THE MID BACK VOWEL MERGER IN GIRONA CATALAN:

Acoustic Evidence

EVA BOSCH-ROURA

New Ways of Analyzing Variation 45
November 3-6 2016

UNIVERSITAT DE BARCELONA

OUTLINE

1. Introduction

2. Background

3. Hypothesis and goals

4. Methods

5. Results

6. Conclusions

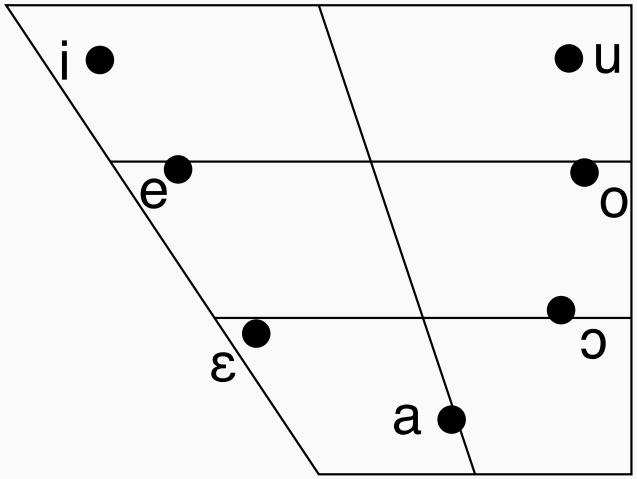
1. INTRODUCTION

CATALAN


Barcelona
l'Alguer
Palma
València
0km 100km200km
Google
Map data ©2016 GeoBasis-DE/BKG (©2009), Google, Inst. Geogr. Nacional

CATALAN VOWELS

- Most Catalan varieties (including the standard) have a 7-stressed vowel system:

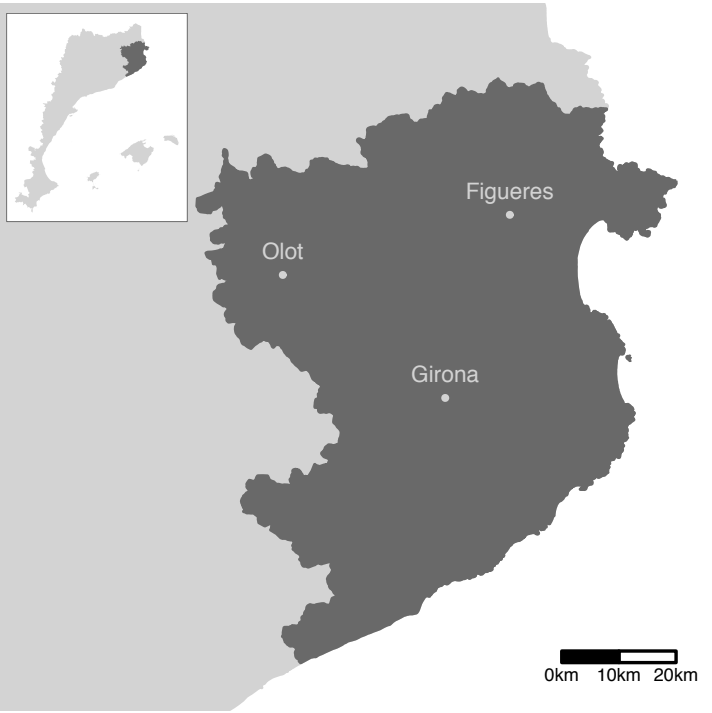


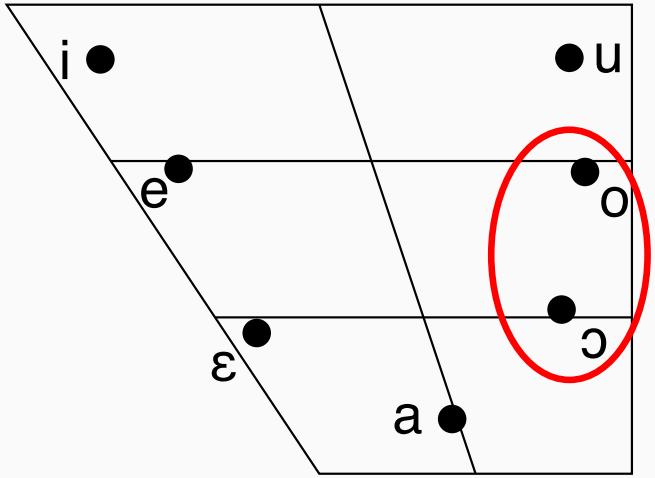
Go check Els sons del català at

http://www.ub.edu/sonscatala/en!

CATALAN VOWELS IN GIRONA

- Most Catalan varieties have a 7-stressed vowel system, but there could be a mid back vowel merger in the subdialect spoken in the area of Girona:





2. BACKGROUND

DIALECTOLOGY ON /o/ AND /ɔ/ IN GIRONA CATALAN

- According to **Coromines (1953)**, ~ 30 words with an etymologically closed *o* in the initial syllable are pronounced with an /ɔ/ in all Catalan varieties, except in Girona and Elna, where /o/ is used.
- Only, with each mention of the theory, the list of words just keeps getting longer (>180 words).

  (Comas 1970: 58-59; Sala 1983: 29; Julià 1986; Puig 1990: 31; Gulsoy 1993: 92; Luna 1995; Monturiol & Domínguez 2002: 28-29; Campmany 2004: 18; Dorca 2008: 24-27, 2010: 152; IEC 2009: 42; Gómez 2011: 51)
- Colomina (1999: 122-124) adds: [ɔ] → [o] / _[n]

ACOUSTIC REMARKS ON /o/ AND /ɔ/ IN GIRONA CATALAN

- Alcover (1908: 328-330), Luna (1995: 283-284), Julià (1986: 132), and Recasens (1996: 131): /ɔ/ *seems* specially *closed* in Girona, and there might not be an /ɔ/-/o/ distinction
- Acoustic studies by Recasens & Espinosa (2009) and Herrick (2003, 2006): no statistical differences between /o/ and /ɔ/ in Girona Catalan speakers (n = 5 + 3)

↓

Girona vowel system: /i/, /e/, /ɛ/, /a/, /O/, /u/

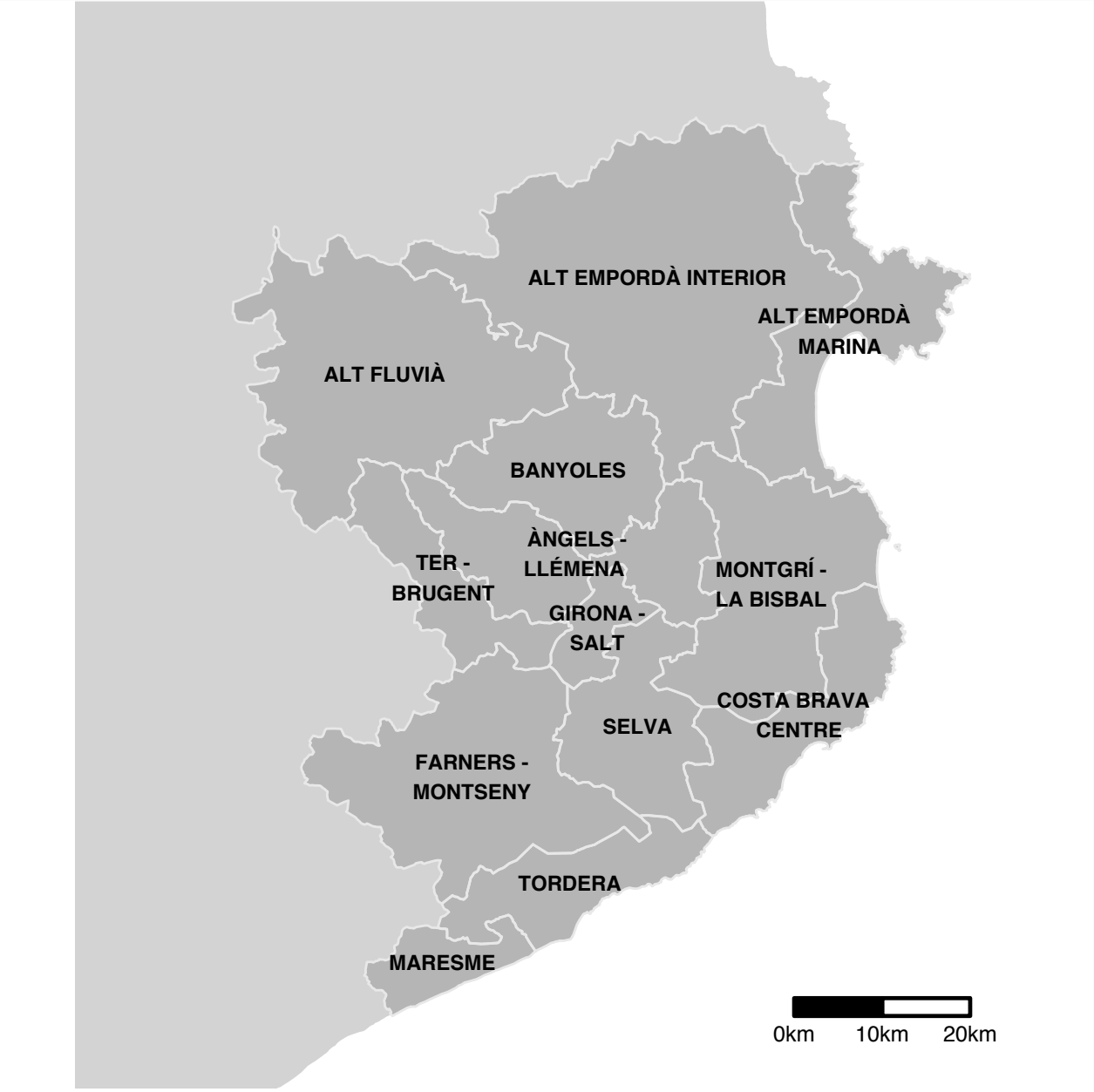
3. HYPOTHESIS AND GOALS

HYPOTHESIS AND GOALS

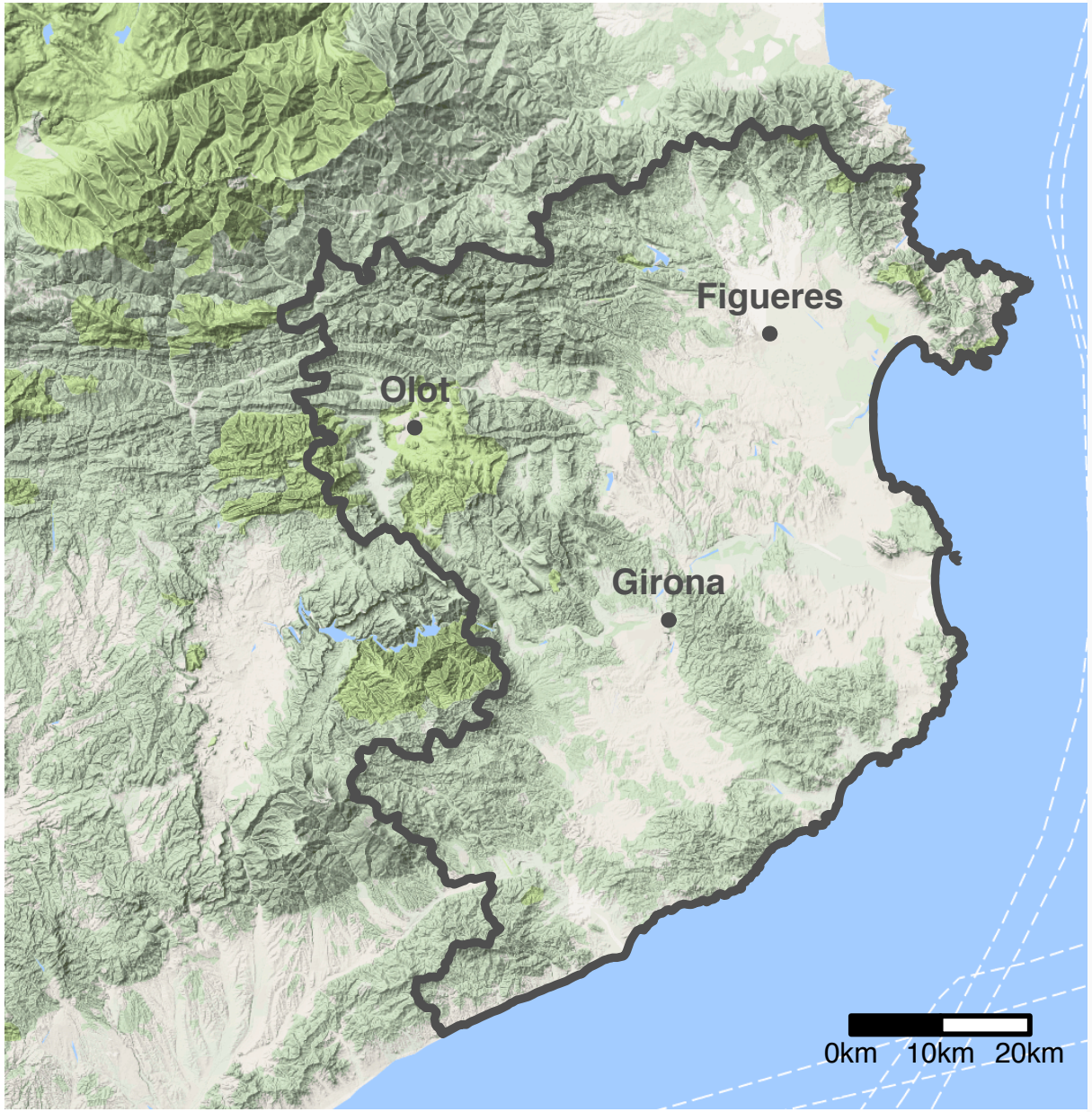
- **Hypotheses:** Mid back vowels are involved in an unconditioned merger in Girona Catalan
- **Goals:**
  - Test the production of /o/ and /ɔ/ —and the whole vowel system
    - in a variety of phonological contexts
    - produced by speakers in different areas of the Girona region
    - produced by speakers of two different age groups
    - produced by male and female speakers
  - Test the perception of /o/ and /ɔ/ —and the whole vowel system

4. METHODS

SURVEY AREAS



SURVEY AREAS



PARTICIPANTS

- **96 speakers** (4, 8, 12 or 16 per survey point, proportionally to population*):
  - 24 m + 24 f, born 1998 (15-17 years old)
  - 24 m + 24 f, born 1948-1957 (57-65 years old)
  - From Catalan-speaking families
  - Native and life-long residents of each survey point
  - Parents born in the same survey area
  - Secondary or higher education

PRODUCTION TESTS

! Many words differentiated by graphical accents in writing, and there are very few minimal pairs:

os/ós, bota/bóta, dona/dóna, mora/móra, (amor/mort)

- **Picture naming task (of sorts):**
  - T1: 7 vowels x 7 contexts (n=49) → syllable, previous and following segment
  - T2: Words mentioned in literature + minimal pairs (n=26)
- **Replication of the reading task in Recasens & Espinosa (2009):**
  - T3: 7 vowels x 4 contexts x 3 repetitions (n=28x3)

* T1+T3: 3,712 tokens, T2: 2,377 tokens

DATA EXTRACTION AND PROCESSING

- Semi-automatic segmentation and alignment with SPPAS (Bigi 2015)
- F1 and F2 values estimated by LPC formant-tracking algorithm in Praat (Boersma & Weenink 2016) and extracted at 7 points throughout each interval with a semi-automatic script (adapted from McCloy & McGrath 2014)
- Data Lobanov-normalized with `vowels` by Kendall & Thomas (2014) in R (R Core Team 2016)

VISUALIZATION, QUANTIFICATION & ANALYSIS

- **Density:** F1xF2 2D KDE contour plots (e.g.: Nycz & Hall-Lew 2013; DiCanio et al. 2015; Amengual 2016; `ggplot2::stat_density2d`, Wickham 2009)

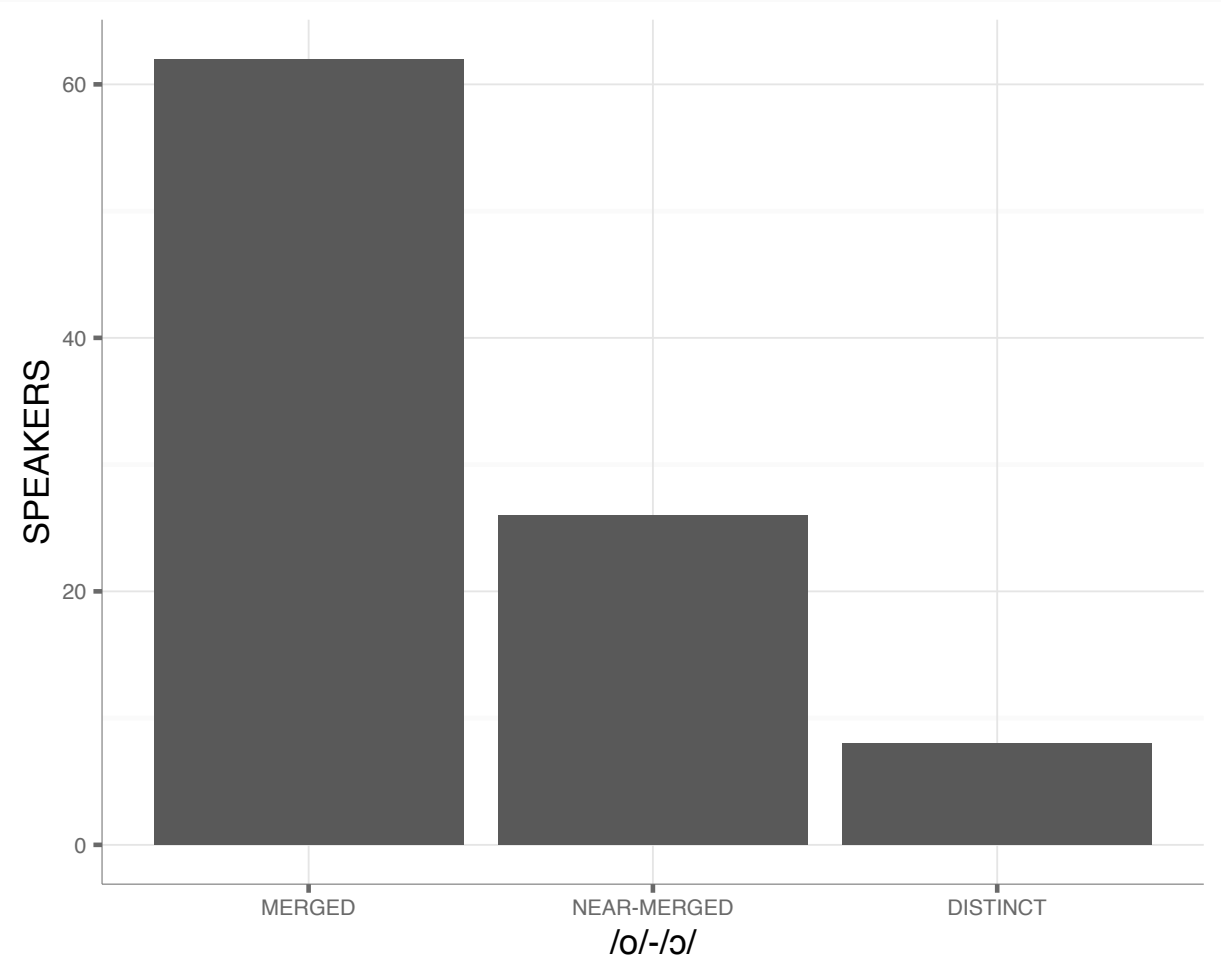
- **Distance:** Euclidean distances /o/-/ɔ/ from F1 & F2 means (e.g.: Baranowski 2013; Dinkin 2011; Evanini 2009; Hay et al. 2009; Recasens & Espinosa 2009; `stats::dist`, R Core Team 2016)
- **Overlap:** Bhattacharyya Coefficient (as proposed by Johnson 2015; `adehabitatHR::kerneloverlap`, Calenge 2006)
- **Trajectory:** SS ANOVA (e.g.: Nycz & Decker 2006; Wassink & Koops 2013; Docherty et al. 2015; Fruehwald 2010; `gss::ssanova`, Gu 2014)

- **Variation:** Mixed-effects models (e.g.: ehem, everyone; `lme4::lmer`, Bates et al. 2015)

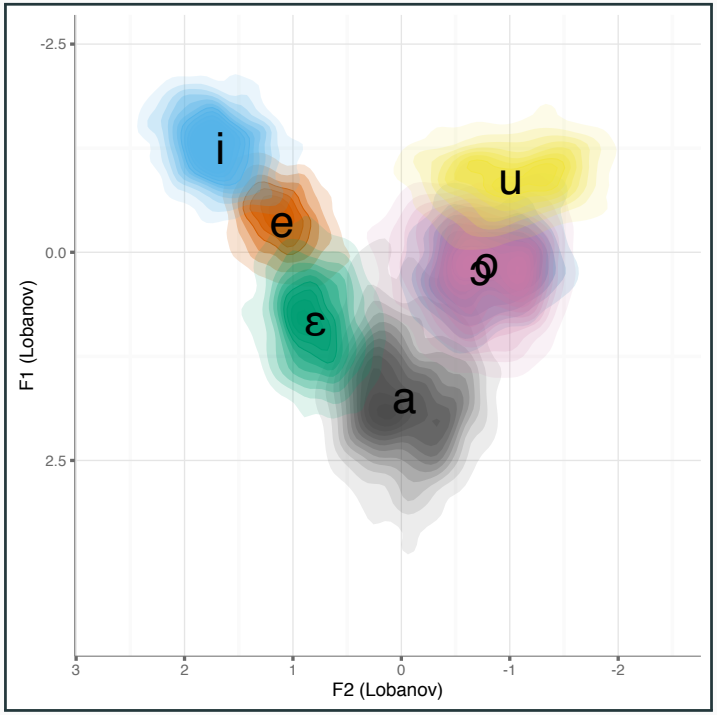
5. RESULTS

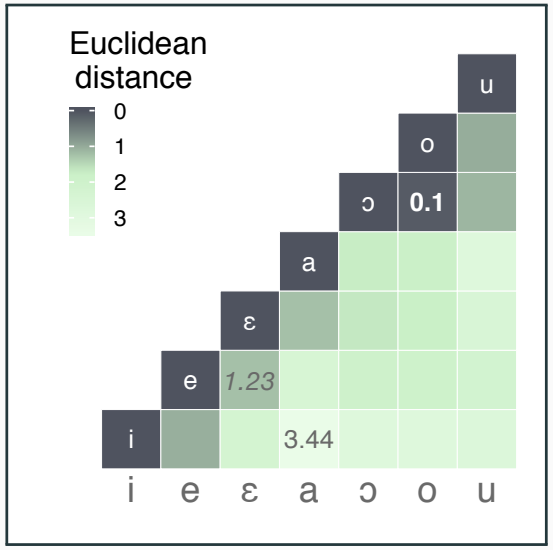
QUANTIFICATION & CLASSIFICATION

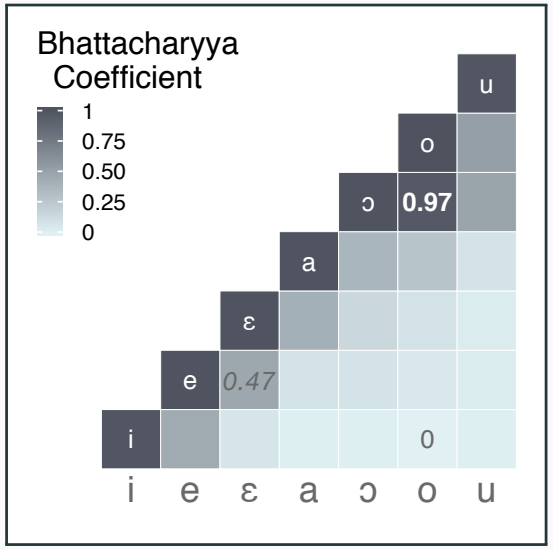
- Individual classification of speakers' mid back vowels:

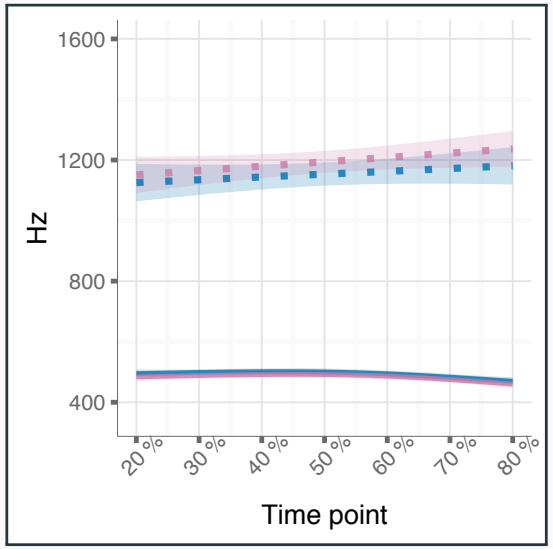


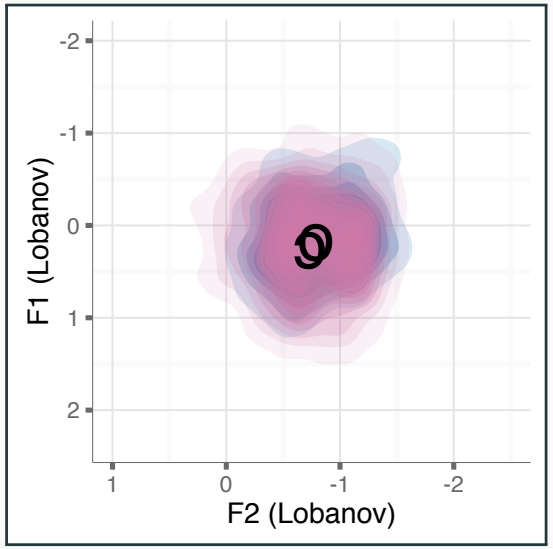
Q & C: MERGED SPEAKERS (62)



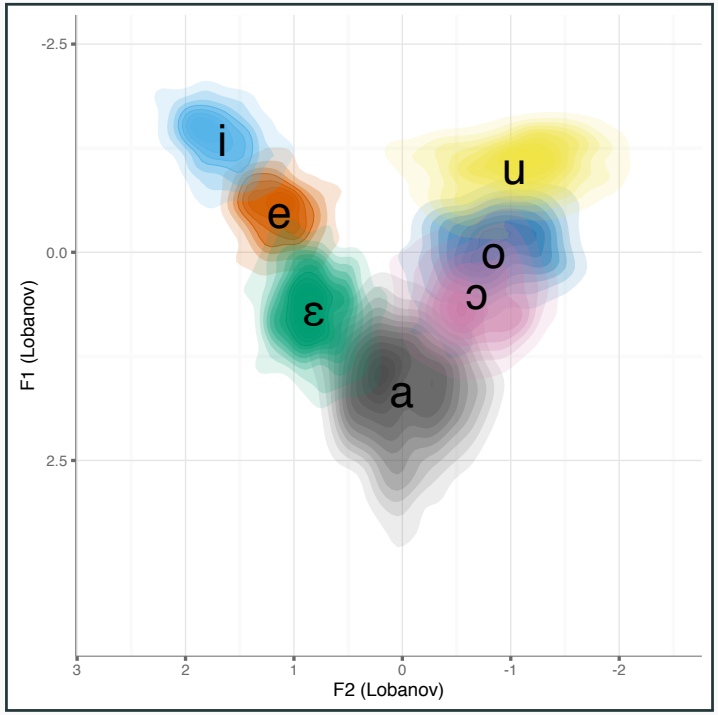


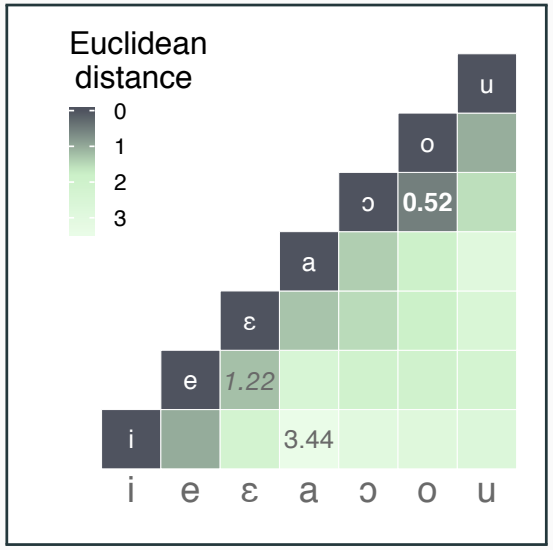


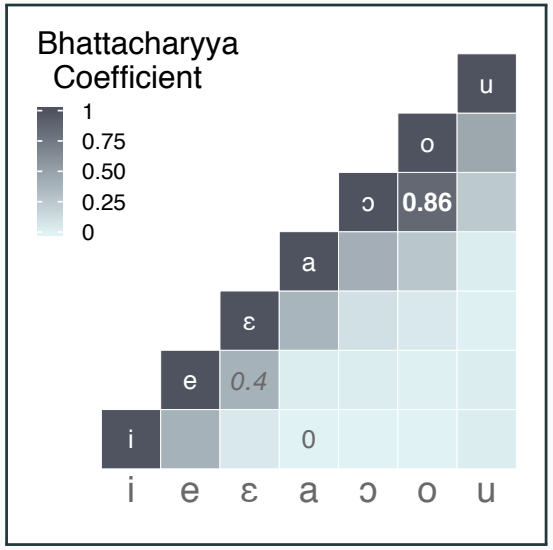


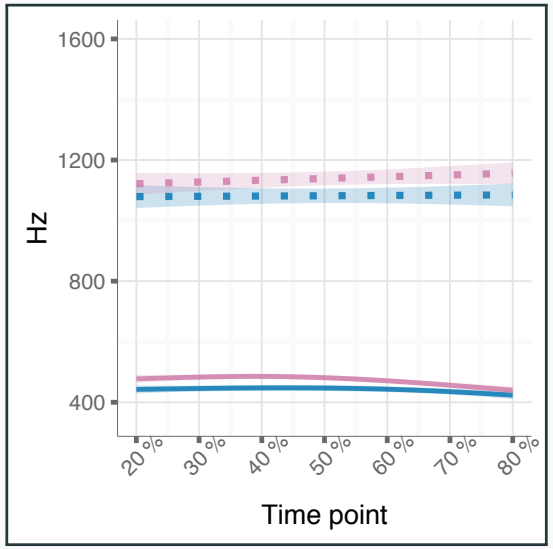


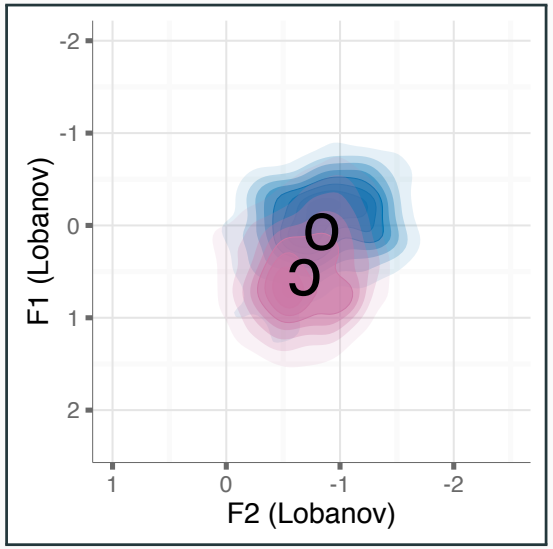
Q & C: NEAR-MERGED SPEAKERS (26)



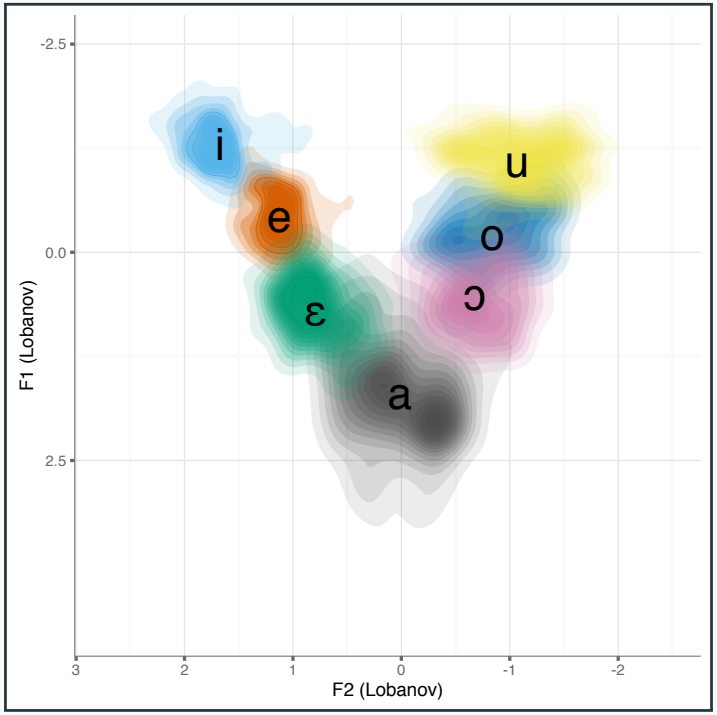


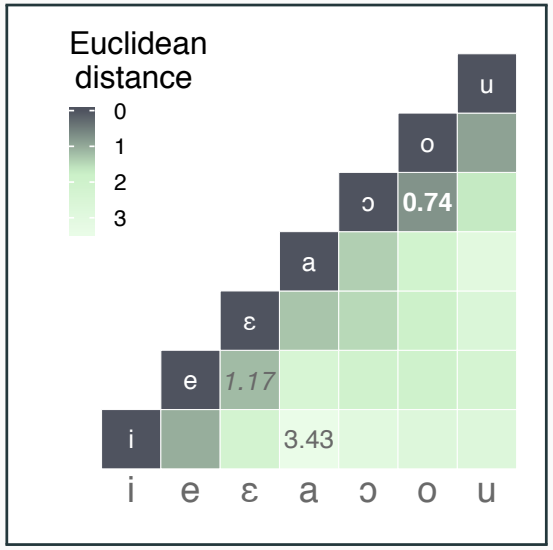


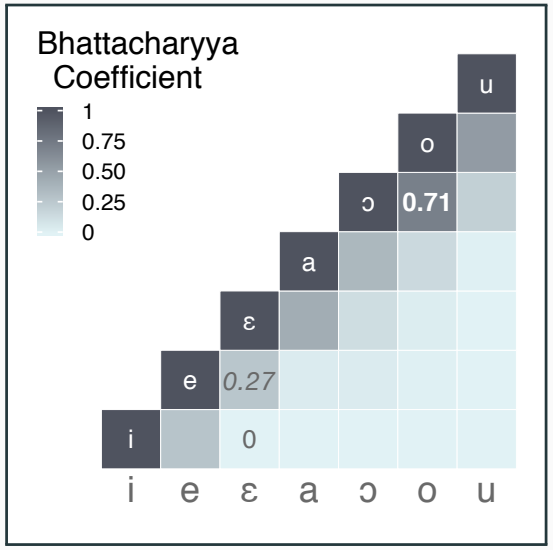


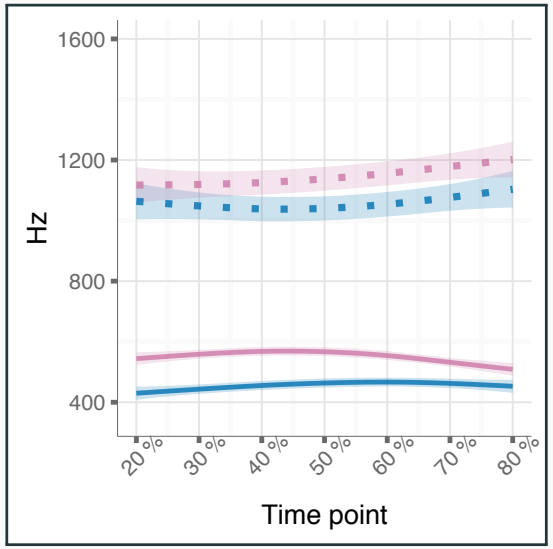


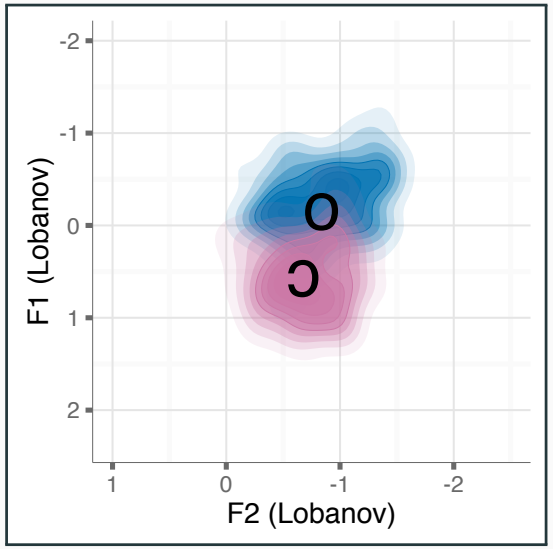
Q & C: DISTINCT SPEAKERS (8)











HOW ACCURATE IS THIS CLASSIFICATION?

F1 LOBANOV ~VOWEL + (1|WORD) + (1|SPEAKER) + (1|PreS) + (1|PostS):

	Estimate	SE	t	Pr(>|t|)	
M1 (M):	0.088	0.091	0.966	0.346	
M2 (NM):	0.41	0.139	2.956	0.0902	**
M3 (D):	0.5891	0.1206	4.884	0.000152	***

DOES SYLLABLE PLAY A ROLE?

- Coromines 1953: /ɔ/ → /o/ in initial syllable
  - Merged /o/-/ɔ/ in FINAL position for
    - near-merged speakers (β=0.1112, SE=0.4902, t=0.227, p=0.842)
    - distinct speakers (β=0.224, SE=0.409, t=0.548, p=0.650)

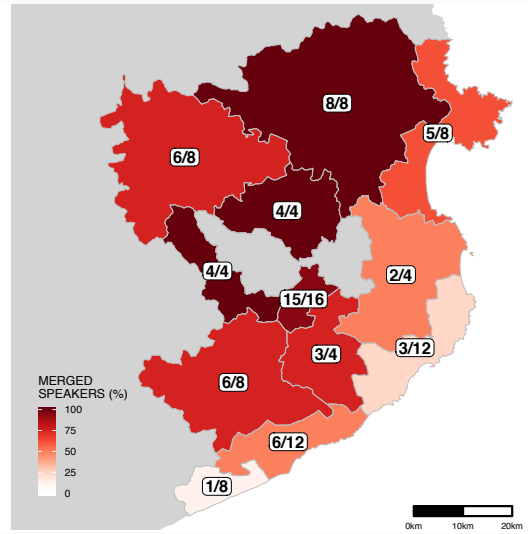
  → just accuracy of the production?

HOW ABOUT *THE LIST*?

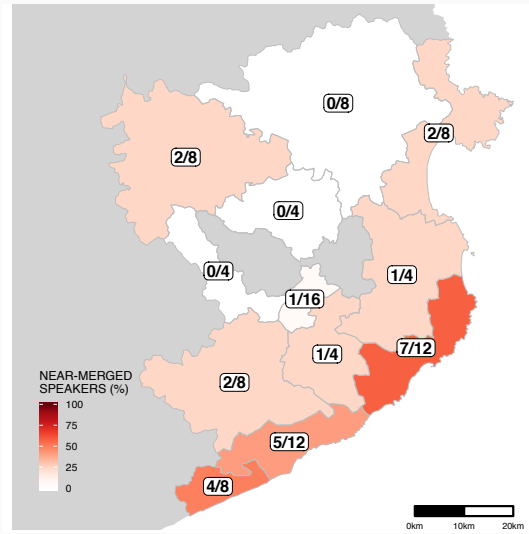
- Words identified by Coromines (1953) and subsequent work, and also minimal pairs, behave just exactly like the mid back vowels produced in T1 and T3 for each of the three (merged, near-merged, distinct) groups of speakers

GEOGRAPHICAL DISTRIBUTION OF THE MERGER I

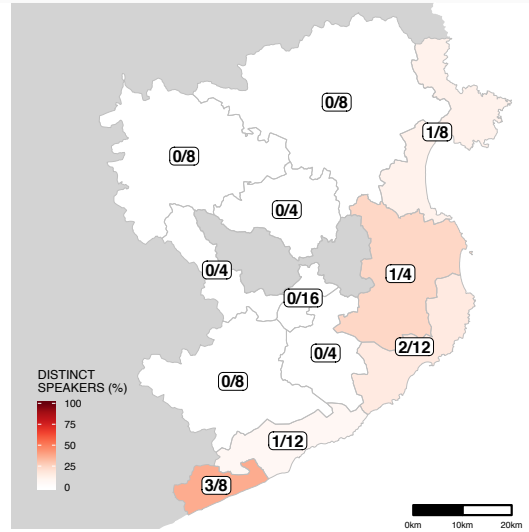
Merged



Near-merged

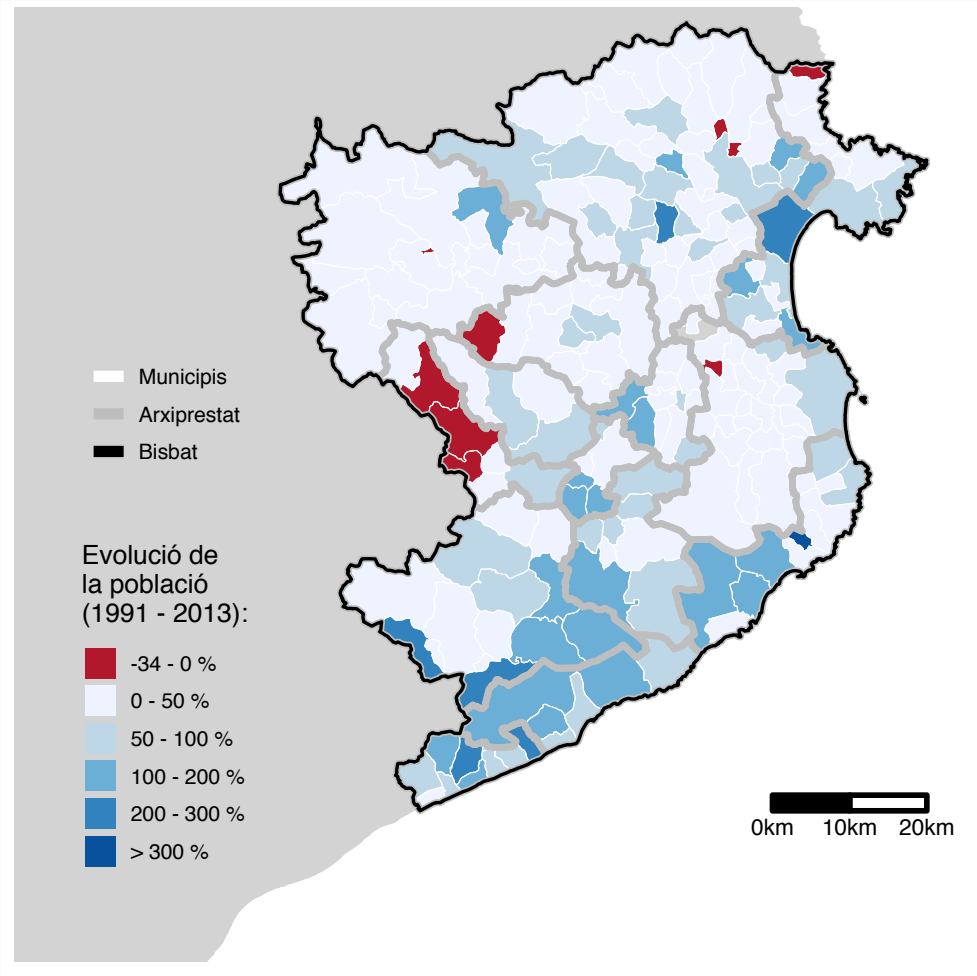


Distinct



GEOGRAPHICAL DISTRIBUTION OF THE MERGER II

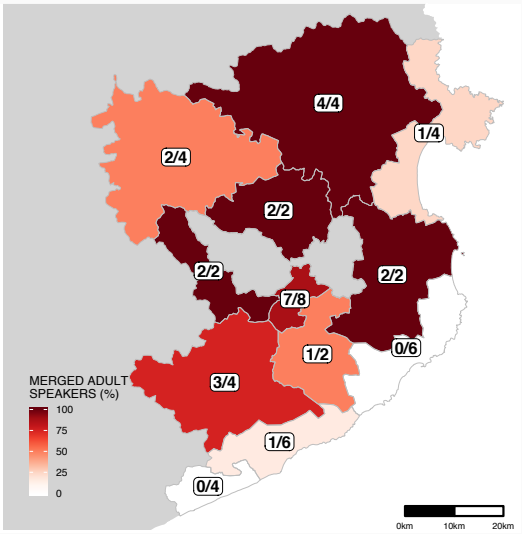
- Major demographic and sociolinguistic differences between interior and coast areas

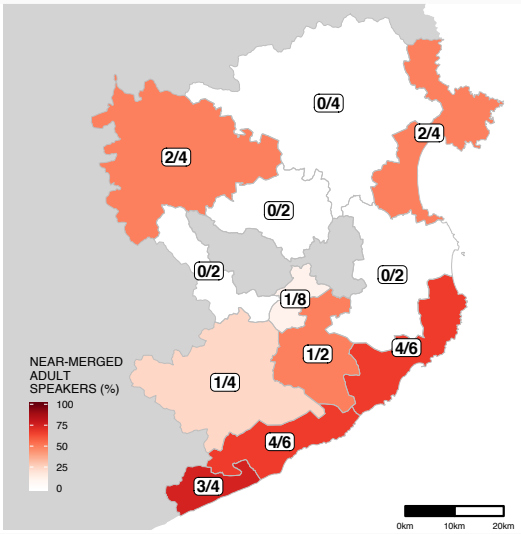


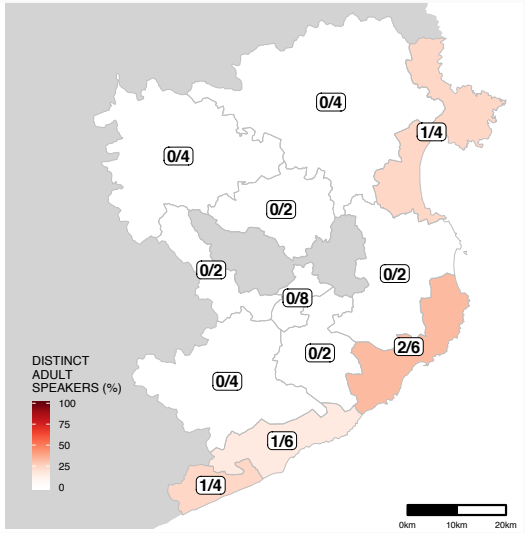
MERGED SPEAKERS BY AREA AND AGE I

Merged | Near-merged | Distinct

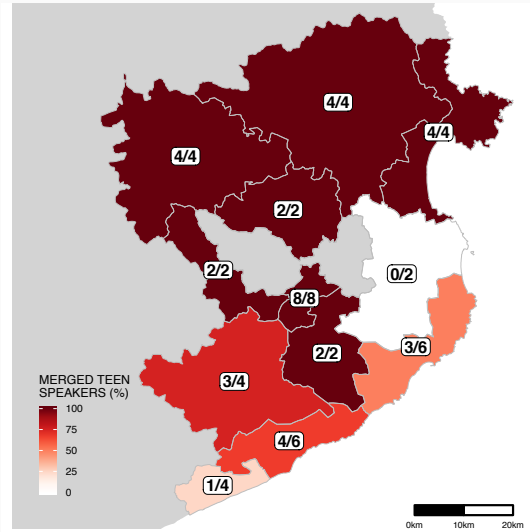
Adults

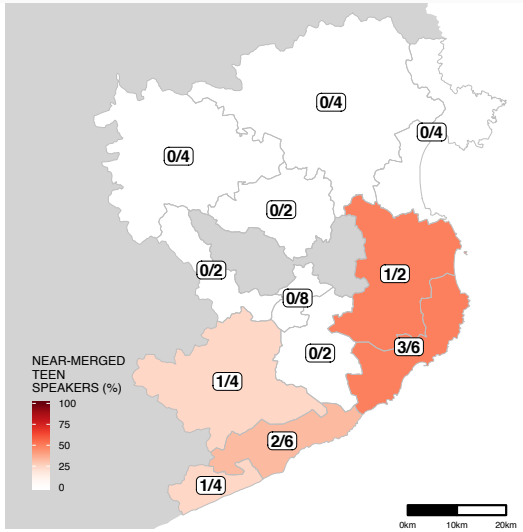


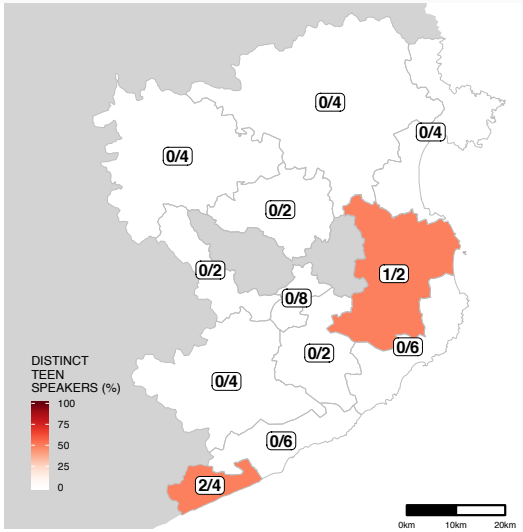




Teens







MERGED SPEAKERS BY AREA AND AGE

- Garde's Principle: "mergers are irreversible by linguistic means" (Labov 1994: 311), but "Given the right social conditions, it is reasonable to think that a distinction can be reintroduced into a speech community in a consistent way" (Labov 1994: 342):
  - Orthography might be playing a big part on the resistance to / denial of the merger → BREAKING NEWS: IEC just decided to remove most diacritic accents affecting minimal/quasi-minimal pairs
- Based on their speech as a whole, the diffusion of the merger among younger speakers in coast areas could be influenced by Spanish (clear instead of dark alveolar /l/, delateralization of alveolopalatal /ʎ/ > /j/, lowering of /ə/) + general comments on language by speakers *and their teachers*

6. CONCLUSIONS

CONCLUSIONS

- The unconditioned merger of /o/ and /ɔ/ in production is complete in the interior part of Girona, and is not complete but expanding in coast areas
- Orthography and the standard variety might be influencing (younger) speakers away from the merger in areas where:
  - the merger is not complete
  - there are important demographic movements and diversity
- Spanish might be influencing younger speakers towards the merger in those same areas → it all depends on each speaker's sensitivity towards language
- Perception results coming soon!

REFERENCES I

Alcover, A. M. (1908). “Una mica de dialectologia catalana”. *BDLC* 4, 291–438. URL: http://taller.iec.cat/filologica/alcover/documents/4BDLC_04.pdf.

Amengual, M. (2016). “Cross-Linguistic Influence in the Bilingual Mental Lexicon: Evidence of Cognate Effects in the Phonetic Production and Processing of a Vowel Contrast”. *Frontiers in Psychology* 7. DOI: 10.3389/fpsyg.2016.00617.

Baranowski, M. (2013). “On the role of social factors in the loss of phonemic distinctions”. *English Language and Linguistics* 17.02, 271–295. DOI: 10.1017/S1360674313000038.

Bates, D., M. Mächler, B. Bolker, & S. Walker (2015). “Fitting Linear Mixed-Effects Models Using lme4”. *Journal of Statistical Software* 67.1, 1–48. DOI: 10.18637/jss.v067.i01.

Bigi, B. (2015). “SPPAS - Multi-Lingual Approaches to the Automatic Annotation of Speech”. *The Phonetician* 111-112, 54–69. URL: http://www.isphs.org/Phonetician/Phonetician_111-112.pdf.

Bivand, R., T. Keitt, & B. Rowlingson (2016). *rgdal: Bindings for the Geospatial Data Abstraction Library*. R package version 1.1-10. URL: https://CRAN.R-project.org/package=rgdal.

Bivand, R. & N. Lewin-Koh (2016). *maptools: Tools for Reading and Handling Spatial Objects*. R package version 0.8-39. URL: https://CRAN.R-project.org/package=maptools.

Bivand, R. & C. Rundel (2016). *rgeos: Interface to Geometry Engine - Open Source (GEOS)*. R package version 0.3-20. URL: https://CRAN.R-project.org/package=rgeos.

REFERENCES II

Boersma, P. & D. Weenink (2016). *Praat: doing phonetics by computer*. URL: http://www.fon.hum.uva.nl/praat/.

Calenge, C. (2006). "The package adehabitat for the R software: tool for the analysis of space and habitat use by animals". *Ecological Modelling* 197, 1035. DOI: 10.1016/j.ecolmodel.2006.03.017.

Campmany, E. (2004). "Diferències fonològiques entre diversos estils de parla al català central septentrional". PhD thesis. Universitat de Barcelona. URL: http://hdl.handle.net/2445/41617.

Colomina, J. (1999). *Dialectologia catalana: introducció i guia bibliogràfica*. Alacant: Universitat d'Alacant. Secretariat de Publicacions i Departament de Filologia Catalana.

Comas, J. (1970). "Caracterització del parlar de Pau: Alt Empordà". BA thesis. Universitat de Barcelona.

Coromines, J. (1953). "Algunes lleis fonètiques catalanes no observades fins ara". *Estudis Romànics* 3, 201–230. URL: http://revistes.iec.cat/index.php/ER/article/viewFile/37651/48276.

DiCanio, C., H. Nam, J. D. Amith, R. C. García, & D. H. Whalen (2015). "Vowel variability in elicited versus spontaneous speech: Evidence from Mixtec". *Journal of Phonetics* 48, 45–59. DOI: 10.1016/j.wocn.2014.10.003.

Dinkin, A. J. (2011). "Weakening resistance: Progress toward the low back merger in New York State". *Language Variation and Change* 23.03, 315–345. DOI: 10.1017/S0954394511000147.

REFERENCES III

Docherty, G., S. Gonzalez, & N. Mitchell (2015). “Static vs Dynamic Perspectives on the Realization of Vowel Nucleii in West Australian English”. In: *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*. Ed. by T. S. C. for ICPhS 2015. University of Glasgow. URL: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0956.pdf.

Dorca, J. (2008). *El parlar de Collsacabra: aproximació i assaig de descripció*. Barcelona: Institut d’Estudis Catalans, 117.

Dorca, J. (2010). “Aproximació geolingüística als parlars de les Guilleries: notes de presentació”. *Quaderns de la Selva* 22, 147–160. URL: http://www.raco.cat/index.php/QuadernsSelva/article/view/225276/334739.

Evanini, K. (2009). “The permeability of dialect boundaries: A case study of the region surrounding Erie, Pennsylvania”. PhD thesis. University of Pennsylvania. URL: http://repository.upenn.edu/edissertations/86.

Fruehwald, J. (2010). “SS ANOVA”. URL: http://www.ling.upenn.edu/~joseff/papers/fruehwald_ssanova.pdf.

Gómez, G. (2011). “Gramàtica del català rossellonès”. PhD thesis. Universitat Autònoma de Barcelona. URL: http://hdl.handle.net/10803/96175.

Gu, C. (2014). “Smoothing Spline ANOVA Models: R Package gss”. *Journal of Statistical Software* 58.5. DOI: 10.1007/978-0-387-98135-2.

REFERENCES IV

Gulsoy, J. (1993). *Estudis de gramàtica històrica*. València: Publicacions de l'Abadia de Montserrat.

Hay, J., K. K. Drager, & P. Warren (2009). "Careful Who You Talk to: An Effect of Experimenter Identity on the Production of the NEAR/SQUARE Merger in New Zealand English". *Australian Journal of Linguistics* 29.2, 269–285. DOI: 10.1080/07268600902823128.

Herrick, D. (2003). "An Acoustic analysis of phonological vowel reduction in six varieties of catalan". PhD thesis. Santa Cruz University of California.

Herrick, D. (2006). "Mid Vowels and Schwa in Eastern Catalan: five non-Barcelona Dialects". In: *New Perspectives on Romance Linguistics. Volume II: Phonetics, Phonology and Dialectology. Selected papers from the 35th Linguistic Symposium on Romance Languages (LSRL), Austin, Texas, February 2005*. Ed. by J.-P. Y. Montreuil. Amsterdam: John Benjamins. Chap. 8, 113–126.

IEC (2009). "2. Les vocals". In: *Gramàtica de la llengua catalana (esborrany provisional)*, 39–49. URL: http://www.iecat.net/institucio/seccions/filologica/gramatica/.

Johnson, D. E. (2015). "Quantifying overlap with Bhattacharyya's affinity (and other measures!)" Communicació presentada a NWAV44. URL: https://danielezrajohnson.shinyapps.io/nwav_44/.

Julià, J. (1986). "Algunes notes sobre el parlar de Cornellà de Terri". In: *Cornellà de Terri: pregons i estudis*. Cornellà de Terri: Ajuntament de Cornellà de Terri, 129–139.

REFERENCES V

Kahle, D. & H. Wickham (2013). “ggmap: Spatial Visualization with ggplot2”. *The R Journal* 5.1, 144–161. URL: http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf.

Kendall, T. & E. R. Thomas (2014). *vowels: Vowel Manipulation, Normalization, and Plotting*. R package version 1.2-1. URL: https://CRAN.R-project.org/package=vowels.

Labov, W. (1994). *Principles of Linguistic Change. Internal Factors*. Malden, MA: Wiley-Blackwell.

Luna, X. (1995). “Vocals tòniques a l’Empordà”. *Annals de l’Institut d’Estudis Empordanesos* 28, 266–303. DOI: 10.2436/93194.

McCloy, D. & A. McGrath (2014). *praat-semiauto: Praat scripts for streamlining manual measurements in acoustic analysis*. URL: https://github.com/drammock/praat-semiauto.

Monturiol, J. & E. Domínguez (2002). *El parlar de la Garrotxa*. 2nd ed. Olot: Ràdio Olot.

Nycz, J. R. & P. D. Decker (2006). *A New Way of Analyzing Vowels: Comparing Formant Contours Using Smoothing Spline ANOVA*. URL: http://research.library.mun.ca/655/.

Nycz, J. R. & L. Hall-Lew (2013). “Best practices in measuring vowel merger”. In: *Proceedings of Meetings on Acoustics*. Vol. 20. San Francisco, 060008. DOI: 10.1121/1.4894063.

Puig, P. (1990). *La Comarca del Baix Montseny: estudi dialectològic i sociolingüístic de la frontera comarcal entre el Vallès Oriental i la Selva: la suposada comarca del Baix Montseny*. Sant Celoni: Impr. Bilbeny.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

REFERENCES VI

Recasens, D. (1996). *Fonètica descriptiva del català: assaig de caracterització de la pronúncia del vocalisme i consonantisme del català al segle XX*. Barcelona: Institut d'Estudis Catalans.

Recasens, D. & A. Espinosa (2009). "Dispersion and variability in Catalan five and six peripheral vowel systems". *Speech Communication* 51.3, 240–258. DOI: 10.1016/j.specom.2008.09.002.

Sala, E. (1983). *El parlar de Cadaqués*. Girona: Diputació, 108.

Santos Baquero, O. (2016). *ggsn: North Symbols and Scale Bars for Maps Created with 'ggplot2' or 'ggmap'*. R package version 0.3.1. URL: https://CRAN.R-project.org/package=ggsn.

Wassink, A. B. & C. Koops (2013). "Quantifying and Interpreting Vowel Formant Trajectory Information". Communicació presentada a Best Practices in Sociophonetics Workshop, NWAV 42. URL: http://faculty.washington.edu/wassink/NWAV2013/2013-Wassink-Koops-slides-7.pdf.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. URL: http://had.co.nz/ggplot2/book.

Huy chexw!

Eva Bosch-Roura – eva.bosch.roura@ub.edu
Presentation available at www.ub.edu/GEVAD

Project FFI2013-46987-C3-1-P (MINECO)
NWAV 45 Student Travel Award