifo

Pàgina 1 de 15 1 - 6 de 88

Domini lingüístic català

Map showing dialectal variations of the verb 'servir' in the present indicative first person singular across the Catalan linguistic domain. The map is divided into regions with different phonetic realizations:

- serβisko (purple)
- sərβéfu (red)
- serβéfu (blue)
- serβéfo (green)

Locations marked on the map include: 11 - Sort, 20 - Salagosa, 31 - Tremp, 38 - Benavard, 48 - Lanarit de l'Itera, 7 - Balaguer, 26 - Lleida, 2 - Solsona, 10 - Berga, 13 - Cervera, 52 - Manresa, 50 - Igualada, 51 - Granollers, 1 - Ripoll, 14 - Olot, 21 - Vic, 9 - Figueres, and 53 - Girona.

Terminado

5 Firefox 2 Microsoft Offi... Bandeja de entr... 6 Explorador d... cod.ppt vitoria.ppt 16:48

- Versió on-line de les bases de dades del COD publicades abans en CD-Rom.
- Mapes automàtics i associació de la seqüència sonora.

Tasques:

1. Transcripcions:
 - a) ortogràfica
 - b) fonètica
2. Tall de la seqüència sonora.
3. Associació de la seqüència sonora amb les transcripcions corresponents.

www.ub.edu/lincat

C O D

Corpus Oral Dialectal *Mapes de variació dialectal*

Maria Pilar Perea (coordinació)
Departament de Filologia Catalana
Universitat de Barcelona

Desenvolupat per **grubit**



El Corpus Oral Dialectal (COD) del català contemporani conté informació dels sis principals dialectes del català, obtinguda d'entrevistes efectuades als caps de comarca -o equivalents- del domini lingüístic català entre 1994 i 1996, amb informants d'entre 30 i 45 anys. La web representa els resultats aconseguits a partir del qüestionari en vuit bases de dades, cadascuna de les quals té una estructura adaptada al contingut: aspectes fonètics rellevants, morfologia verbal regular, clítics pronominals, articles, possessius, pronoms personals forts, demostratius i locatius.

[Entra a l'aplicació](#)

Finançat pels projectes "Explotació d'un corpus oral dialectal" [HUM2004-01504/FILO] i "Explotació d'un corpus oral dialectal: anàlisi de la variació lingüística i desenvolupament d'aplicacions informàtiques per a la transcripció automatitzada (ECOD)" [HUM2007 65531/FILO].

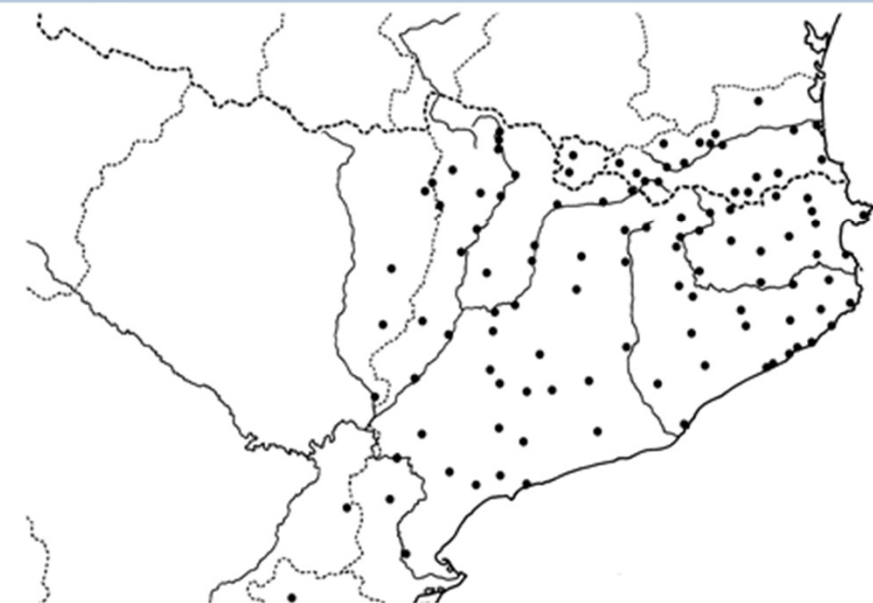
Mapes de variació dialectal



- Artides
- Clitics
- Demonstratius
- Locatius
- Possessius
- Pronoms
- Verbs



[a]



Mapes de variació dialectal

Verbs **beure** IIB present indicatiu 1a p. sing.

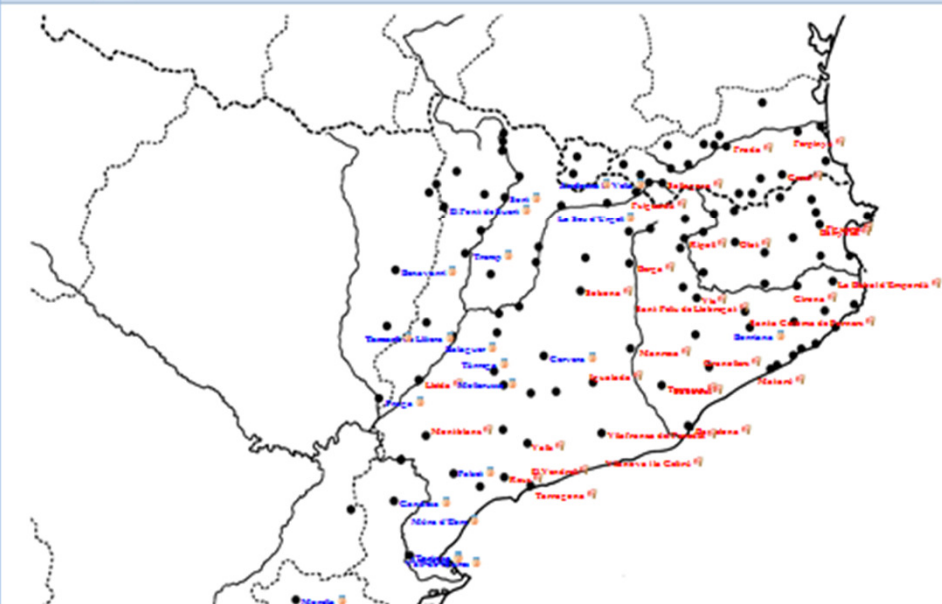
So	Verb	F. estàndard	Conjugació	Temps	Persona	F. fonètica
	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ
	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ
	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ
	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ
	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ

111

Pàgina 1 de 15

1 - 6

(a)



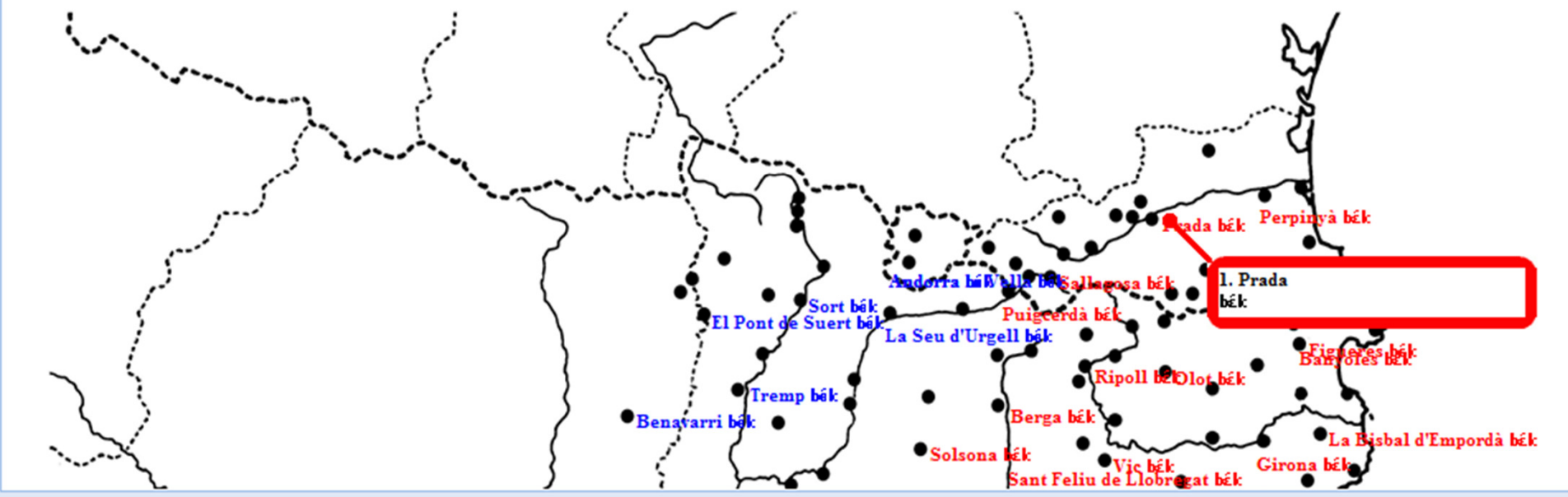
Mapes de variació dialectal

Verbs beure IIB present indicatiu 1a p. sing.

So	Verb	F. estàndard	Conjugació	Temps	Persona	F. fonètica
🔊	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ
🔊	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ
🔊	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ
🔊	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ
🔊	beure	bec	IIB	Present indicatiu	1a p. sing	béɰ

Pàgina 1 de 15

[a]



b) Textos orals espontanis

Accessibilitat on-line de diferents col·leccions de textos: Transcrits en versió fonètica i fonootogràfica i arxiu sonor corresponent (MP3):

Dipòsit digital de la Universitat de Barcelona (<http://diposit.ub.edu/dspace/handle/2445/10413>), i també des de la p. web del projecte (<http://www.ub.edu/lincat>) i la del “Corpus de Català Contemporani de la Universitat de Barcelona (CCUB)” (<http://www.ub.edu/ccub/>):

- **2008-09: 13 textos de diferents dialectes** publicats com a Viaplana & Perea, ed. (2003) en paper i CD-Rom.
- **2009: 12 textos del balear** (Pons & Viaplana, ed.).
- **2010: 14 textos de distintos dialectos** (Perea & Viaplana, ed.).
- **2010: 8 textos del noroccidental** (Carrera & Viaplana, ed.).

<http://www.ub.edu/cccub>

Corpus de Català Contemporani de la Universitat de Barcelona (CCCUB)

El CCCUB és un arxiu de corpus de llengua catalana oral contemporània que ha estat confegit per a l'estudi de la variació dialectal, social i funcional en la llengua catalana.

Aquesta pàgina posa a disposició dels investigadors la transcripció i una part dels arxius sonors dels textos del CCCUB, que han estat dipositats al [Dipòsit Digital de la Universitat de Barcelona](#).

Alguns documents ja havien estat publicats amb anterioritat en llibre i CD-Rom. En aquests casos, els CD-Rom corresponents contenen materials addicionals que no han estat dipositats al Dipòsit Digital de la UB: bases de dades, arxius sonors i la descripció detallada dels materials.

La publicació a internet dels materials del CCCUB és un procés obert: alguns dels documents publicats amb anterioritat en CD-Rom no han completat encara el procés de publicació a la xarxa; a més, es preveu afegir nous materials inèdits al CCCUB. Consulteu les novetats.

NOVETATS (31.01.2011):

- [Textos orals del nord-occidental \(I. Carrera i J. Viaplana\): Móra d'Ebre, Sort, Tamarit de Llitera, Tàrraga.](#)

- [Corpus Oral Dialectal \(COD\)](#)
- [Corpus de Varietats Socials \(COS\)](#)
- [Corpus Oral de Conversa Col·loquial \(COC\)](#)
- [Corpus Oral de Registres \(COR\)](#)

Corpus Oral Dialectal ([COD](#))

[Pàgina principal del CCCUB](#)

Viaplana, Joaquim i Maria Pilar Perea (ed.) (2003): [Textos orals dialectals del català sincronitzats. Una selecció](#). Barcelona: [Promociones y Publicaciones Universitarias](#) (PPU). [Llibre i CD-ROM]
ISBN: 84-477-0864-0

Viaplana, Joaquim; Maria-Rosa Lloret, Maria-Pilar Perea i Esteve Clua (2007): [COD. Corpus Oral Dialectal](#). Barcelona: [Promociones y Publicaciones Universitarias](#) (PPU). [CD-ROM]
ISBN: 978-84-477-0990-8

Pons, Clàudia i Joaquim Viaplana (ed.) (2009): [Corpus oral dialectal \(COD\). Textos orals del balear](#). Dipòsit Digital de la UB.

Perea, Maria-Pilar i Joaquim Viaplana. [Corpus Oral Dialectal \(COD\). Selecció de textos](#). Dipòsit Digital de la UB [Arxiu obert]

Carrera-Sabaté, Josefina i Joaquim Viaplana (ed.): [Corpus Oral Dialectal \(COD\). Textos orals del nord-occidental](#). Dipòsit Digital de la UB.

Copyright: [Grup d'Estudi de la Variació](#), [Departament de Filologia Catalana](#), [Universitat de Barcelona](#), 2008

COD, Textos orals del balear (2009)

[Pàgina principal del CCCUB](#)

[Pàgina principal del COD](#)

- [Ciutadella](#)
- [Eivissa 1](#)
- [Eivissa 2](#)
- [Felanitx](#)
- [Formentera 1](#)
- [Formentera 2](#)
- [Manacor](#)
- [Maó](#)
- [Palma](#)
- [Pollença 1](#)
- [Pollença 2](#)
- [Sóller](#)

Pons, Clàudia i Joaquim Viaplana
(ed.) (2009): *Corpus oral dialectal
(COD). Textos orals del balear.*
Dipòsit Digital de la UB.

Copyright: [Grup d'Estudi de la Variació](#), [Departament de Filologia Catalana](#), [Universitat de Barcelona](#), 2008

Buscar en el Repositorio digital

Búsqueda avanzada

- [Página de inicio](#)
- [Qué es?](#)
- [Cómo publicar en el Repositorio](#)
- [Política de Acceso Abierto UB](#)
- [Estadísticas](#)

Listar

- [Comunidades y colecciones](#)
- [Fecha Publicación](#)
- [Autor](#)
- [Título](#)
- [Materia](#)

Servicios

- [Alertas](#)
- [Mi repositorio](#)
Usuarios autorizados
- [Editar perfil](#)
- [Ayuda](#)
- [Aviso legal](#)

Dipòsit Digital de la UB > Recerca > Grup d'Estudi de la Variació (GEV) - Corpus de Català Contemporani de la Universitat de Barcelona (CCCUB) > Co

Por favor, use este identificador para citar o enlazar este ítem:

Título : Corpus Oral Dialectal (COD). Ciutadella1

Autor : [Pons, Clàudia](#)
[Viaplana, Joaquim, 1942-](#)

Matèria: Linguistics
Spoken Catalan
Dialectology
Balearic
Menorquí language
Ciutadella (Balearic Islands)
Lingüística
Català parlat
Dialectologia
Balear
Menorquí
Ciutadella (Balears)

Fecha de publicación : 29-sep-2009

Resumen : Aquest document conté la transcripció fonètica, la fonoortogràfica i l'arxiu de so d'un fragment de conversa lliu del Corpus de Català Contemporani de la Universitat de Barcelona (CCCUB), un arxiu de corpus de llengua catalana amb finalitat de contribuir a l'estudi de la variació dialectal, social i funcional en la llengua catalana. Una selecció de www.recercat.net), i també és accessible a través del web del CCCUB: <http://www.ub.edu/ccub>.

URI : <http://hdl.handle.net/2445/11627>

Aparece en las colecciones: [Corpus de Català Contemporani de la Universitat de Barcelona \(CCCUB\)](#)
[Corpus Oral Dialectal \(COD\)](#)

Ficheros en este ítem:

Fichero	Descripción	Tamaño	Formato
COD Ciutadella Fon.pdf		51,88 kB	Adobe PDF
COD Ciutadella For.pdf		39,62 kB	Adobe PDF
COD Ciutadella MP3.mp3		9,52 MB	MP3



Mostrar el registro Dublin Core completo de este ítem
















[Recomiende este ítem](#)

[Estadísticas de este documento](#)

" Estadísticas del ítem " [Corpus Oral Dialectal \(COD\). Ciutadella1](#) "Período [[este mes](#)] [[este año](#)] [[último año](#)] [[todos los años](#)]

estadísticas

Visitas	Descargas
120	80

País	Visitas	Descargas
 unknown	50	27
 Spain	19	22
 United States	18	13
 United Kingdom	0	6
 Germany	5	5
 China	0	3
 France	1	2
 Taiwan	1	1
 Lithuania	0	1
 Korea, Republic of	10	0
 Italy	8	0
 Ukraine	5	0
 Czech Republic	1	0
 Sweden	1	0
 Ireland	1	0

Dominio	Visitas	Descargas
unknown	63	30
yandex.ru	19	12
ecs.soton.ac.uk	0	6
Red-83-43-254.dynamidP.rima-tde.net	1	4
fil.ub.es	1	4
137.99.40.188.clients.your-server.de	0	4
Red-81-39-247.dynamidP.rima-tde.net	1	3

2. Anàlisi de les dades del COD

a) Anàlisi dialectomètrica i representació de la distància lingüística

Coord.: E. Clua

i) Complementar i contrastar la classificació de les varietats del català obtinguda mitjançant la metodologia del *Cluster Analysis* amb l'aplicació d'altres tècniques classificatòries

- Classificació **dendromètrica** a partir de la informació de les bases de dades del COD, amb anàlisi **fonològica** prèvia
- **Comparació** d'aquests resultats amb els obtinguts a partir de les formes **fonètiques** (en línia amb el treball de les escoles dialectomètriques de **Salzburg** i de **Gröningen**)

(V. treballs de Clua i Valls)

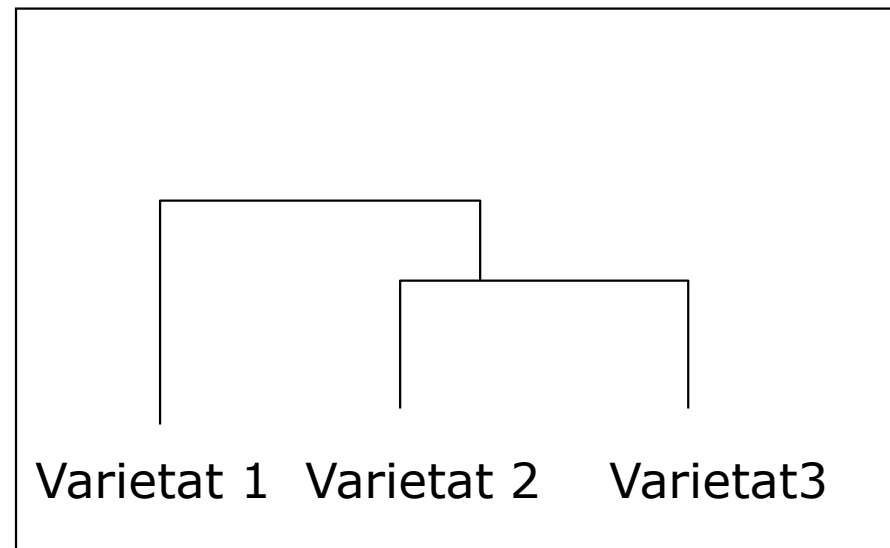
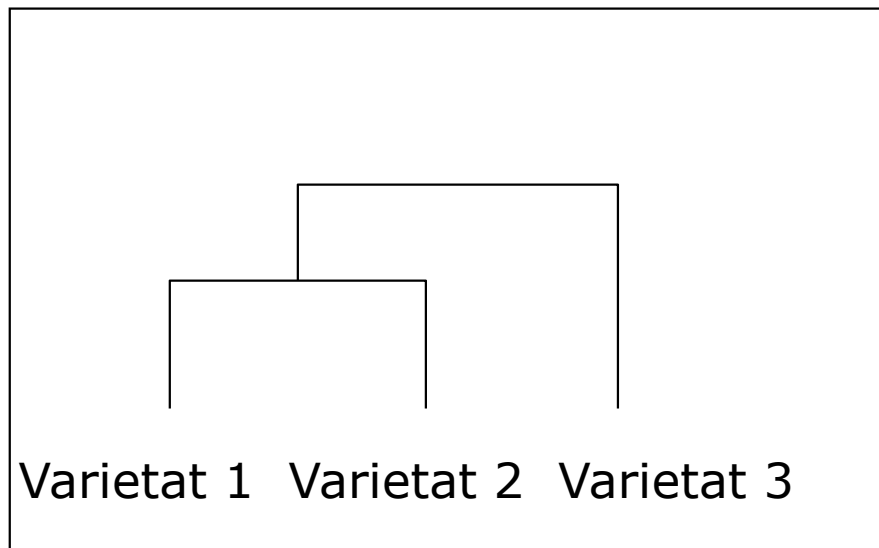
Exemple de pronoms clítics del valencià

Ortografia	Varietat 1	Varietat 2	Varietat 3
<i>et parle</i>	[et] parle	[te] parle	[te] parle
<i>t'anime</i>	[t] anime	[t] anime	[t] anime
<i>animant-te</i>	animant [te]	animant [te]	animant [te]
<i>anima't</i>	anima [t]	anima [t]	anima [te]

Distància lingüística basada en dades fonètiques:

	Varietat 1	Varietat 2	Varietat 3
Varietat 1			
Varietat 2	1		
Varietat 3	2	1	

Representació de la distància lingüística basada en dades fonètiques

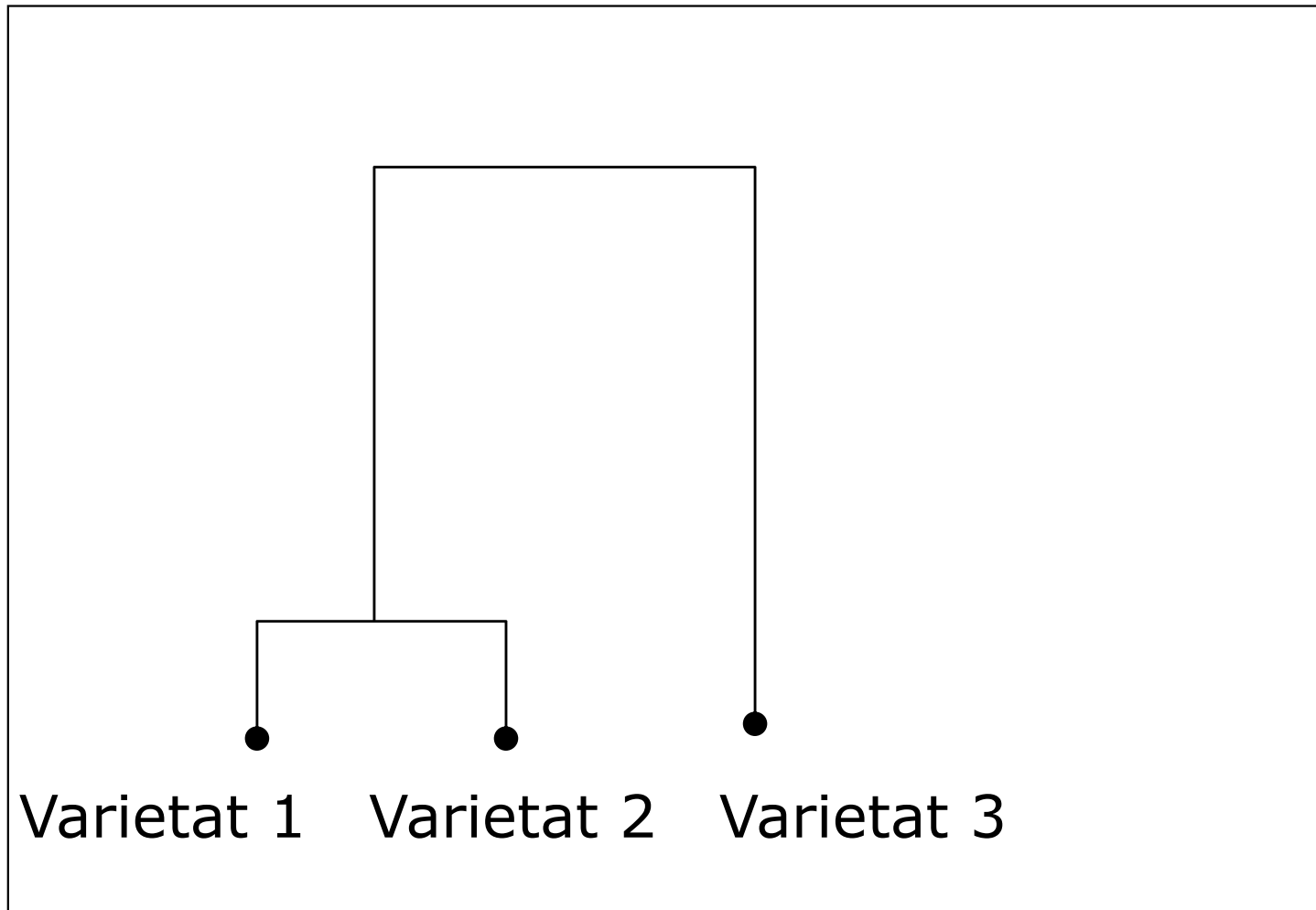


Ortografia	Varietat 1	Varietat 2	Varietat 3
<i>et parle</i>	[et] parle	[te] parle	[te] parle
<i>t'anime</i>	[t] anime	[t] anime	[t] anime
<i>animant-te</i>	animant [te]	animant [te]	animant [te]
<i>anima't</i>	anima [t]	anima [t]	anima [te]
Interpretació fonològica	/t/ epèntesi a perifèria	/t/ epèntesi a dreta	/te/ elisió vocals en contacte

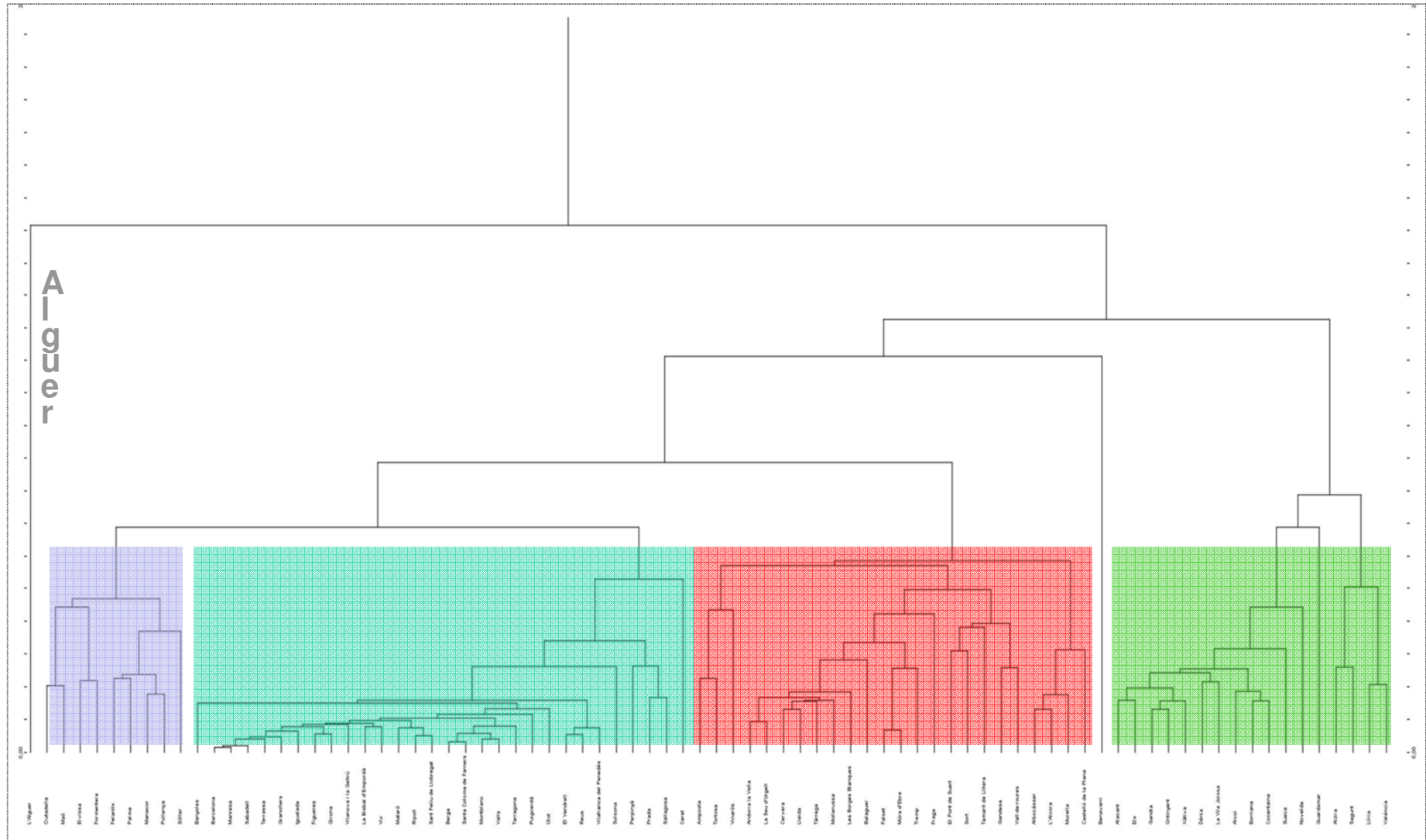
Distància lingüística basada en formes fonològiques:

	Varietat 1	Varietat 2	Varietat 3
Varietat 1			
Varietat 2	0		
Varietat 3	4	4	

Representació de la distància lingüística basada en l'anàlisi fonològica



Anàlisi dialectomètrica del COD (anàlisi fonològica prèvia de les dades)



V. balears

V. centrals i septentr.

V. nord-occident.

V. valencianes

ii) Aplicació de les dades del COD al *Visual Dialectometry Method* (VDM): representacions estadístiques i representacions espacials mitjançant mapes cloropets

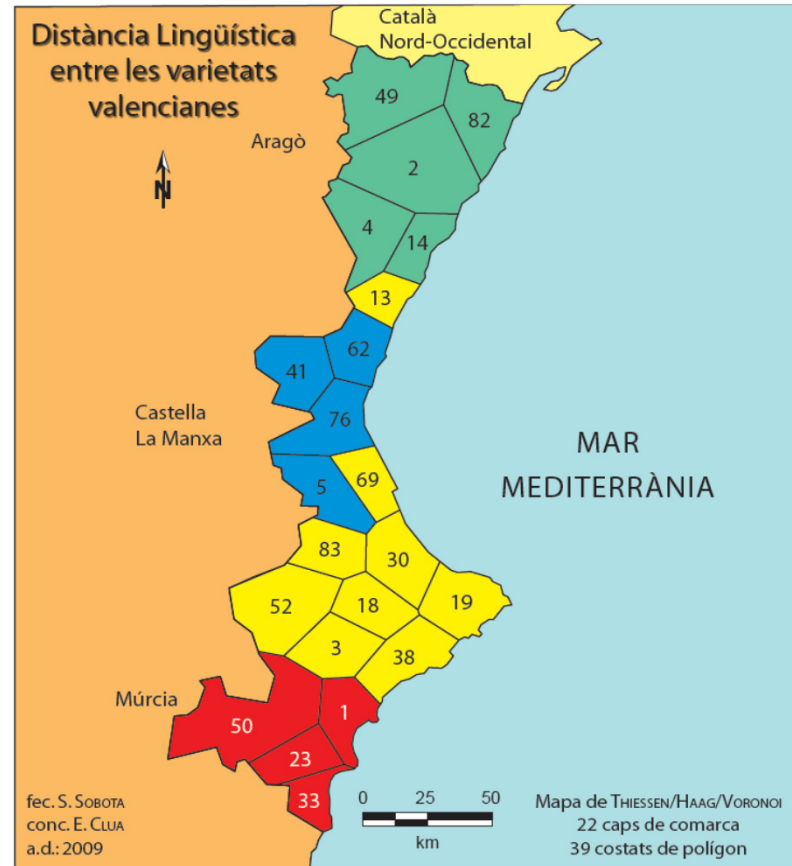
Mapes cloropets: Reflecteixen la distància lingüística de forma cromàtica:

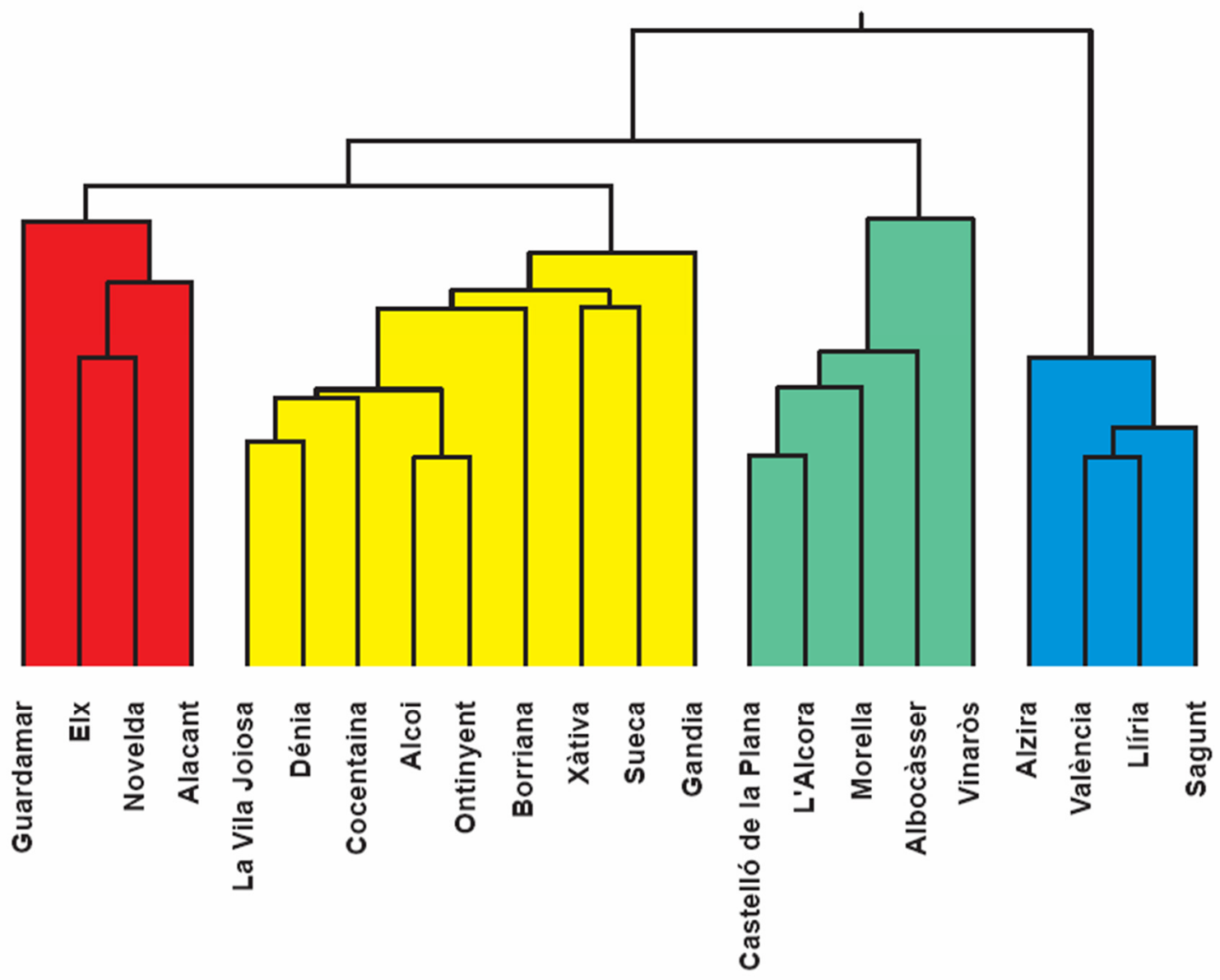
colors més **calents** → varietats més **similars**

colors més **freds** → varietats més **distants**

- En col·laboració amb l'escola de Salzburg (Goebl *et al.*), s'ha creat una **Applet** que permet visualitzar els resultats interactivament on-line

(De moment, amb l'anàlisi de les varietats valencianes)





http://www.iula.upf.edu/rec/vdm_val/

Corpus Oral dialectal (COD)

Maria-Rosa Lloret, Joaquim Viaplana, Maria Pilar Perea i Esteve Clua

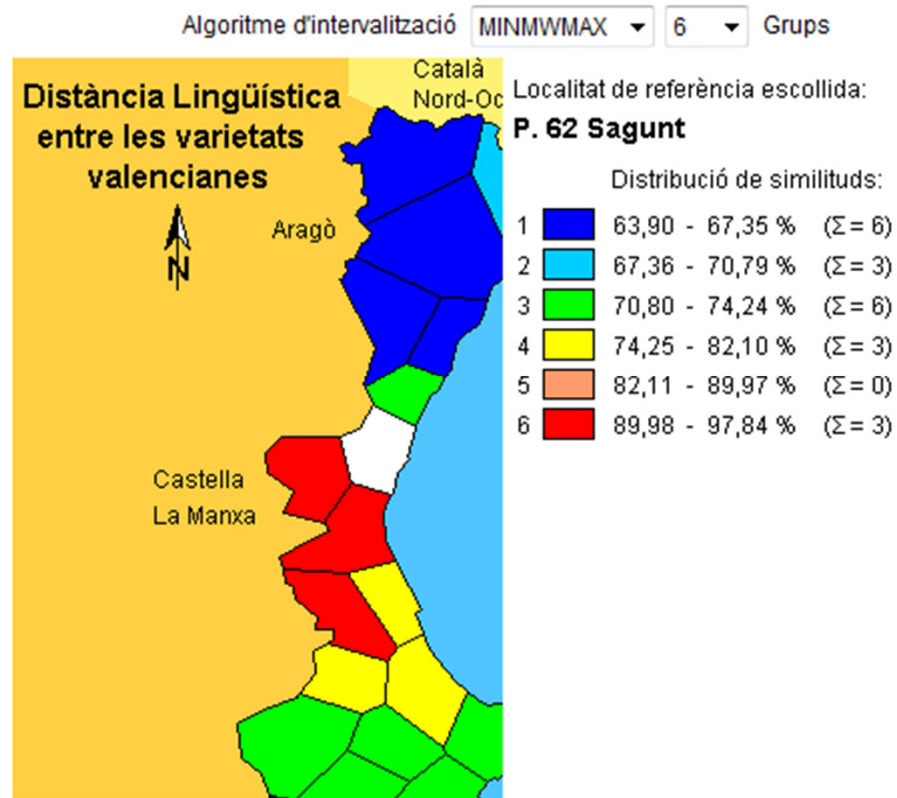
Applet Dialectomètrica

Distància Lingüística entre les varietats valencianes

Coordinació: Esteve Clua

Desenvolupat per: Slawomir Sobota

Global: Mapa de similitud

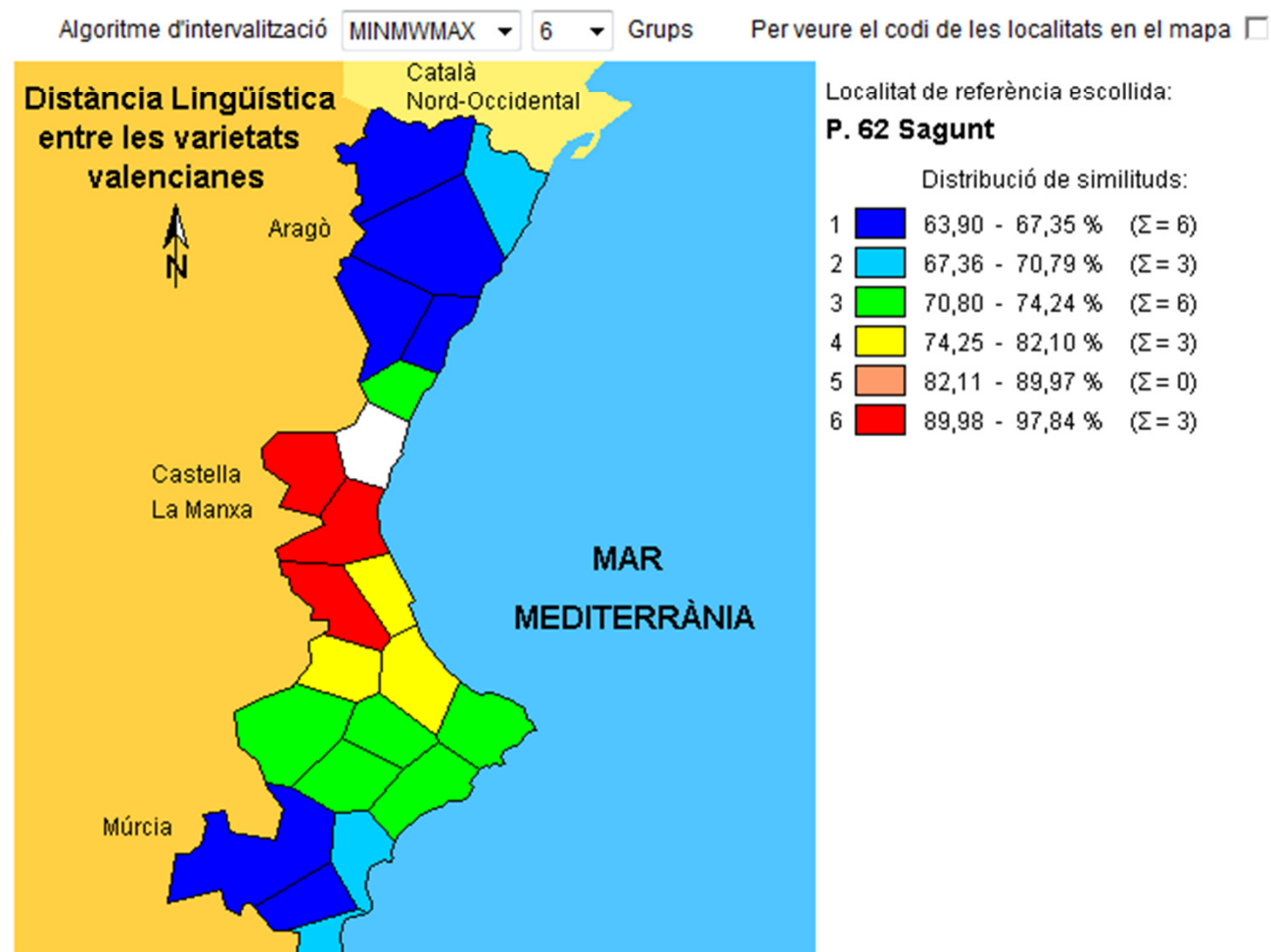


Distància Lingüística entre les varietats valencianes

Coordinació: Esteve Clua

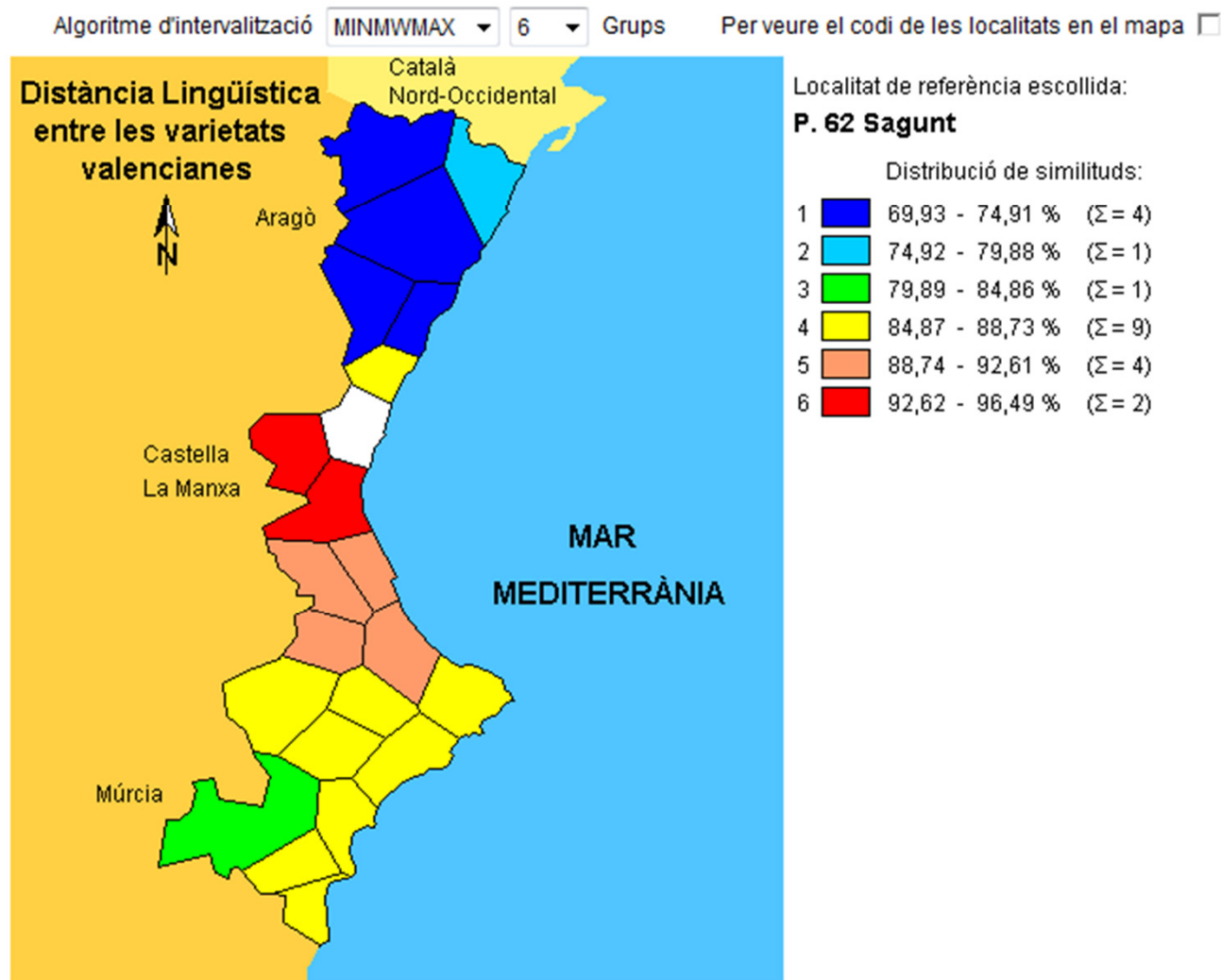
Desenvolupat per: Slawomir Sobota

Global: Mapa de similitud



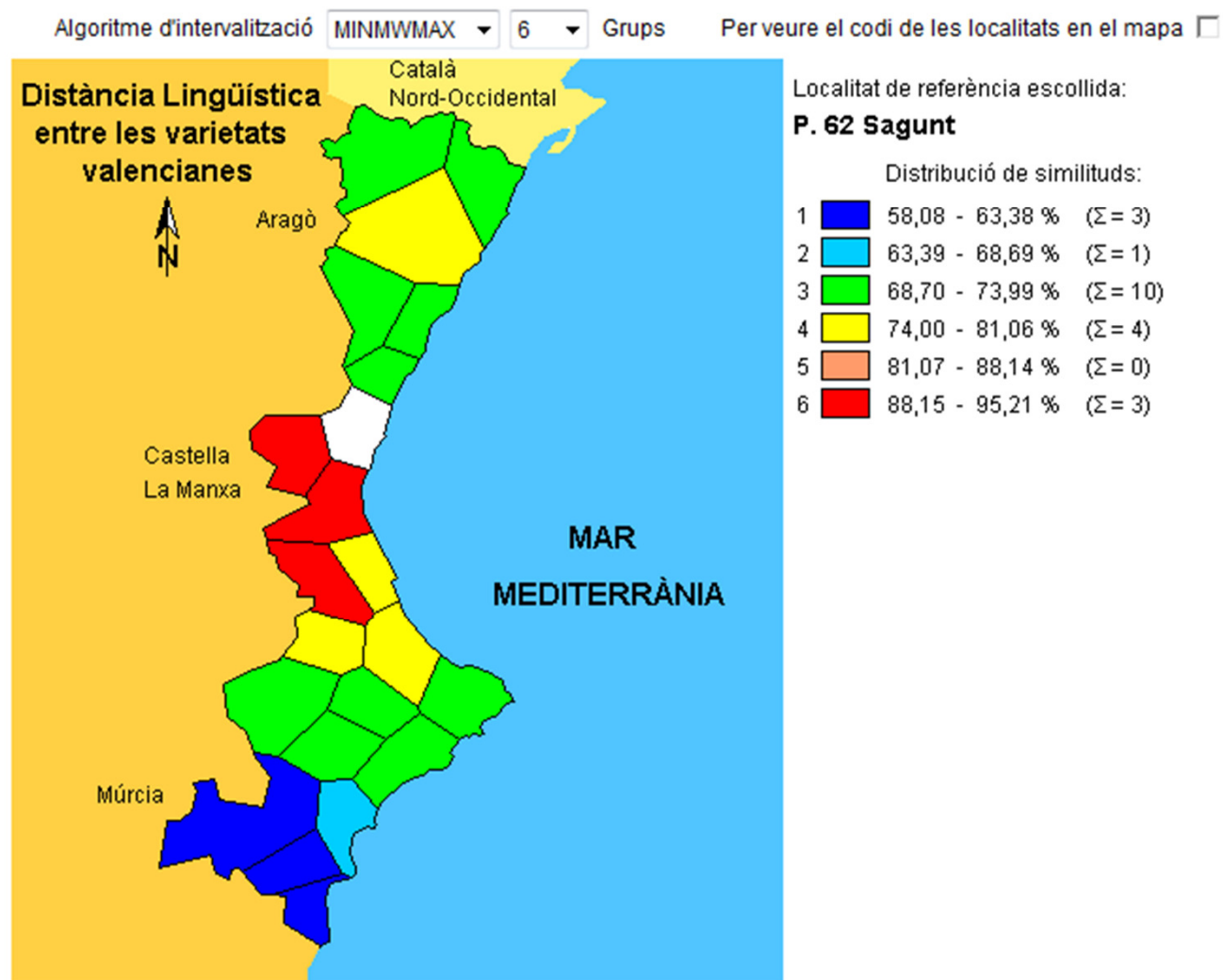
Distància Lingüística entre les varietats valencianes

Flexió verbal: Mapa de similitud



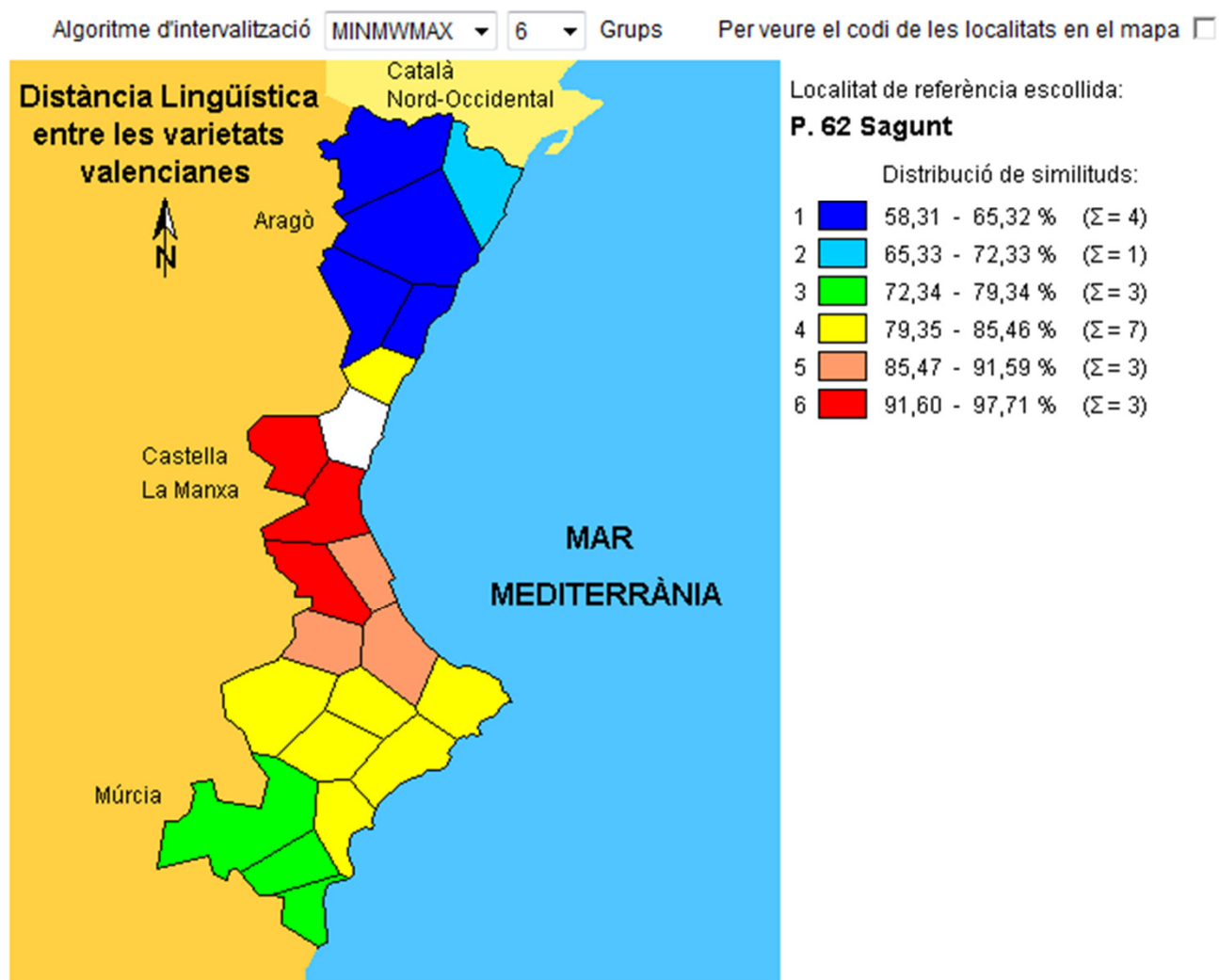
Distància Lingüística entre les varietats valencianes

Flexió nominal: Mapa de similitud



Distància Lingüística entre les varietats valencianes

Processos fonològics: Mapa de similitud



2. Anàlisi de les dades del COD:

**b) Anàlisi de fenòmens lingüístics
amb relació a la variació dialectal**

Coord.: M-R. Lloret

i) Estudi de mecanismes que permetin relativitzar els efectes de les generalitzacions lingüístiques i analitzar la variació lingüística.

Àmbit fonètic

- VOCALS: Descripció i estudio acústic de:
 - a) l'**harmonia vocàlica** en **valencià** (v. treballs d'Herrero i Jiménez)
 - b) les **vocals mitjanes tòniques** en català **central** (v. treballs de Callejón) i **nord-occidental** (v. treballs de Carrera).

Resultats més importants: diferències microdialectals en el timbre de les vocals; **conseqüències per a l'harmonia;** aspectes **sociolingüístics**.

- CONSONANTS: Estudi de la producció i percepció de realitzacions **fricatives** i **africades** de les palatals en català **nord-occidental** (v. treballs de Carrera).

Resultats més importants: aspectes **sociolingüístics** que determinen la producció i la percepció.

Àmbit fonològic

- Estudi comparatiu entre diversos **dialectes del català** i fenòmens afins en **altres llengües romàniques**:
 - a) VOCALS: processos d'**harmonia vocàlica**, **obertura** de vocals àtones i **epèntesis** (v. treballs de Jiménez i Lloret).
 - b) CONSONANTS: **assimilacions** de sonoritat i entre consonants (v. treballs de Jiménez, Lloret i Pons).

Resultats més importants: efectes de **prominència** (segmental i estructural) en els fenòmens fonològics; importància de la **similitud** articulatòria en les assimilacions; adequació de la **Teoria de l'Optimitat (TO)** per explicar la variació existent.

- Descripció i anàlisi de fenòmens consonàntics de l'**alguerès** relatius al **rotacisme** i la distribució de **líquides** (v. treballs de Cabrera), i a les **africades** (v. treballs de Bosch, Jiménez i Lloret).

Resultats més importants: **noves** dades sobre aquesta varietat; efectes de **prominència** analitzats en el marc de la **TO**.

Àmbit morfològic (flexiu)

- Estudis sobre diversos **dialectes del català**:
 - a) Naturalesa morfològica de les **vocals finals** dels elements nominals (v. treballs de Viaplana) i dels **segments verbals postradicals** velars i palatals (v. treballs d'Adam, Lloret, Marí i Viaplana, i tesi de Querol).
 - b) Predictibilitat en **al·lomorfia** (v. treballs de Lloret i tesi de Querol).
 - c) Sobreaplicació i bloqueig de fenòmens fonològics per **efectes paradigmàtics** (v. treballs d'Adrover, Lloret i Pons).

Resultats més importants: **analogies condicionades** per qüestions fonològiques i per efectes d'homogeneització paradigmàtica; **migració morfològica** condicionada; adequació de la **Teoria de l'Optimitat (TO)** i de la **Morfologia Natural** per explicar aquests fenòmens d'interfície.

- Descripció de la variació existent en la **flexió verbal** del **català septentrional** (v. treballs d'Adam, Adrover i Perea).

Resultats més importants: **noves** dades sobre aquesta varietat; estudi de **canvis en curs**; adequació de la **Morfologia Natural** per explicar aquests canvis.

Àmbit sintàctic

- Explotació de textos orals per a l'estudi sintàctic i morfofonològic de la distribució de les **preposicions *a/en/amb*** en diferents **dialectes del català**.

Resultats més importants: **imbricacions** dels tres àmbits formalitzats en el marc de la **Teoria de l'Optimitat** (v. treball d'Albareda).

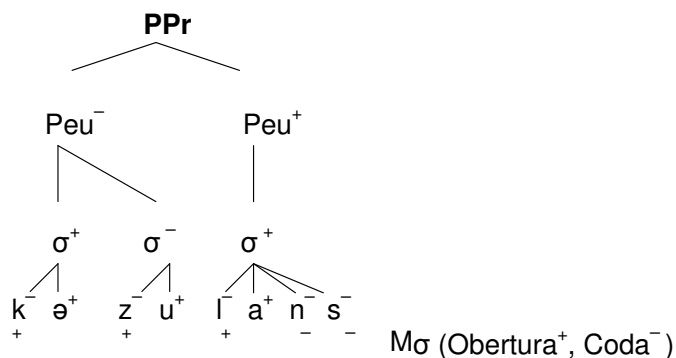
ii) Estudi de mecanismes que permetin diferenciar les característiques de l'estil formal vs. informal i de les varietats geogràfiques vs. estàndard.

- Diferències entre l'estil **formal** i l'**informal** (v. tesi de Campmany).
Resultats més importants: efectes de **marcatge** (naturalitat) en l'estil informal; adequació de la **Teoria de l'Optimitat** per explicar la variació.
- Constatació de la progressiva **desdialectalització**:
 - a) General en algunes àrees **nord-occidentals**: estudi comparatiu global de les característiques obtingudes a partir de les dades del COD (v. treballs de Clua, Viaplana i Valls) i estudi de zones de frontera específiques (v. tesi Rosàs i Valls).
 - b) Pèrdua de l'article salat, derivat de IPSU, en alguns **dialectes del català** (v. tesi de Busquet).
 - c) Pèrdua de característiques específiques de l'**alguerès** (v. treballs de Bosch i Cabrera).**Resultats més importants:** estudi de causes **sociolingüístiques** en els processos de desdialectalització.

EXEMPLE 1: Les 2 epèntesis de l'alguerès: a/i

Correlació entre prominència estructural i segmental:
posicions fortes/febles i elements més/menys sonants (Lloret&Jiménez)

Prominència de les posicions dins de la paraula prosòdica (PPr):



PPr	M	M	M	M	M	N	M	M
Peu	M	N	M	M	M	N	M	M
σ	M	N	M	N	M	N	M	M
Mσ	N		N		N		M	M

M = Marge (posició feble)
N = Nucli (posició forta)

/kaz+ulan+z/ : [kəz+u'lan+s] 'casolans'

Prominència de les posicions dins i fora de la paraula prosòdica (PPr):

/skriv/: [as'krif] 'escric'
/pɔk/: [pɔk] 'poc'
/skriv##pɔk/: [as,krifi'pɔk] 'escric poc'

	PPr+	PPr-	PPr+
	askrif	i	pɔk
Enunciat	N	M	N

• Al Nucli (posició forta), epèntesi [a] (vocal més sonant, més oberta)

• Al Marge (posició feble), epèntesi [i] (vocal menys sonant, més tancada)

EXEMPLE 2: Estudi comparatiu sobre vocalisme

Correlació entre prominència estructural i segmental:
posicions fortes/febles i elements més/menys sonants (Lloret&Jiménez)

Varietat	Tret que determina la prominència	Efecte de la prominència	Exemples
Gallec	Posició pretònica (amb efectes paradigmàtics)	Manteniment de l'obertura	belleza, asudo, asino, abxectivo gobernación vs. gobernar, botar vs. botar
Català (Mallorca)	Síl·laba àtona inicial (amb efectes paradigmàtics)	No reducció a [ə] (manteniment de [e])	pedreta, pedrota vs. pàdrera vs. papàret, papàrot
Català, gallec, italià	Síl·laba tònica	Obertura de vocals en préstecs i cultismes	estap, etica, tècnica (IEC [jɛk] cat.) bola vs. bōla ('esfera') (gall.)
Català	Síl·laba tònica inicial	Obertura de <i>o</i>	olla, hora, flor
Castellà, català...	Síl·laba inicial	Preferència per la vocal [a]	NOVACULA > navaja, navalla
Castellà, català...	Síl·laba inicial	Manteniment de hiats	bi.òleg vs. ra.diò.leg (cat.)
Valencià (Canals)	Posició inicial absoluta de paraula (síl·laba àtona)	Manteniment de l'obertura	abrim, amplim vs. portem
Català occidental	Posició inicial absoluta de paraula	Vocal epentètica: [a]	[a]spina vs. pobr[e]
Valencià (Pedreguer)	Posició proclítica	Vocal epentètica: [a]	[at] parle vs. parlant-[te]

En teoria de l'optimitat, exemples 1 i 2 mostren
efectes **Marcatge Posicional** i **Fidelitat Posicional**
(amb efectes de **fidelitat paradigmàtica** entre formes superficials)

EXEMPLE 3: Estudi comparatiu sobre harmonia vocàlica

Artulatori o perceptiu? (Jiménez & Lloret)


UNIFORMITAT GESTUAL (base articular):

- **Fidelitat Posicional:**
- El tret que s'escampa apareix en una posició forta i s'escampa cap a posicions febles **per minimitzar l'esforç artulatori.**

- Ex.: Tònic → Àton 'ε e → 'ε ε
- 

LEGITIMACIÓ PERCEPTIVA (base funcional):

- **Marcatge Posicional:**
- El tret que s'escampa apareix en una posició feble i s'escampa cap a posició més forta **per esdevenir més perceptible.**

- Ex.: Àton → Tònic 'ε e → 'e e
- 

EXEMPLE 3: Estudi comparatiu sobre harmonia vocàlica (Jiménez & Lloret)

	ARTICULATORI	PERCEPTIU	ELEMENTS AFECTATS
Espanyol (andalús)	0 general: 'kɔmikɔ ^h variació: 'trɛβolɛ ^h ~ 'trɛβɔlɛ ^h	1	a) Sí·l·laba accentuada ('trɛβolɛ ^h) b) Sí·l·laba accentuada i posttònica interna ('trɛβɔlɛ ^h) [+ benefici articuladori] c) Paraula (mɔ'mɛntɔ ^h) [+ benefici articuladori]
Asturià (Lena)	0 'peʃaru	1	Sí·l·laba accentuada ('peʃaru vs. 'paʃara)
Asturià (montañés Tudanca)	2	1	Sí·l·laba accentuada i posttònica interna (centralització) (o'rɛgAnU)
Asturià (montañés pasiego)	2	1	Paraula (centralització) (rɪ'frjAtɪkU)
Català (valencià)	1	2	Sí·l·laba posttònica adjacent (tova'ʎɔlɔ vs. 'mɛðika)
Català (sud d'Alacant)	1	2	Paraula (tɔvɔ'ʎɔlɔ)
Català (Tortosa)	1	2	Sí·l·labes pretòniques adjacents (ri'vista, ini'mik, disi'kat vs. 'bimen, 'diye)

1 = factor detonant de l'harmonia; 2 = benefici secundari de l'harmonia; 0 = factor/benefici no rellevant en l'armonia 68

- Alguns dels treballs disponibles a:

<http://www.uv.es/foncat/>

Coord.: J.Jiménez

El treball analític continua en el projecte coordinat:

- FFI2010-22181-C03-03 (IP: Esteve Clua)

**Distància lingüística en els eixos espacial i temporal:
aspectes fonològics i morfològics del català**

- FFI2010-22181-C03-02 (IP: Maria-Rosa Lloret)

**Descripció i interpretació de la variació dialectal:
aspectes fonològics i morfològics del català**

- FFI2010-22181-C03-01 (IP: Joan Mascaró)

**Anàlisi teòrica dels processos segmentals i
morfofonològics del català**

FFI2010-22181-C03:

**Estudi de la fonologia i la morfologia del
català: descripció, teoria i variació**

Moltes gràcies per la seva atenció

Moitas grazas pola súa atención

mrosa.lloret@ub.edu