

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

APRIL 29 - MAY 1, 2021

LINGUISTIC
SYMPOSIUM ON
ROMANCE
LANGUAGES

51

Sponsored by:

The School of Literatures, Cultures, & Linguistics

The Departments of Linguistics, Spanish & Portuguese, and French & Italian

The Lemann Center for Brazilian Studies

<https://publish.illinois.edu/lsl51atUIUC>



Welcome to the 51st Linguistics Symposium on Romance Languages (LSRL 51) hosted by the University of Illinois at Urbana-Champaign. In this book, you will find an abstract for each of the 4 plenary talks, 42 oral presentations, and 13 poster presentations. This year's selection of scholars truly highlights the international scope of LSRL by presenting from countries which collectively span an impressive portion of the globe: Argentina, Austria, Belgium, Brazil, Canada, Chile, France, Germany, Italy, Mexico, Peru, Spain, Switzerland, the UK, and the US. LSRL's global reach stems from years of top-notch research in all major areas of linguistics, the impact of which typically extends beyond the Romance languages themselves and bears on key issues of interest to the wider field of Linguistics. The consistency of such impactful research owes itself to the community of scholars who graciously offer up their time and expertise to evaluate multiple abstracts for the conference. This year we received over 100 abstracts, and we would like to thank our scientific committee for ensuring yet another round of innovative and thought-provoking presentations. We would also like to thank our sponsors for their support in making LSRL 51 a reality. Finally, a sincere thank you to all presenters, without whom, there would be no LSRL.

The LSRL 51 organizing committee:

Jonathan E. MacDonald

Zsuzsanna Fagyal

Ander Beristain

Robin Turner

April 2021





The LSRL 51 Organizing Committee would like to extend a special thanks to our scientific committee and volunteers who lent their time and expertise to help make this year's program a success.

Scientific committee: Mireille Tremblay, Jeffrey Steele, Leonardo Savoia, Yves Roberge, Gemma Rigau Oliver, Rajiv Rao, Fabio Montermini, Egle Mocciaro, Natalia Mazzaro, Diane Massam, Rita Manzini, John Lipski, Virginia Hill, Francisco Moreno Fernández, Richard Cameron, Andrea Calabrese, Laura Bafile, Angeliki Athanasopoulos, Evangelia Adamou, Raquel Fernández Fuertes, Fernando Martínez-Gil, Francesco Romano, Lydia White, Marina Vigarío, Dominique Sportiche, Petra Sleeman, Liliana Sánchez, Nuria Sagarra, Edward Rubin, Johan Rorryck, Rebecca Ronquest, Dan Olson, Jairo Nunes, Jaume Mateu, Randall Gess, Charlotte Galves, Gorka Elordieta, Carmen Dobrovie-Sorin, Isabelle Charnavel, Mónica Cabrera, Melissa Bowles, Aurora Bel, Jennifer Austin, J. Clancy Clements, Anna María Escobar, Ana Carvalho, Juana Licerias, Cinzia Russi, Eulàlia Bonet, Elena Valenzuela, Irene Vogel, Barbara Vance, Valerie Saugera, Marcos Rohena-Madrado, Francesc Roca, Lori Repetti, Antxon Olarrea, Erin O'Rourke, Julia Herschensohn, Tim Gupton, Carolina González, Anna Gavarró, Diego Pescarini, Olga Fernández Soriano, Manuel Díaz-Campos, Sonia Cyrino, Sonia Colina, Laura Colantoni, Anna Cardinaletti, Travis Bradley, Judy Bernstein, Marc Authier, Deborah Arteaga, Alfredo Garcia-Pardo, Scott Alvord, Peggy Renwick, Ana-Maria Martins, Julio Villa-Garcia, Myriam Uribe-Etxebarria, Miguel Simonet, Ryan Shosted, Cristina Schmitt, Andrés Saab, Lara Reglero, Acrisio Pires, Anna Pineda, Francisco Ordóñez, Silvina Montrul, Mairi McLaughlin, Luis López, Manuel Leonetti, José Ignacio Hualde, Ion Giurgea, Angel Gallego, Rebecca Foote, Ian Roberts, Itxaso Rodriguez, Karen Zagona, Antonio Fábregas, Maria Teresa Espinal, Justin Davidson, Roberta D'Alessandro, Maria Cristina Cuervo, José Camacho, Barbara Bullock, Hélène Blondeau, Julie Auger, Karlos Arregi, Grant Armstrong, Patricia Amaral, Bryan Donaldson, Gyula Zsombok, Michelle Sheehan, Amanda Dalola, Grant Goodall, Erik Willis, Timothy Face

Session Chairs: Patricia Amaral, Barbara Bullock, Salvatore Callesano, Justin Davidson, Suzanne Franks, José Ignacio Hualde, Jill Jegerski, Douglas Kibbee, Matthew Maddox, Silvina Montrul, Erin O'Rourke, Chilin Shih, Imanol Suárez-Palma, Aida Talic, Yan Tang, James Yoon

Conference volunteers: Lorena Alarcon, Caio Albernaz, Étienne André, Willi Asamoah, Izaro Bedialauneta, Sarah Clark, Joshua Dees, Martine Gallardo, Kacie Gastanaga, Walther Glodstaf, Lindsey Graham, Charlotte Prieu, Yinglun Sun, Ari Theodoropoulos, Kara Yarrington, Katie VanDyne



Linguistic Symposium on Romance Languages 51, April 29 – May 1, 2021. UIUC

Time zone guide (LSRL times are in CST)

PST (Los Angeles) 7:00	CST (Chicago) 9:00	EST (New York City) 10:00	BST (Brasilia) 11:00	GMT (London) 15:00	CET (Paris, Madrid) 16:00
------------------------------	--------------------------	---------------------------------	----------------------------	--------------------------	---------------------------------

DAY 1	Thursday, April 29th	
	Room 1 (R1)	Room 2 (R2)
8:30-9:00 CST	R1: Opening remarks Jonathan MacDonald (Chair of LSRL51) L. Elena Delgado (Director of the School of Literatures, Cultures, and Linguistics) Mariselle Meléndez, James Yoon, Zsuzsanna Fagyal (Heads of sponsoring departments)	
9:00-10:30 CST	R1: Clitics, doubling, and person restrictions Session chair: Aida Talić (UIUC) Zoom host: Walther Glödstaf 1. “Distinguishing between the accounts of the A/A’-distinction: The view from Argentinian Spanish Clitic Doubling” (Fong) h:00 2. “Agree, Expletives, and Person restrictions in Bolognese” (Rubin) h:30 3. “Grammatical encoding of agency in middle contexts in Ibero-Romance” (Suárez-Palma) h:00	R2: Vowels, consonants, and words in speech Session chair: José Ignacio Hualde (UIUC) Zoom host: Izaro Bedialauneta-Txurruka 1. “Sound change and rhythm in Altiplateau Mexican Spanish” (Marchini & Ramsammy) h:00 2. “The interaction between post-aspiration and stress in Sevillian Spanish” (Gilbert) h:30 3. “Loanword adaptation in Portuguese-American speech” (Bland) h:00
10:30-11:30 CST	R1: Poster Session 1 Zoom host: Robin Turner 1. “The acquisition of stop consonants by Chinese learners of Spanish: A longitudinal and cross-sectional study” (Bravo Diaz) 2. “The grammar and use of the epistemic future: A comparison between two Italo-Romance speech communities” (Silvestri) 3. “Simple clitics in Italo-Romance: Dialectal variation and phrasal phonology” (King) 4. “Clausal determiner as inherent case: Evidence from Spanish” (Gallego & Ruiz Alonso) 5. “The influence of language typology in multilinguals’ language learning strategies” (Sáez Fajardo) 6. “The lexis of Sardinian within the Romance scene: A minority-language case with valuable insights into historical linguistics” (Gandarillas)	<i>GatherTown</i> networking session open concurrently
11:30-12:30 CST	Lunch break (<i>GatherTown</i> remains open if you want to grab a bite together with other attendees!)	

Linguistic Symposium on Romance Languages 51, April 29 – May 1, 2021. UIUC

Time zone guide (LSRL times are in CST)

PST (Los Angeles) 7:00	CST (Chicago) 9:00	EST (New York City) 10:00	BST (Brasilia) 11:00	GMT (London) 15:00	CET (Paris, Madrid) 16:00
------------------------------	--------------------------	---------------------------------	----------------------------	--------------------------	---------------------------------

DAY 1	R1: Movement in multiple domains Session chair: James Yoon (UIUC) Zoom host: Martine Gallardo	R2: Vowels and acoustic cues Session chair: Yan Tang (UIUC) Zoom host: Lindsey Graham
12:30-2:00 CST	1. “Long passives of causatives and perception verbs in Italian: Implications for phase theory” (Sheehan & Casalicchio) h:30 2. “Rethinking Romance vocatives: Mapping address inversion in the nominal edge” (Corr) h:00 3. “A unified analysis of failed operator movement in Spanish in terms of anti-locality” (Petersen O’Farrill) h:30	1. “Target vowel asymmetry in Brazilian Veneto metaphony” (Garcia & Guzzo) h:30 2. “Sociophonetic variation in the Spanish production of stressed and unstressed /e, o/ by Spanish-Galician bilinguals” (de la Fuente Iglesias) h:00 3. “Do we really need context? The role of acoustic cues in idiomatic disambiguation” (Lares) h:30
2:00-2:10 CST	Coffee break	
2:10-3:40 CST	R1: Studying the past and present in corpora Session chair: Barbara Bullock (UT Austin) Zoom host: Lorena Alarcon 1. “The phonetic reduction of Spanish stressed vowels in spontaneous speech” (Perry, Kelley, & Tucker) h:10 2. “How fast did Cicero speak? The speech rate of Classical Latin verdun its Romance descendants” (Stelzer) h:40 3. “The Pluperfect First hypothesis: The compound pluperfect as a necessary precondition of the perfect-to-perfective shift” (Balla-Johnson) h:10	R2: Acquisition: L2 & Bilingual Session chair: Silvina Montrul (UIUC) Zoom host: Kacie Gastanaga 1. “Inhibition vs Lexicon in implicature generation among bilingual speakers” (Lingwall Odio & Grinstead) h:10 2. “Forgotten allophones and their gradient friends: Complexity in the acquisition of Spanish voiced stops” (Campos-Astorkiza, DeLeon, Locascio, & Sullivan) h:40 3. “How frequent are these verbs? The role of lexical frequency in children’s acquisition of verb morphology at different ages of acquisition” (Goldin, López-Otero, & Hur) h:10
3:40-3:45 CST	Coffee break	
3:45-4:45 CST	R1: Plenary speaker (Ryan Shosted) Talk title: <i>The Oropharyngeal Character of Nasal Vowels: Raising in Brazilian Portuguese and Romanian</i> Session chair: Chilin Shih (UIUC) Zom host: Ander Beristain	

--END OF DAY 1--

Time zone guide (LSRL times are in CST)

PST (Los Angeles) 7:00	CST (Chicago) 9:00	EST (New York City) 10:00	BST (Brasilia) 11:00	GMT (London) 15:00	CET (Paris, Madrid) 16:00
------------------------------	--------------------------	---------------------------------	----------------------------	--------------------------	---------------------------------

DAY 2	Friday, April 30th	
	Room 1 (R1)	Room 2 (R2)
9:00 – 10:00 CST	R1: Plenary speaker (Heather Burnett) Talk title: <i>Political Dimensions of Écriture Inclusive in Parisian Universities</i> . Session chair: Douglas Kibbee (UIUC) Zoom host: Ander Beristain	
10:00- 11:30 CST	R1: Argument and Event structure Session chair: Imanol Suárez-Palma (UF) Zoom host: Aristides Theodoropoulos 1. “The aspectual properties of <no + event deverbal nominal> in Spanish” (Ros García) h:00 2. “On the argument and event structure of the Catalan suffix <i>-ejar</i> ” (Acedo-Matellán & Gibert-Sotelo) h:30 3. “Adjectival passives with a progressive reading” (Gibert-Sotelo & Marín) h:00	R2: Prosody and its interfaces Session chair: Erin O’Rourke (UA) Zoom host: Williams Asamoah 1. “Acquiring pragmatic prosody: Evidence from 3-to-4-year-old Catalan speaking children” (Pronina, Hübscher, Vilà-Giménez, & Prieto) h:00 2. “Prosody and visual perception on the processing of gapping sentences in Brazilian Portuguese” (Oliveira da Silva & Alves Fonseca) h:30 3. “Prosodic constraints on wh-extraction from infinitival clauses” (Muñoz Pérez, Verdecchia, & Carranza) h:00
11:30- 12:30 CST	R1: Business meeting (lunch welcomed)	Lunch break
12:30- 1:30 CST	R1: Poster Session 2 Zoom host: Robin Turner 1. “On the development of negative <i>en DET vida</i> in Spanish” (Howe & Cabezas Zapata) 2. “An acoustic account of rhotic assibilation in Chihuahua Spanish” (Mazzaro) 3. “The active/passive alternation in French: A statistical approach to corpus data” (Da Cunha & Abeillé) 4. “DOM co-occurrence restrictions and their repair strategies: Evidence from Galician and Romanian” (Gravely & Irimia) 5. “Singular, plural, modal” (Mihoc) 6. “The role of non-distinctive features in speech perception: Voicing assimilation and the perception of contrasts in Spanish” (Campos-Astorkiza) 7. “Intransitive causatives in English and Romance” (Mangialavori Rasia & Ausensi) 8. “Constraints on V-IO DO and V-S-DO in Romance. IOs as “subjects”: A Labeling Theory approach” (Gallego)	<i>GatherTown</i> networking session open concurrently

Time zone guide (LSRL times are in CST)

PST (Los Angeles) 7:00	CST (Chicago) 9:00	EST (New York City) 10:00	BST (Brasilia) 11:00	GMT (London) 15:00	CET (Paris, Madrid) 16:00
------------------------------	--------------------------	---------------------------------	----------------------------	--------------------------	---------------------------------

DAY 2	R1: Subjects: Null and Overt Session chair: Matthew Maddox (UNL) Zoom host: Caio Albernaz Siqueira		R2: Strengthening and Weakening Session chair: Justin Davidson (UC Berkeley) Zoom host: Ander Beristain	
	1. “Which factors affect the production of Overt pronouns in Catalan?” (Mayol) h:30	2. “Structure preservation and contact effects: Subjects in Brazilian Veneto” (Guzzo & Franken) h:00	1. “Strengthening and weakening at domain edges: Evidence from Spanish” (Ramsammy & King) h:30	2. “Experimental evidence of palatal approximant strengthening in Medellín Spanish” (Leslie) h:00
1:30-3:00 CST	3. “PRO as the subject of generic null impersonals in Brazilian Portuguese” (Carvalho) h:30	3. “Lenition and contrastiveness: Word-final stops in Catalan” (Hualde & Zhang) h:30		
3:00-3:10 CST	Coffee break			
3:10-4:40 CST	R1: Syntax and Semantics Session chair: Patrícia Amaral (IU) Zoom host: Joshua Dees		R2: Segments & Syllable structure Session chair: Suzanne Franks (UIUC) Zoom host: Jennifer Zhang	
	1. “Los Pedros y los Picapiedra: Resolving the semantic ambiguity of family names” (Jambrović) h:10	2. “Optional <i>se</i> constructions and flavours of applicatives in Spanish” (Arunachalam & Martin) h:40	1. “Accounting for participants’ best linguistic behavior: Vowel nasalization, nasal consonant weakening and hypercorrect /s/” (Bongiovanni) h:10	2. “Do Spanish codas completely resyllabify? A look at Mexican Spanish” (Repiso-Puigdelliura) h:40
	3. “Two types of causatives and DOM” (Goncharov & Irimia) h:10		3. “Nasal airflow evidence for complete resyllabification in Spanish” (Beristain) h:10	
4:40-4:45 CST	Coffee break			
4:45 – 5:45 CST	R1: Plenary speaker (Jairo Nunes) Talk title: <i>Crosscategorical φ-deficiency in Brazilian Portuguese.</i> Session chair: Jonathan MacDonald Zoom host: Katie VanDyne			

--END OF DAY 2--

Time zone guide (LSRL times are in CST)

PST (Los Angeles) 7:00	CST (Chicago) 9:00	EST (New York City) 10:00	BST (Brasilia) 11:00	GMT (London) 15:00	CET (Paris, Madrid) 16:00
------------------------------	--------------------------	---------------------------------	----------------------------	--------------------------	---------------------------------

DAY 3	Saturday, May 1st	
	Room 1 (R1)	Room 2 (R2)
	R1: Experimental Syntax and Semantics Session chair: Jill Jegerski (UIUC) Zoom host: Étienne André	R2: Variation and innovation in corpora Session chair: Salvatore Callesano (UIUC) Zoom host: Charlotte Prieu
9:00-10:30 CST	1. “The role of C and T in the THAT-trace Effect: Evidence from Code-switching” (Hoot & Ebert) h:00	1. “Translator style as a sociolinguistic variable: Variation in news translation from English to Romance” (McLaughlin & Davidson) h:00
	2. “The online comprehension of re-complementation in Spanish: Competing models and individual differences in working memory capacity considered” (Frank) h:30	2. “Not so inclusive: The debate on inclusive-writing on French Wikipedia” (Zombok) h:30
	3. “On the acquisition of psych predicates: When intervention makes ‘Subject’ extraction harder” (Mateu) h:00	3. “Lexical and grammatical semantic shifts: A distributional analysis of ‘embargo’ (Amaral, Hu, & Kübler) h:00
10:30 – 11:30 CST	R1: Plenary speaker (Anna María Escobar) Talk title: <i>Language contact as a trigger for language change: Grammaticalization pathways of the Spanish Present Perfect.</i> Session chair: Zsuzsanna Fagyal (UIUC) Zoom host: Robin Turner	
11:30 – 12:30 CST	R1: LaTeX Workshop (Yinglun Sun)	
12:30-1:00 CST	R1: Closing remarks	

--END OF LSRL 51--

LSRL 51 Organizing Committee:
Dr. Jonathan MacDonald (Chair)
Dr. Zsuzsanna Fagyal
Ander Beristain
Robin Turner

Political Dimensions of Écriture Inclusive in Parisian Universities
(joint work with Céline Pozniak)

Heather Burnett
Université de Paris 7-Denis Diderot

Écriture inclusive (EI) (i.e. shortening expressions like "étudiants et étudiantes" to "étudiant·e·s", "étudiant(e)s" etc.) has long been the topic of public debates in France. These debates have become more intense in recent years, often focusing on the higher education system and culminating in the formulation of three separate laws banning it for public administration. In this paper, we investigate the foundations of these conflicts through a large quantitative corpus study of the (non)use of EI in Parisian undergraduate brochures. Our results suggest that Parisian university professors use EI not only to ensure gender neutral reference, but also as a tool to construct their political identities. We show that both the use of EI and its particular forms are conditioned by how brochure writers position themselves on non-gender related issues within the French university's political landscape, which explains how conflicts surrounding a linguistic practice have become understood as conflicts about larger issues in French society. Our paper thus provides new information to be taken into account in the formulation and promotion of non-sexist language policies, and sheds light on how feminist linguistic activism and its opposition are deeply intertwined with other kinds of social activism in present-day France.





***Language contact as a trigger for language change:
Grammaticalization pathways of the Spanish Present Perfect***

Anna María Escobar
University of Illinois at Urbana-Champaign

The semantic trajectories of contact-induced grammatical change are still not well understood (cf. Heine & Kuteva 2005; Kuteva et al. 2019). This presentation examines the development of the Spanish Present Perfect (PP) in an ex-colonial region where Spanish is in contact with Indigenous languages and argues for the inclusion of linguistic factors connected to subjectivity and information structure in the study of the PP, alongside the traditional temporal and aspectual factors. Perfects in the world's languages derive from three main sources (BE/HAVE, COME, FINISH, Bybee, Perkins & Pagliuca 1994), but HAVE perfects (e.g. Spanish) are considered the least common pattern (WALS, Dahl & Velupillai 2013; Drinka 2017). Bybee et al. (1994) posit that only the stative source branches into two distinct grammaticalization pathways of the PP: the temporal (towards past/perfective) or the evidential (leading to (in)direct evidential). Although Romance languages are described as following the temporal pathway (Harris 1982; Squartini & Bertinetto 2000, de Acosta 2011), there are exceptions. Dacia-Romanian (Drinka 2017), Judeo-Spanish (Varol 2006), and Andean Spanish varieties (Escobar & Crespo 2020) are argued to follow the evidential path, in each case triggered by contact with a language or languages of families (Turkic or Quechua-Aymaran) that have evidential markers. Through grammatical analyses of data from collected interviews, the effect of subjectivity and information structure on the grammaticalization pathway of the PP in Andean Spanish varieties is established.

References:

- Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago and London: The University of Chicago Press.
- Dahl, Östen & Viveka Velupillai. 2013. The Perfect. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Plank Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/68>, Accessed on 2021-03-28.)
- de Acosta, Diego. 2011 Rethinking the genesis of the Romance periphrastic perfect. *Diachronica* 28 (2), 143-185.
- Drinka, Bridget. 2017. *Language contact in Europe: The Periphrastic Perfect through history*. Cambridge: Cambridge University Press.
- Escobar, Anna María & Claudia Crespo. 2020. La gramaticalización de la subjetividad en el español andino: el pretérito perfecto compuesto con valor evidencial. In: Andrade Ciudad, Luis & Sessarego, Sandro (eds.) *Los castellanos del Perú: historia variación y contacto lingüístico*. London & New York: Routledge. Pp. 156-205.
- Harris, Martin. 1982. The 'past simple' and 'present perfect' in Romance. In: Harris, Martin & Nigel, Vincent (eds.) *Studies in the Romance Verb*. London: Croom Helm. Pp. 42-70.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Squartini, Mario & Pier Marco Bertinetto. 2000. The simple and compound past in Romance languages. In: Dahl, Östen (ed.) *Tense and aspect in the languages of Europe*. Berlin: Mouton de Gruyter. Pp. 403-439.
- (Bornes-)Varol, Marie-Christine. 2006. El judeoespañol en contacto: el ejemplo de Turquía. *Revista Internacional de Lingüística Iberoamericana* IV, 2 (18), 99-114.

Crosscategorical ϕ -deficiency in Brazilian Portuguese

Jairo Nunes
Universidade de São Paulo

A cluster of diachronic changes in Brazilian Portuguese (BP) documented in the 19th and 20th centuries has set it aside within Romance. BP has lost its third person accusative clitics (e.g. Nunes 1993, Cyrino 1997, Galves 2001), its third person dative clitics (e.g. Torres Morais and Berlinck 2006), its third person possessive pronouns (e.g. Negrão and Müller 1996, Cerqueira 1996), and external possessor constructions (e.g. Barros 2006, Torres Morais and Salles 2016); its directional verbs came to select the preposition *em* ‘in’ instead of a ‘to’ (e.g. Wiedemer 2013); and its null subjects (e.g. Duarte 1995, Rodrigues 2004, Ferreira 2009) and null possessors (e.g. Rodrigues 2004, Floripi and Nunes 2009) have become considerably restricted. In addition, BP came to allow “topic-subject” constructions where an apparent topic functions as a subject in controlling verbal agreement (e.g. Pontes 1987, Nunes 2017, Kato and Ordóñez 2019), as well as hyper-raising constructions, where a subject undergoes A-movement from a finite clause or inflected infinitival, triggering verbal agreement in both the embedding and the embedded clause (e.g. Ferreira 2009, Martins and Nunes 2010, Nunes 2020). In this talk, I argue that these seemingly independent changes are actually related to a general process of underspecification of ϕ -features on phase heads and pronominal elements in BP. I also show that this process also affects the acquisition of third person accusative clitics in formal BP via schooling, leading to an exceptional pattern of clitic placement (e.g. Nunes 2015).



References:

- Barros, Evângela. 1996. Construções de posse com clítico no PB: Percurso diacrônico. Doctoral dissertation, Universidade Federal de Minas Gerais.
- Cerqueira, Vicente Cruz. 1996. A sintaxe do possessivo no Português Brasileiro. Doctoral Dissertation, Universidade Estadual de Campinas.
- Cyrino, Sonia. 1997. O objeto nulo no português do Brasil: Um Estudo Sintático-Diacrônico. Londrina: Editora UEL.
- Duarte, Maria Eugênia. 1995. A perda do Princípio “Evite Pronome” no português brasileiro. Doctoral dissertation, Universidade Estadual de Campinas.
- Ferreira, Marcelo. 2009. Null subjects and finite control in Brazilian Portuguese. In J. Nunes (ed.): *Minimalist essays on Brazilian Portuguese Syntax*, 17-49. Amsterdam/Philadelphia: John Benjamins.
- Floripi, Simone and Jairo Nunes. 2009. Movement and resumption in null possessor constructions in Brazilian Portuguese. In J. Nunes (ed.): *Minimalist essays on Brazilian Portuguese syntax*, ed. by Jairo Nunes, 51-68. Amsterdam/Philadelphia: John Benjamins.
- Galves, Charlotte. 2001. Ensaio sobre as gramáticas do português. Campinas: UNICAMP.
- Kato, Mary and Francisco Ordóñez 2019. Topic subjects in Brazilian Portuguese and clitic left dislocation in Dominican Spanish. *Syntax* 22: 229-247.
- Martins, Ana Maria and Jairo Nunes. 2010. Apparent hyper-raising in Brazilian Portuguese: Agreement with topics across a finite CP. In P. Panagiotidis (ed.): *The complementiser phase: Subjects and wh-dependencies*. Oxford: Oxford University Press.
- Negrão, Esmeralda and Ana Müller. 1996. As mudanças no sistema pronominal do português brasileiro: substituição ou especialização de formas. *D.E.L.T.A.* 12:125-152.
- Nunes, Jairo. 1993. Direção de cliticização, objeto nulo e pronome tônico na posição de objeto em português brasileiro”. In I. Roberts and M. A. Kato (eds.): *Português brasileiro: Uma viagem diacrônica*, 207-222. Campinas: Editora da UNICAMP.
- Nunes, Jairo. 2015. De clítico a concordância: O caso dos acusativos de terceira pessoa em português brasileiro. *Cadernos de Estudos Linguísticos* 57: 61-84.
- Nunes, Jairo. 2017. Circumventing ϕ -minimality: On some unorthodox cases of A-movement in Brazilian Portuguese. In R. E. V. Lopes, J. Avelar, and S. Cyrino (eds.): *Romance Languages and Linguistic Theory 12: Selected papers from the 45th Linguistic Symposium on Romance Languages*. Amsterdam/Philadelphia: John Benjamins.
- Nunes, Jairo. 2020. Hiperálçamento em português brasileiro. *Cuadernos de la ALFAL* 14.2: 199-227.
- Pontes, Eunice. 1987. O tópico no português do Brasil. Campinas: Pontes.
- Rodrigues, Cilene. 2004. Impoverished Morphology and A-movement out of Case Domains. Doctoral dissertation, University of Maryland.
- Torres Morais, Maria Aparecida and Rosane Berlinck. 2006. A caracterização do objeto indireto no português: Aspectos sincrônicos e diacrônicos. In R. V. Mattos e Silva and T. Lobo (eds.): *Para a história do português brasileiro: Novos dados, novas análises*, vol. VI.1, 73-106. Salvador, Editora da UFBA.
- Torres Morais, Maria Aparecida and Heloísa Salles. 2016. The external possessor construction in European Portuguese and Brazilian Portuguese. In M. A. Kato and F. Ordóñez (eds.): *The morphosyntax of Portuguese and Spanish in Latin America*, 204-235. Oxford, Oxford University Press.
- Wiedemer, Marcos. 2013. Variação e gramaticalização no uso de preposições em contextos de verbo de movimento no português brasileiro. Doctoral dissertation, Universidade Estadual Paulista, São José do Rio Preto.



The Oropharyngeal Character of Nasal Vowels: Raising in Brazilian Portuguese and Romanian

Ryan Shosted

University of Illinois at Urbana-Champaign

Nasalization presents a particularly interesting opportunity for sound change to be “drawn from a pool of synchronic variation” (Ohala 1989). Vowels under the influence of velopharyngeal opening (whether phonemically nasal or phonetically nasalized) are well-known for their acoustic complexity; the listener’s response to this challenge is well-studied (Hawkins 1985, among others). Velopharyngeal opening increases the complexity of the vocal tract resonator, adding extra formants and anti-formants to the spectrum associated with relatively simpler oral vowels. The bandwidths of the spectral peaks also tend to widen with greater sound absorption. The sinuses may add still more resonant complexity. It has long been acknowledged that this presents challenges to the analysis of nasal and nasalized vowels and many creative solutions have recently been proposed to deal with the detection of nasality (e.g. Styler 2017; Carignan 2021).

Another question that has received increasing attention in the last decade is this: what is the oropharyngeal character of nasal vowels? Analyzing the spectrum of nasal vowels inherently yields an ambiguous result. Are the formant values detected indicative of the true position and degree of constriction in the vocal tract or of the degree of velopharyngeal opening? In truth, the formant values represent both, so special care must be taken to situate oral and nasal vowels in the same frequency space. To clarify: when we transcribe the symbol [ã], for example, what do we mean? There are two possibilities: (1) That the vowel has all the oropharyngeal characteristics of oral [a] with velopharyngeal opening; and (2) that a vowel of some unknown oropharyngeal quality sounds like [ã], because it has experienced velopharyngeal opening. The first vowel has a motor plan equivalent to [a] (plus velopharyngeal opening) and may therefore experience the same kinds of synchronic and diachronic pressures as its oral counterpart. But the second is a different vowel that merely sounds like [a].

Arai (2005) was the first to show that oral and nasal vowels manifest different tongue positions. Later studies (e.g., Engwall et al. 2006; Carignan et al. 2011; Shosted et al. 2012; Barlaz et al. 2018; Cler et al. 2021, among others) confirmed this and presented increasing evidence that oropharyngeal shape in oral and nasal vowels can differ significantly across a variety of languages, including French and Brazilian Portuguese. What ramifications does this have for our understanding of sound change in Romance?

In this talk, I will review the difficulties inherent in the phonetic study of nasalization; discuss findings on the oropharyngeal character of nasal vowels generally; and discuss vowel nasalization in two Romance languages, one for which I present articulatory data (Brazilian Portuguese) and one for which I present philological data (Romanian). In both languages, the effects of velopharyngeal opening have raised the etymological low vowel considerably. Indeed, the increase is so great that it hardly makes sense to call these vowels “low” any longer. In both cases, an influential contact language possessed a central vowel: the high central vowel in the case of Tupinambá / Língua Geral and a mid-to-high central vowel in the case of Old Church Slavonic / Bulgarian. I argue that the natural tendency to raise low vowels that experience velopharyngeal opening, coupled with the influence of a high central vowel in a contact language, brought about the extreme nasal vowel raising observed in these two Romance languages.

References:

- Barlaz, M., Shosted, R., Fu, M. Fu, and Sutton, B. 2018. Oropharyngeal articulation of phonemic and phonetic nasalization in Brazilian Portuguese. *J. Phon.* 71: 81–97.
- Carignan, C., Shosted, R., Shih, C., and Rong, P. 2011. Compensatory articulation in American English nasalized vowels. *J. Phon.* 39: 668–682.
- Carignan, C. 2021. A practical method of estimating the time-varying degree of vowel nasalization from acoustic features. *J. Acous. Soc. Am.* 149: 911–922.
- Hawkins, S., and Stevens, K. N. Acoustic and perceptual correlates of the non-nasal–nasal distinction for vowels. *J. Acoust. Soc. Am.* 77, 1560–1575, 1985.
- Ohala, J. J. 1989. Sound change is drawn from a pool of synchronic variation. In L. E. Breivik and E. H. Jahr (eds.), *Language Change*, pp. 173–198. Berlin: De Gruyter Mouton.
- Shosted, R., Carignan, C., and Rong, P. 2012. Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi. *J. Acous. Soc. Am.* 131(1): 455–465.
- Styler, W. 2017. On the acoustical features of vowel nasality in English and French. *Journal of the Acoustical Society of America* 142: 2469–2482.

On the argument and event structure of the Catalan suffix *-ejar*

Víctor Acedo-Matellán (*U. of Oxford*) & Elisabeth Gibert-Sotelo (*U. Rovira i Virgili*)

0. This paper explores the properties of the Catalan verbalizing suffix *-ejar*. Most recently studied by Ultra-Massuet & Castroviejo (2013), we believe that there are still interesting observations to make about its semantics and syntax. After presenting the rich variety of outputs that this suffix allows [1], we focus on two generalizations. The first one has to do with the consistent eventive nature of verbs derived with this suffix, even from adjectival bases that count as individual-level predicates, like colour or size adjectives [2]. The second one is the robust argument- and event-structure properties that these verbs show when they are built on proper names or adjectives [3]. Adopting a nanosyntactic perspective, we propose that this suffix is associated with an activity interpretation via its lexicalizing the subeventive structure of a caused process; the different nuances in event structure emerge from 1) its taking either a root or a stative projection as complement and 2) the interactions between the semantics of the root and that of the arguments [4].

1. *-ejar* attaches productively to bases of different categories: nouns (*fein-ejar* ‘do work’; cf. *feina* ‘work’), adjectives (*brut-ejar* ‘look dirty’; cf. *brut* ‘dirty’), verbs (*menj-ot-ejar* ‘eat little and with no appetite’; cf. *menjar* ‘eat’), adverbs (*ran-ejar* ‘be close to a certain threshold’; cf. *ran* ‘close to the base of something’), numerals (*seixant-ejar* ‘be around sixty’; cf. *seixanta* ‘sixty’), and even proper names, a category that can hardly be claimed to involve any eventive variable (*obam-ejar* ‘be or act like Obama’; cf. *Obama*). Besides, (1) produces both transitive (1) and intransitive (2) verbs. [CTILC]

(1) *Arribarem a migdia* (1) and *air-ejar* (2) verbs.
arrived.1PL at noon and air-ej-IPFV.1PL the house
‘We would arrive at noon and air the house.’

(2) Tornant, el glaç blanqu-*ej-a* pel camí. [CTILC]
come_back.GER the ice white-*ej*-PRS.3SG along the path
‘When we come back, the ice shines white along the path’.

2. Verbs with *-ejar* are eventive even when they take bases that are typical individual-level predicates, such as colour adjectives. This is evidenced when a verb like *groguejar* ‘show yellow’ is submitted to standard eventivity tests, like compatibility with perception reports (3a), locative modifiers (3b), and manner adverbials (3c) (Maienborn 2007):

(3) a. *Vèiem grogu-*ej-ar* una vinya.* [CTILC]
see.IPFV.1PL yellow-*ej*-INF a vine.
‘We would see a vine show yellow’.

b. El poble que grogu-*ej-a* *allà baix* amb el sol ponent. [CTILC]
the village that yellow-*ej*-PRS.3SG there down with the sun setting
‘The village showing yellow down there in the setting sun’.

c. Les pomposes raïmades grogu-*eg-en* *amb llustre de mel.* [CTILC]
the opulent grapes yellow-*ej*-PRS.3PL with brilliance of honey
‘The opulent grapes shine yellow with the brilliance of honey’.

Importantly, the eventiveness of *ejar*-verbs is independent of whether they are dynamic or not. Thus, *groguejar* ‘show yellow’ is a non-dynamic verb, in spite of its eventiveness, as its inability to license *do*-proforms (4a) or imperatives (4b) indicates (cf. Silvagni 2017:168ff):

(4) a. #El que fa el poble és grogu-*ej-ar*.
what does the village is yellow-*ej*-INF.
‘What the village does is to show yellow.’

b. #Grogu-*ej-a*!
yellow-*ej*-IMP

Lexical and Grammatical Semantic Shifts: A Distributional Analysis of *embargo*

Patrícia Amaral, Hai Hu and Sandra Kübler, *Indiana University*

Cross-linguistically, it is known that concessive connectives are complex, semantically derived expressions (König, 1988) resulting from regular mechanisms of syntactic and semantic change. However, it is hard to find measures of semantic relatedness and compositionality related to meaning that do not depend exclusively on intuition. This paper investigates the development of the Spanish concessive connective *sin embargo* ‘however’ (originally a PP ‘without obstacle or hindrance of’) and compares it to the noun *embargo*, originally ‘obstacle, impediment’. Using the tools of distributional semantics (Hamilton et al., 2016; Boleda, 2020), we examine the lexical change of *embargo* to its current meaning ‘prohibition of commercial relations’ and the grammatical change leading to the creation of *sin embargo*. Changes in lexical meaning are often the product of socio-historical factors and differ from changes resulting from grammaticalization (Traugott and Dasher, 2002; Campbell, 2007). We show that (i) word embeddings provide a fruitful methodology to investigate both types of meaning change and (ii) this methodology provides replicable ways to capture the compositionality of these expressions at early stages of their reanalysis.

In distributional semantics, word meaning is captured by the “company words keep” (Firth, 1957). Semantic change can thus be investigated through change in the word co-occurrence patterns, specifically by identifying the most similar words given by word embedding models at different points in time (Hamilton et al., 2016). In this paradigm, similar words are both synonyms/antonyms and words that are syntagmatically close. We traced the corpus neighbors of both the noun and the multi-word expression *sin embargo* in two time periods: for the Medieval/Classical period we used a set of Spanish chronicle texts, *Chronicles*, ranging from the 13th-16th c. (7 mio. words) from the *Digital Library of Old Spanish Texts* (<http://www.hispanicseminary.t&c/ac/index.htm>). For the contemporary data, we used the Spanish Billion Words corpus, SBW (1,5 bio. words, <https://crscardellino.github.io/SBWCE/>).

To separate the two semantic shifts, we had to treat *sin embargo* as a single word in corpus preparation since the embeddings algorithm does not recognize multi-word expressions. Unlike in the older period, *sin embargo* is nowadays much more frequent than the noun, which affects the calculation of the neighbors; without this step, its meaning would have merged with the meaning of *embargo* and dominated the choice of neighbors.

The lexical change of *embargo* is reflected in the change in similar words: in *Chronicles* the neighbors are words meaning ‘barrier, obstruction’, while in contemporary Spanish we find ‘block-age’ and ‘sanctions’ and names referring to historical embargos, like Irangate, Sadam (Hussein) or Unita (cf. Figure 1). As for the grammatical change, *sin embargo* was originally a PP ‘without obstacle/impediment’, a (clause-internal) verb modifier. Accordingly, almost all its neighbors in the old texts are verbs. The PP grammaticalized into a concessive connective, with loss of compositionality and increase in syntactic scope (Portolés Lázaro, 1995; Garachana Camarero, 1998); in SBW, its neighbors are words meaning ‘but, nevertheless, however’. In the Medieval/Classical period, *sin embargo* retains compositionality, cf. Amaral and Delicado Cantero (2018), as shown by some of the shared neighbors with the noun, including *estorbo/(d)estorbar* ‘hinder’, *quebran-tar*, *nocir* ‘harm’, and *recabdo/recabdar* ‘achieve’. The word embeddings also provide empirical evidence for the role of negative contexts in favoring the reanalysis of concessives: among the

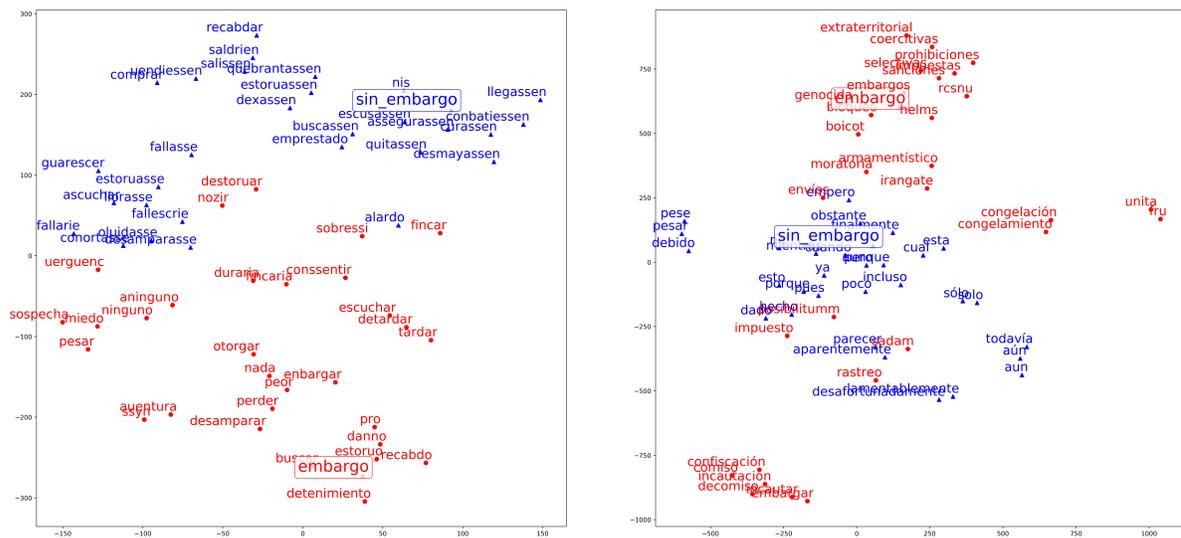


Figure 1: 30 most similar words to *embargo* and *sin embargo* in Medieval/Classical (left) and contemporary Spanish (right). Red: neighbors of *embargo*; blue: neighbors of *sin embargo*.

neighbors of *embargo* are negative words (e.g. *sin*, *nada*, *ninguno*), which confirms claims by Pérez Saldanya and Salvador (2015), a.o..

Our study provides methodological insights that can be applied to other cases of lexical and grammatical change: in the latter case, multi-word expressions must be treated as single words because of the frequency difference between time periods. From a theoretical perspective, our results provide a measure of compositionality of *sin embargo* in the older period (overlap in neighbors with the noun it contains) and insight into contexts favoring the creation of concessives.

References

- Amaral, P. and Delicado Cantero, M. (2018). Subcategorization and change: A diachronic analysis of *sin embargo* (*de que*). In McDonald, J., editor, *Contemporary Trends in Hispanic and Lusophone Linguistics: Selected Papers from the Hispanic Linguistics Symposium 2015*, pages 31–48. John Benjamins, Amsterdam.
- Boleda, G. (2020). Distributional semantics and linguistic theory. *Annual Review of Linguistics*, 6:13.1–13.22.
- Campbell, L. (2007). *Historical Linguistics: An Introduction*. The MIT Press, Boston.
- Firth, J. R. (1957). *Papers in Linguistics*. Oxford University Press, London.
- Garachana Camarero, M. (1998). La evolución de los conectores contraargumentativos: la gramaticalización de *no obstante* y *sin embargo*. In Zorraquino, M. A. M., editor, *Los marcadores del discurso: teoría y análisis*, pages 193–212. Arco Libros, Madrid.
- Hamilton, W. L., Leskovec, J., and Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pages 1489–1501, Berlin, Germany.
- König, E. (1988). Concessive connectives and concessive sentences. cross-linguistic regularities and pragmatic principles. In *Explaining language universals*, pages 146–166. Blackwell.
- Pérez Saldanya, M. and Salvador, V. (2015). Oraciones subordinadas concessivas. In *Sintaxis Histórica de la Lengua Española, 3a Parte*, pages 3699–3839. Fondo de Cultura Económica–UNAM.
- Portolés Lázaro, J. (1995). Diferencias gramaticales y pragmáticas entre los conectores discursivos *pero*, *sin embargo* y *no obstante*. *Boletín de la Real Academia Española*, 75:231–269.
- Traugott, E. C. and Dasher, R. (2002). *Regularity in Semantic Change*. Cambridge University Press, Cambridge.

Optional *se* constructions and flavours of applicatives in Spanish

Sudha Arunachalam (NYU) and Fabienne Martin (HU Berlin)

A. In Spanish, the reflexive clitic *se* routinely serves as a non-selected argument in transitive structures. Some authors claim that non-selected *se* appears in the context of consumption verbs only, see (1) (Zagona 1996, Campanini & Schäfer 2011). Others argue that it can also appear with a larger set of core and non-core transitive verbs, see (2-4).

- (1) a. Juan (*se*) comió las fresas. b. El mar (*se*) come la costa.
 Juan REFL eat.PFV.3SG the strawberries the sea REFL eat.PST.3SG the coast
- (2) Ayer Juan (*se*) cocinó una paella para sus invitados. (Sanz 2000)
 yesterday J. REFL cook.PFV.3SG a paella for his guests.
- (3) Josep (*se*) lavó todos los platos de la cena. (Sanz & Laka 2002)
 Josep REFL wash.PFV.3SG all the plates of the dinner
- (4) El submarino (*se*) hundió dos acorazados. (Otero 1999, Armstrong 2013)
 the submarine REFL sank.PFV.3SG two battleships
- (5) El perro/niño (*#se*) comió/bebió el hueso/la leche durante una hora.
 the dog/child REFL eat/drink.PFV.3SG the bone/the milk during an hour
- (6) a. Mi perro (*#se*) ha bebido aceite de cocina. b. El niño (*#se*) comió veneno!
 my dog REFL drink.PFV.3SG oil of cook the child REFL eat.PFV.3SG poison

A unified account of the semantic and pragmatic import of the reflexive in **optional *se* constructions (OSCs)** is notoriously challenging to pin down. Five inferences have been claimed to distinguish the *se*-marked and unmarked variants of OSCs, and they are often argued to co-occur with some verb types only.

#1. Telicity. *Se* in OSCs has famously been argued to require or enforce a telic interpretation of the underlying predicate with consumption verbs (Nishida 1994, see (5)-(6)), but also with all others (Sanz & Laka 2002). As such it is often analyzed as an aspectual operator, similar to particles of exhaustivity like English *up* (other authors claim, however, that bare nouns in the *se*-variant of OSCs (cf. (6)) are acceptable in some contexts; see Rivas 2011, Armstrong 2013). **#2. Easy-goingness/emphatic involvement.** The *se*-marked version in OSCs has been claimed to present the agent as involved, willful and enjoying the reported event, what Armstrong 2013 captures by positing that *se* spells-out a Voice head introducing a specific kind of ‘easy-going agents’. This inference is not triggered with verbs of consumption, which also take inanimate subjects in OSCs ((1b)). **#3. Benefactiveness/affectedness.** The *se*-variant in OSCs has been reported to convey that the DP’s referent is affected by the event (Arce-Arenales 1989) or benefits from it, what Rigau 1994 accounts for by positing that *se* is a benefactive argument. **#4. Counter-expectation.** *Se* in OSCs has been shown to present the event as surprising (Strauss 2003). In particular, de la Mora’s 2011 corpus study shows that *se* is strongly favoured when the object is not expected to be ingested. This inference has not been traced back to the other properties of *se* yet. **#5. Direct agency.** The *se*-variant of OSCs has been experimentally shown (Hodgson 2001) to convey that the agent executes the action on their own, without the intervention of another enabling entity, what Zagona 1996 translates in localist terms (the agent must remain *with* the object). The goal of this paper is to provide on the basis of new experimental data a unified analysis of OSCs, doing justice to the morpho-syntactic properties of the reflexive—it needs an antecedent and shows phi-feature agreement with it—and to explain which verb type/context triggers which inference(s).

B. Experiment 1 was an online pilot acceptability judgment task on a scale from 1 (fully acceptable) to 7 (fully unacceptable) on (among others) the *se*-variant of (1b-6), with 70 speakers of Peninsular Spanish, and 41 of American Spanish. Breaking the results down to a binary decision by dividing the scale into two disjunctive ranges ([1-3] for acceptable; [4-7] for non-acceptable), we obtained the percent of speakers accepting sentences (1b-6) in Table (7).

	(1b)	(2)	(3)	(4)	(5) <i>perro</i>	(5) <i>niño</i>	(6a)	(6b)
(7) American Sp. (N=41)	83%	41%	27%	17%	65%	51%	36%	29%
Peninsular Sp. (N=70)	84%	30%	35%	10%	38%	32%	45%	29%

Experiment 2 was a truth value judgment task. We showed 44 Spanish speakers (most from South America) the video clips used in Arunachalam & Kothari’s 2011 study on Hindi vs. English perfective sentences, depicting either a PARTIALLY complete event (e.g., eating half of a cookie) or a FULLY complete event (e.g., eating all of a cookie). We used one consumption verb (*comer*), one creation verb (*dibujar*) and 5 CoS verbs (*arrancar*, *cubrir*, *apagar*, *cerrar*, *llenar*), always with a quantized object. We manipulated two variables within subjects: FULL/PART and +/-*se*. Except for *comer*, there were more true judgments for *-se* sentences than *+se* ones in both FULL/PART conditions (*+se*/PART: 46%, *-se*/PART: 61%; *+se*/FULL: 70%; *-se*/FULL: 97%). This

suggests that event (in)completion is not what distinguishes the +*se* vs. –*se* variants of OSCs with verbs of creation and CoS verbs. For the consumption verb *comer*, 100% subjects liked both +*se*/–*se*-sentences in the FULL condition, but even in the PART situation, 73% judged +*se*-sentences as true (88% in –*se*-ones). This is much more than for English subjects tested on the same material by Arunachalam & Kothari 2011, only 17% of which accepted sentences such as *Ana ate up the cookie* in a PART situation. This suggests that *comerse* differs from *eat up*, as argued in §F below.

C. We propose that the variety of inferences identified above reflects the variety of applicative heads hosted by OSCs. In OSCs containing a consumption verb, *se* is merged as the indirect argument in the specifier of a low applicative head (below the verb) expressing an internalization of the theme by the dative argument (see (8a) from Campanini & Schäfer 2011). For a first set of speakers (that we call ‘Low-APPL speakers’), the Appl in OSCs can be low *only*. However, we propose that for a second set of speakers (‘Low/HIGH-APPL-speakers’), OSCs can also host a high applicative head (above the verb) with various semantic flavours, see (8b-c), including with consumption verbs.

- (8) a. $\llbracket \text{Appl}_{INTO} \rrbracket \rightsquigarrow \lambda x \lambda y \lambda P. \mathbf{theme}(e, x) \wedge P(e) \wedge \exists s(\mathbf{cause}(e, s) \wedge \mathbf{into}(x, y, s))$
 b. $\llbracket \text{Appl}_{BEN} \rrbracket \rightsquigarrow \lambda P \lambda x \lambda e. P(e) \wedge \mathbf{beneficiary}(e, x)$ (Pylkkänen 2002)
 c. $\llbracket \text{Appl}_{EXP} \rrbracket \rightsquigarrow \lambda P \lambda x \lambda e. P(e) \wedge \mathbf{experiencer}(e, x)$ (Myler 2014)

When *se* serves as the argument of Appl (8a), the incorporative meaning of consumption verb OSCs is not simply derived from world-knowledge about what eating events are (as the case in the unmarked variant), but is encoded *structurally* and thus focused on. We argue that this accounts for the fact that *se* is strongly favoured when the object is not expected to be eaten (de la Mora 2011, inference #4), at least for speakers for which *se* is truly optional in (1). But it also explains why many speakers in fact *require se* in OSCs with consumption verbs (17/32 speakers tested by D’introno et al. 2007): for them, if the incorporative meaning entailed by consumption verbs *can* be encoded in the syntax, it *must* be so.

Other verbs than consumption verbs do not combine with low Appl (Campanini & Schäfer 2011). For LOW-APPL-speakers, sentences such as (2-4) thus turn out ungrammatical. As Table 7 shows, the percent of speakers accepting (2)-(4) is indeed much lower than for (1), confirming the existence of this LOW-APPL group. But it is still far from negligible in the case of (2)-(3). We take this to confirm the co-existence of LOW-APPL and LOW/HIGH-APPL-dialects, cross-cutting the Amer. Sp vs. Penins. Sp divide.

D. In (2)-(4), *se* is the argument of the high Appl (8b) or (8c). The subject is then coindexed *via* reflexivisation with a beneficiary or experiencer, and thus needs to be animate. This accounts for the low acceptability of (4), whose subject is inanimate. With Appl_{BEN} (8b), inferences #2/#3 are triggered. We argue that Appl in (8c) is similar to the head in experiencer *have* sentences (Harley 1998, Myler 2014). It introduces an experiencer, identified by *se* with the agent (‘Josep had himself cook the dinner’). Since the agent also experiences the event, they must be a ‘direct’ executor (and not a simple instigator), which is at the source of the ‘direct agency’ inference #5. Inference #2 is also easy to explain via the experiential semantics of Appl in (8c). Finally, the dispreference for the +*se*-variant with all verbs except *comer* observed in Experiment 2 can be linked to the fact that the agent in the videoclips used in Arunachalam and Kothari 2011 is neutral, and not pictured as an easy-going beneficiary/experiencer.

E. Spanish bare nouns are property-denoting and combine with the verb via pseudo-incorporation (Espinal & McNally 2011). As a result, bare nouns cannot feed the first argument of the Low Appl (8a). This explains the unacceptability of (6a/b) for Low APPL speakers. However, a bare noun can form with the verb a VP feeding the property *P* in (8b/c). Hence why (6a/b) are acceptable for LOW/HIGH-speakers (an average of 35% of our 111 speakers for (6a/b)).

F. Experiment 2 showed that *comerse* is judged true in a partial-event situation by 73% of Spanish subjects, while *eat up* was accepted only by 17% of English speakers in the same situation. Also, sentences (5) are accepted by 46,5% of speakers on average, while *eat up* strongly rejects atelic uses. We account for this aspectual difference as follows. *Comerse la pizza* is more strongly telic than *comer la pizza* for *se* turns a mono-eventive verb into a bi-eventive predicate (see (9a)). Still, bi-eventive predicates such as (9a) are known to tolerate atelic uses in the appropriate context (Kearns 2007, Piñón 2007). By contrast, atelic uses are banned for English *eat up the pizza*. This is accounted by Piñón’s 2011 analysis of *eat up* repeated in (9b/c), according to which *eat up* denotes *pairs* of eventualities whose second member is an event boundary. This rules out cumulative uses for (9b), still possible for (9a).

- (9) a. $\llbracket \text{J comerse la pizza} \rrbracket \rightsquigarrow \lambda e. \mathbf{eat}(e) \wedge \mathbf{agent}(e, \mathbf{J}) \wedge \mathbf{theme}(e, \mathbf{p}) \wedge \exists s(\mathbf{cause}(e, s) \wedge \mathbf{into}(\mathbf{p}, \mathbf{J}, s))$
 b. $\llbracket \text{J eat up the pizza} \rrbracket \rightsquigarrow \lambda \langle e, b \rangle. \mathbf{eat}^+(\langle e, b \rangle) \wedge \mathbf{agent}(\langle e, b \rangle, \mathbf{J}) \wedge \mathbf{theme}(\langle e, b \rangle, \mathbf{p})$ (Piñón 2011)
 c. $\forall \langle e, b \rangle (\mathbf{V}^+(\langle e, b \rangle) \rightarrow b = \mathbf{right-boundary-of}(e) \wedge b \sqsubset e)$

The Pluperfect First Hypothesis

The compound pluperfect as a necessary precondition of the perfect-to-perfective shift
William Balla-Johnson, *University of North Georgia*

The Romance compound tenses have garnered much attention from linguists over the past century. However, the pluperfect has often been overlooked in favor of the present perfect in particular, but also the future and the conditional. The pluperfect deserves attention as an object of study in its own right since it has been observed to not only follow, but to lead changes that subsequently occur in the present perfect, as well as develop independently from other tenses.

This project examines the development of the compound pluperfect [CPP] (imperfect auxiliary + past participle, ex. Sp. *había cantado* ‘I had sung’) in the western European Romance languages and specifically what relation this process had with the perfect-to-perfective shift (also called *aoristic drift* (Bybee, et al. 1994)) in the Romance compound past [CP] (present auxiliary + past participle, ex. Sp. *(yo) he cantado* lit. ‘I have sung’). Descriptive data is presented from several large historical corpora of Spanish, French, and Portuguese which show that the generalization of the compound pluperfect was complete before any subsequent perfect-to-perfective shift in the areas where it occurred. The generalization of the compound pluperfect is thus suggested to be a necessary precondition of the perfect-to-perfective shift of the compound past.

This perfect-to-perfective shift is seen in the evolution of the semantics of the French compound past *j’ai chanté*, from originally ‘I have sung’ (perfect), to now ‘I sang’ (perfective), which was traditionally only able to be expressed by the simple past [SP] (also called the preterit, ex. Fr. *je chantai*). Once this shift was complete, the compound past entered into variation with the simple past, which resulted in the functional loss of the latter.

Although it happened far earlier, the loss of the Latin synthetic pluperfect (CANTAVRAM ‘I had sung’) and the generalization of the Romance compound pluperfect has been recognized as analogous to the generalization of the compound past at the expense of the simple past which took place in early modern French and is currently partially underway in peninsular Spanish (Squartini 1999).

The proposal presented here, dubbed the **Pluperfect First Hypothesis**, states that any perfect-to-perfective shift is only able to occur in the compound past after the compound pluperfect underwent an identical semantic extension. As the aspect of the compound pluperfect extended from perfect-in-past to perfective-in-past (cf. Squartini 1999), it came into competition with the original Romance synthetic pluperfect (called the ‘-ra form’, ex. Pt. *cantara* ‘I had sung’). After a period of variation, the compound pluperfect was then generalized at the expense of the -ra form in all of western Romance, with the exception of Galician and literary Portuguese.

Figure 1 below shows the semantic extension of the pluperfect and the compound past. As evidenced by their staggered positions on the timeline, in every language here examined, the compound pluperfect becomes the generalized form for the perfective-in-past (Stage 5, on the right) before the compound past’s first extension into perfective aspect (Stage 3, on the left). Each development has not necessarily fully attained in each language, but in every period and in every main variety in western European Romance, the aspectual development of the pluperfect is more advanced.

Geographically, the areas in which the compound pluperfect was first generalized show the largest perfect-to-perfective shift in the compound past. Specifically, modern French and Italian show the most advanced shift (Squartini & Bertinetto 2000), and likewise had a generalized perfective compound pluperfect earlier than in other languages where the shift is less advanced (Occitan, Catalan, Spanish), or altogether absent (Portuguese, Galician). Pictured below in Figure

2, this work adds a dynamic and chronologically expanding variable to the history of western European Romance. Shown below in Figure 2, between the 11th and the 19th centuries, the generalized compound pluperfect is shown to gradually spread westward from French and the northern Italian varieties, to Occitan, Catalan, and Spanish, and finally into Portuguese.

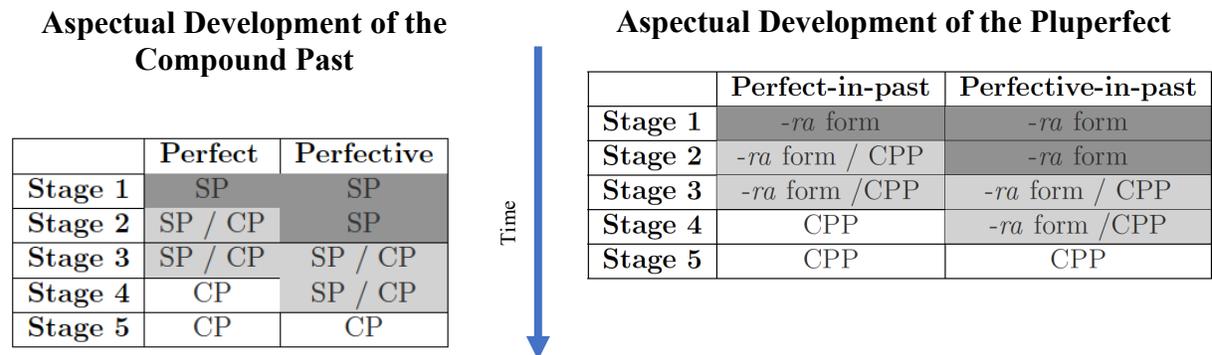


Figure 1 – Aspectual development of the Compound Past [CP], Simple Past [SP], Compound Pluperfect [CPP] and the *-ra* forms along a timeline

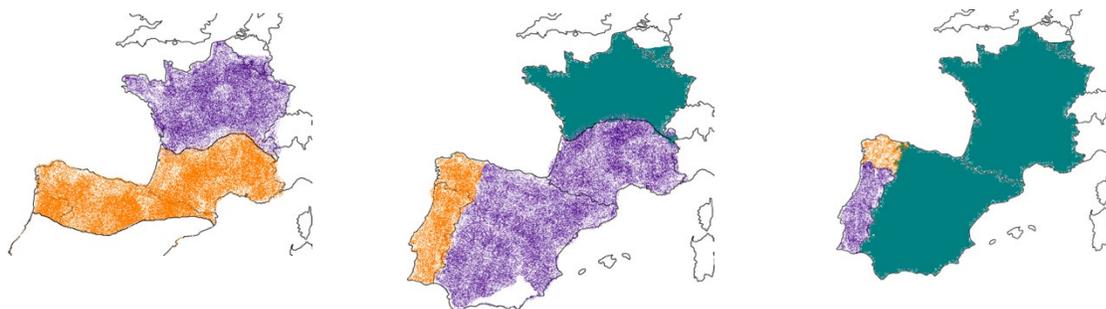


Figure 2 - Western Romance pluperfects c. 1000 (left), c. 1400 (center), c. 1800 (right). **Orange**: productive *-ra* pluperfect indicative; **Purple**: variation between a compound pluperfect and *-ra* forms; **Green**: generalization of the compound pluperfect

In sum, this project presents evidence that the development of a generalized compound pluperfect and the degree of perfect-to-perfective shift experienced by the compound past are not simply parallelisms in Romance. Rather, the early development of a generalized perfective-in-the-past compound pluperfect provided a model of a compound tense with perfective aspect. Then, through analogy the compound past also began to be used in perfective, in addition to perfect, contexts (as in early modern French and modern Spanish). However, the generalization of a compound pluperfect did not always lead to the perfect-to-perfective shift of the compound past. The generalization of a compound pluperfect is thus named a necessary precondition of the perfect-to-perfective shift of the compound past.

Selected References:

- Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Squartini, Mario. 1999. On the semantics of the Pluperfect: Evidence from Germanic and Romance. *Linguistic Typology* 3(1). 51-90.
- Squartini, Mario & Pier Marco Bertinetto. 2000. The simple and compound past in romance languages. In Osten Dahl (ed.), *Tense and Aspect in the Languages of Europe*, Mouton-de Gruyter.

Nasal airflow evidence for complete resyllabification in Spanish

Ander Beristain (anderb2@illinois.edu), *University of Illinois at Urbana-Champaign*

Introduction and Background. The conceptualization of the syllable as a unit for articulatory organization is neither a novel concept nor one that has received full support from its beginnings. Several studies have investigated the way in which syllabic structure affects gestural overlapping of the articulators of the segments in a given sequence. It has been postulated that greater degree of coarticulation and weakening will occur in syllable-final (rather than syllable-initial) segments (Krakow, 1999), alluding to the notion that certain processes are more likely to affect segments in a given syllable type and employing diachronic processes to exemplify that. Krakow (1999) addressed this issue with regards to coarticulatory vowel nasalization in English and demonstrated the correlation between syllabic structure and gesture overlap regarding labial and velum movement. According to her results, a greater degree of nasal coarticulation (i.e., earlier lowering of the velum) occurred in tautosyllabic /Vn./ sequences as opposed to heterosyllabic /V.nV/ sequences due to biomechanical reasons and an easier velum lowering transition from vowel to coda nasal as opposed to the articulatory effort of an onset nasal followed by a fully oral vowel.

Traditionally, it has been assumed that word-final consonants in Spanish undergo *resyllabification* before a vowel in the next word (Harris and Kaisse, 1990, p. 131). For instance, *un amigo* would undergo the following rules: /un amigo/ → [un. a.mi.ɣo → u|n.a.mi.ɣo → ~~linking~~yo (attach onset)]. Remarkably, in this view, weakening phenomena that affect word-internal coda consonants may also affect word-final consonants even if they are resyllabified as onsets preceding a vowel in the following word, which has been interpreted as evidence for either rule ordering (Harris, 1983) or constraint ranking within Optimality Theory (Colina, 1997). In varieties where /s/ reduces to [h] in the coda (but not in word-internal position), word-final /s/ is also aspirated when resyllabified in /Vs#V/, suggesting /mas amor/ → [mah. a.mor → ma.ha.mor] vs. /VsV/, e.g. /masa/ → *[ma.ha]. Similarly, in velarizing dialects, word final /n/ is velarized in /Vn#V/, e.g. /ban a kasa/ → [baŋ a. ka.sa → ba.ŋa.ka.sa] vs. /VnV/, e.g. /bana/ → *[ba.ŋa] ‘vain’.

The existence of obligatory and complete resyllabification of /C#V/ sequences in Spanish has been called into question more recently. A number of recent studies (Hualde & Prieto, 2014; Strycharczuk & Kohlberger, 2016, among others) have concluded that a complete resyllabification does not take place with regard to sibilant fricatives as the durational and voicing properties of canonical intervocalic onset positioned fricatives differ from the derived ones. However, aspects such as the coordination of oral and velum gestures in resyllabification have been disregarded.

Research Question and Hypothesis. We investigate the existence of complete resyllabification by comparing the degree of vowel nasalization in /Vn.C/ (where V and /n/ are unambiguously tautosyllabic, e.g. *can.sar*), across syllable boundaries in /V.nV/ (e.g. *pa.na*) and /V.n#V/ (e.g. *pan asado*). Crucially, the resyllabification hypothesis predicts /Vn.C/ > /V.nV/ = /V.n#V/.

Methodology. 9 native speakers of Peninsular Spanish participated in a read-aloud task in Spanish from which aerodynamic data were extracted using two pressure transducers connected to a vented mask with one oral and one nasal cavity. Each participant produced 40 target tokens under four conditions: /CV.CV/ (oral control), /CVn./ (nasalized V, tautosyllabic *n*), /CV.nV/ (nasalized V, heterosyllabic *n*), and /CV.n#V/ (nasalized V, resyllabified *n*). 40 time-normalized points were extracted from each vocalic segment in order to observe nasal airflow patterns across time. Thus, 20,160 datapoints were submitted to the statistical analysis which employed a Generalized Additive Mixed Model (GAMM) in R.

Results. Regarding anticipatory vowel nasalization, the onset of this is significantly greater in the tautosyllabic (/CVn./) sequence (85%) than in the other two structures. Heterosyllabic (/CV.nV/) and resyllabified (/CV.n#V/) sequences show their onsets at 92% and 91%, respectively, and no significant statistical results are found when compared; yet, significant results are found when they are compared to the tautosyllabic sequence (Figure 1). Regarding carryover nasalization, vowels in canonical /nV/ and derived /n#V/ sequences show symmetrical (and non-significant) nasalization patterns when compared (Figure 2).

Conclusion. Evidence was found to consider the syllable as articulatory organization. As expected, tautosyllabic sequences exhibited greater coarticulatory nasalization degree than heterosyllabic sequences. Derived (resyllabified) heterosyllabic sequences present symmetric nasalization patterns with canonical ones, thus showing that a complete resyllabification occurs, contrary to the results in Hualde and Prieto (2014) and Strycharczuk and Kohlberger (2016).

Figures

Figure 1: Anticipatory vowel nasalization patterns

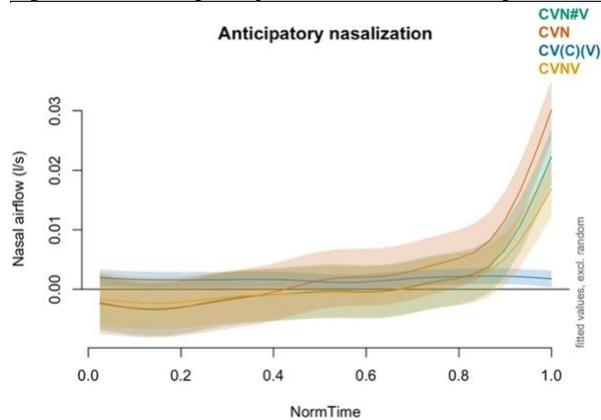
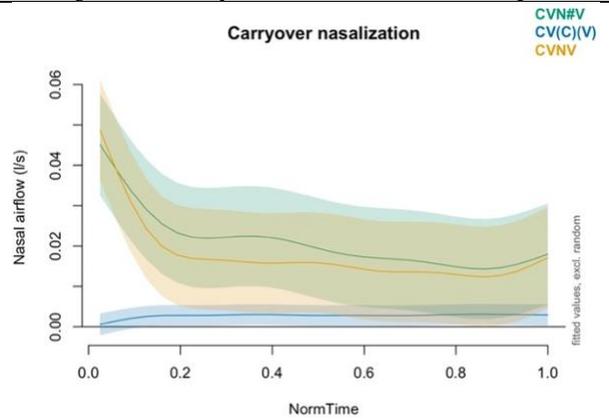


Figure 2: Carryover vowel nasalization patterns



References

- Colina, S. (1997). Identity constraints and Spanish resyllabification. *Lingua*, 103, 1–23.
- Harris, J. (1983). Syllable structure and stress in Spanish: A nonlinear analysis. MIT Press.
- Harris, J. W., & Kaisse, E. M. (1999). Palatal vowels, glides and obstruents in Argentinian Spanish. *Phonology*, 19, 117–190.
- Hualde, J. I., & Prieto, P. (2014). Lenition of sibilants in Catalan and Spanish. *Phonetica*, 71, 109–127.
- Krakow, R. A. (1999). Articulatory organization of syllables: A review. *Journal of Phonetics*, 27, 23–54.
- Strycharczuk, P., & Kohlberger, M. (2016). Resyllabification reconsidered: On the durational properties of word-final /s/ in Spanish. *Journal of the Association for Laboratory Phonology*, 7(1), 3.

Loanword Adaptation in Portuguese-American Speech

Justin Bland, The Ohio State University

In languages that do not permit /sC-/ onsets, loanwords with such onsets must be repaired, often by inserting a vowel. This vowel may be inserted before the cluster as in (1a), which is termed prothesis, or within the cluster as in (1b), which is termed anaptyxis (Fleischhacker 2005):

- (1) a. English *stop* > Spanish [es.top]
b. English *stop* > Japanese [su.top.puu]

Standard European Portuguese (EP) and Brazilian Portuguese (BP) traditionally repair all /sC-/ onsets with prothesis, for example *ski* [iʃ.'ki] 'ski' and *eslavo* [iʒ.'la.vu] 'Slav' (Mateus & d'Andrade 2000; Bisol 1999). However, phonetic transcriptions of the speech of EP-speaking immigrants to the US, which I term Luso-American Portuguese (LAP), suggest that these speakers tend to repair sibilant + stop onsets with prothesis (2a) and sibilant + sonorant onsets with anaptyxis (2b):

- (2) a. English *steamer* > LAP [iʃ.'ti.mə]
b. English *snow* > LAP [si.'nɔ]

Such a pattern of loanword adaptation appears unexpected given a general Ibero-Romance preference for repairing all /sC-/ onsets with prothesis; however, it turns out to be common cross-linguistically, occurring in languages as diverse as Egyptian Arabic, Bengali, Hindi, Sinhalese, and Wolof (Broselow 2015; Fleischhacker 2005).

This pattern then requires an explanation that takes into account both standard EP phonology as well as possible differences between EP and LAP. In this study, I first examine and propose an analysis of the loanword adaptation of complex onsets in standard EP and BP. I then present and analyze a small corpus of American English loanwords with /sC-/ onsets in LAP. Finally, I propose an explanation for the repair pattern observed in LAP based on the reanalysis of what Hall (2003) terms intrusive vowels as epenthesis.

Standard EP and BP typically adapt /sC-/ onsets via prothesis (e.g. *spa* [iʃ.'pa] 'spa') and other illicit complex onsets via anaptyxis (e.g. *pneu* [pi.'new] 'tire') (Collischonn & Wetzels 2016; Keller 2010; Mateus & d'Andrade 2000). Collischonn & Wetzels (2016) and Keller (2010) propose two different explanations for this pattern. I take a middle ground, drawing from but ultimately diverging from both analyses. Specifically, I suggest that complex onset adaptation in standard EP and BP can be explained as a tension between faithfulness constraints prohibiting word-medial insertion and markedness constraints prohibiting illicit codas. On the one hand, Keller's (2010) use of the OUTPUT-CONTIGUITY constraint, which prohibits word-medial insertion, results in an overall preference for prothesis. On the other hand, as Collischonn & Wetzels (2016) note, Portuguese prohibits most codas, with the notable exception of /s/. I propose that a combination of these arguments best explains complex onset adaptation in standard EP and BP. That is, prothesis is used in words like *spa* [iʃ.'pa] because prothesis is preferred overall and it results in a licit /s/ coda. But anaptyxis must be used in a word like *pneu* [pi.'new], because even though prothesis is preferred, its use in *pneu* *[ip.'new] would result in an illicit /p/ coda.

Moving from standard EP to LAP, I then present a small corpus of American English loanwords that I compiled from several earlier descriptions of immigrant varieties of EP (Borges 1960; Pap 1949; Dias 1989). According to this data, 74% of words with sibilant + stop onsets were adapted through prothesis (e.g. *steamer* [iʃ.'ti.mə]), while 63% of words with sibilant + sonorant onsets were adapted

through anaptyxis (e.g. *snow* [si.'nɔ]). With one exception, the remaining /sC-/ onsets were left unrepaired. Importantly, only one sibilant + stop onset was repaired through anaptyxis, and no sibilant + sonorant onsets were repaired through prothesis. This data suggests a general pattern that sibilant + stop onsets are repaired through prothesis and sibilant + sonorant onsets are repaired through anaptyxis in LAP.

Finally, I propose an explanation for the pattern observed in LAP. Following Boersma & Hamann (2009) and others, I argue that loanword adaptation is influenced by both phonology and perception—in this case, the perception of intrusive vowels. Hall (2003) demonstrates that in certain contexts a phonetic vowel-like element, which she calls an intrusive vowel, can surface between two consonants, for example in pronunciations like *puh-lease* for *please*. According to Hall, sonorants are a common trigger for intrusion, making variable intrusion possible in English sibilant + sonorant onsets but not in sibilant + stop onsets. Since /sC-/ onsets are illicit in Portuguese, listeners may be particularly attentive to vocalic cues for epenthesis in such contexts, and if vowel intrusion is present in a word like *snow*, they may reinterpret the intrusion as epenthesis, resulting in anaptyxis: [si.'nɔ]. On the other hand, since vowel intrusion is not possible in a word like *steamer*, LAP speakers must rely on the Portuguese phonological system, which specifies prothesis in this context: [iʃ.'ti.mɐ]. Complex onset repair in LAP is thus influenced both by the perception of vocalic cues and by the phonology of Portuguese.

In summary, this study examines the adaptation of loanwords with /sC-/ onsets in both standard Portuguese and an immigrant variety of EP spoken in the US. I propose an analysis of /sC-/ onset adaptation in standard EP based on a tension between the tendency to avoid medial insertion and the need to avoid illicit codas. I then present a corpus of American English loanwords in LAP, which shows a pattern that differs from standard Portuguese. Finally, I argue that the pattern observed in LAP can be explained as an interaction between standard Portuguese phonology and the reinterpretation of perceptual cues for vowel intrusion as vowel epenthesis.

References

- Boersma, P., & Hamann, S. (2009). Loanword adaptation as first-language phonological perception. In A. Calabrese & W. L. Wetzels (Eds.), *Loan phonology* (pp. 11-58). John Benjamins.
- Borges, N. O. da C. (1960). Influência anglo-americana no falar da ilha de S. Miguel (Açores). *Revista Portuguesa de Filologia*, Suppl. 2, 1-83.
- Broselow, E. (2015). The typology of position-quality interaction in loanword vowel insertion. In Y. Hsiao & W. Lee (Eds.), *Capturing phonological shades within and across languages* (pp. 292-319). Cambridge Scholars.
- Collischonn, G., & Wetzels, W. L. (2016). Syllable structure. In W. L. Wetzels, S. Menuzzi, & J. Costa (Eds.), *The handbook of Portuguese linguistics* (pp. 86-106). Wiley-Blackwell.
- Dias, E. M. (1989). *Falares emigrantes: Uma abordagem ao seu estudo*. Instituto de Cultura e Língua Portuguesa.
- Fleischhacker, H. A. (2005). *Similarity in phonology: Evidence from reduplication and loan adaptation*. (Doctoral dissertation.) University of California Los Angeles.
- Hall, N. (2003). *Gestures and segments: Vowel intrusion as overlap*. (Doctoral dissertation.) University of Massachusetts Amherst.
- Keller, T. (2010). O alinhamento relacional e o mapeamento de ataques complexos em português. *Letras de Hoje*, 45, 61-70.
- Mateus, H., & d'Andrade, E. (2000). *The phonology of Portuguese*. Oxford University Press.
- Pap, L. (1949). *Portuguese-American speech: An outline of speech conditions among Portuguese immigrants in New England and elsewhere in the United States*. King's Crown.

Accounting for participants' best linguistic behavior: Vowel nasalization, nasal consonant weakening and hypercorrect /s/

Silvina Bongiovanni, *Michigan State University*

It is not uncommon in laboratory studies for speakers to modify their speech (e.g., by hyper-articulating, attempting a “supra-regional” norm, or hypercorrecting) to ‘speak correctly’, especially if interacting with out-of-community researchers. Often the result of this heightened attention to speech is facilitated by elicitation techniques, though other times it is rooted in linguistic insecurity and (perceived) higher linguistic standing of the interlocutor. In any case, researchers working within laboratory approaches have had to balance their methods with (lack of) naturalness of the speech sample. For example, in a recent study analyzing vocalization of /l/ and /r/, Willis & Ronquest (2021) use percentage of use of /s/ to account for stylistic-modified speech. Results indicate that percentage of use of /s/ (retained/aspirated vs. elided) predicted the use of canonical variants of liquid phonemes. In other words, speakers with a higher tendency to retain or aspirate /s/ also produced canonical variants of liquid phonemes (i.e. [l] and [r] respectively). The present study extends this idea to the examination of anticipatory vowel nasalization and nasal consonant weakening. Is it the case that with production of full /s/ nasal consonants are less likely to weaken and exhibit shorter temporal extent of nasalization in the pre-nasal vowel? This study compares production data from Argentine and Dominican speakers.

Two key findings in previous work are relevant to the research reported here. First, though Argentine and Dominican Spanish are /s/-leniting dialects, though Dominican Spanish is more advanced in the debuccalization process. Thus, full /s/ it is not the most expected realization, especially among Dominicans. Second, dialects also differ in terms of nasal consonant weakening and anticipatory vowel nasalization. Previous work comparing Argentine and Dominican Spanish has shown that the latter exhibits earlier onset of nasalization, consistent with an interpretation of phonologized nasalization for this dialect group (Bongiovanni 2021, in revision). Interestingly, this work has also found that similar degrees of nasal consonant weakening (measured in terms of duration and consonant-to-vowel ratio) across dialect groups, which was unexpected. These findings thus raise the genuine possibility that results may have been the artifact of heightened attention to speech due to the nature of lab speech combined with well-known differences in linguistic (in)securities of speaker groups (Büdenbender 2010, Toribio 2000). The present study returns to this data set and examines the production of nasal consonant weakening and anticipatory vowel nasalization as a function of (broadly defined) attention to speech, which is measured in terms of the production of /s/.

The data analyzed come from 28 speakers of Dominican Spanish and 26 of Argentina Spanish, recorded with a nasometer, a split channel set of microphones that record nose and mouth signals separately. Onset of nasalization was operationalized as the first time point to exceed a 15% of the range between the minimum and the maximum nasal energy reading (Bongiovanni 2021, Delvaux, Demolin, Harmegnies, & Soquet 2008; Solé 1992). Nasal consonant weakening was analyzed in terms of temporal reduction, i.e. measured in duration (Beddor 2007, 2009). I hypothesize that if speakers are putting forth their ‘best’ linguistic behavior, their production of /s/ will likely not exhibit signs of lenition. For this purpose, the duration of /s/ was used as a proxy for attention to speech. It is expected that a longer /s/ (that is, frication, [s] or [h]), will co-occur with more constricted nasal consonants and shorter temporal extent of nasalization.

Findings indeed show that longer /s/ co-occur with longer nasal consonants, a statistical trend in both dialect groups (Figure 1). However, when duration of /s/ was compared to the time-course of nasalization, it was only the Argentine data which showed statistical earlier onset of nasalization with shorter /s/ (Figure 2). These results confirm previous findings regarding covariation of nasal consonant weakening and anticipatory vowel nasalization (Bongiovanni, in revision). Because nasalization has phonologized as a feature of the vowel (and no longer a marker of nasal consonant weakening), it does not varied as was hypothesized.

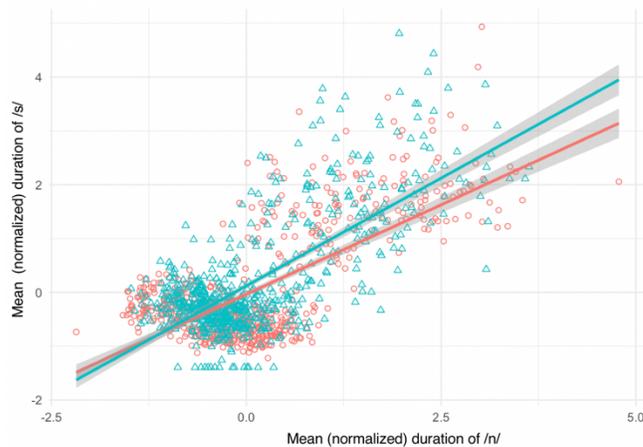


Figure 1 – Mean duration of /n/ as a function of mean duration of /s/

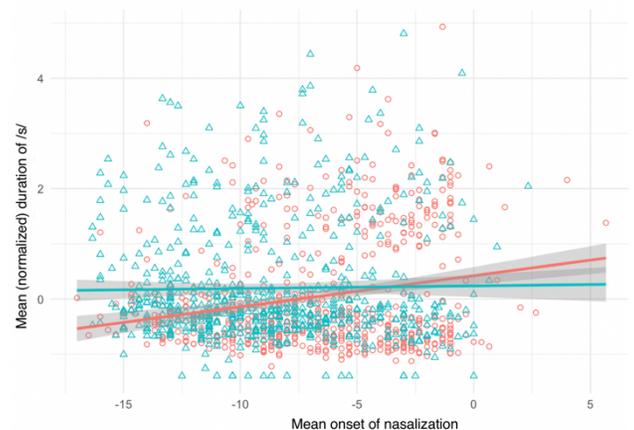


Figure 2 – Onset of nasalization as a function of mean duration of /s/. Time point zero represents the vowel-nasal boundary

Green = Dominican data; pink = Argentine data

References

- Beddor, P. (2009). A coarticulatory path to sound change. *Language*, 85(4), 785–821.
- Beddor, P. (2007). Nasals and nasalization: The relation between segmental and coarticulatory timing. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th international congress of phonetic sciences (ICPhS 2007)*, Saarbrücken, 6-10 August, 2007 (pp. 249-254) Dudweiler: Pirrot.
- Bongiovanni, S. (2021). Acoustic investigation of anticipatory vowel nasalization in a Caribbean and a non-Caribbean dialect of Spanish. *Linguistics Vanguard*, 7(1).
- Bongiovanni, S. (in revision). On covariation between nasal consonant weakening and anticipatory vowel nasalization: Evidence from a Caribbean and a non-Caribbean dialect of Spanish.
- Büdenbender, E.-M. S. (2010). Comparing Dominican linguistic (in)security in the Dominican Republic and in the diaspora. In C. Borgonovo, M. E., Echevarría & P. Prévost (Eds.), *12th Hispanic Linguistics Symposium* (pp. 148–159). Somerville, MA: Cascadilla Proceedings Project.
- Delvaux, V., Demolin, D., Harmegnies, B., & Soquet, A. (2008). The aerodynamics of nasalization in French. *Journal of Phonetics*, 36(4), 578–606.
- Solé, M.-J. (1992). Phonetic and phonological processes: The case of nasalization. *Language and Speech*, 35(1-2), 29–43.
- Toribio, A. J. (2000). Nosotros somos dominicanos: Language and self-definition among Dominicans. In A. Roca (Ed.), *Research on Spanish in the United States: Linguistic issues and challenges* (pp. 252–270). Somerville, MA: Cascadilla Press.
- Willis, E., & Ronquest, R. (2021). A sociophonetic explanation of coda liquids and vocalization in Cibao Dominican Spanish. In Manuel Díaz-Campos (Ed.), *Variationist Approaches to Sociolinguistics*.

**The Acquisition of Stop Consonants by Chinese Learners of Spanish:
A Longitudinal and Cross-sectional Study**

Celia Bravo Diaz (bravo065@umn.edu), *University of Minnesota*

While there is a vast amount of research that explores the acquisition of stop consonants in L2 Spanish by native speakers of English (Bongiovanni et al, 2015; Casillas, 2020; Díaz-Campos, 2004; Zampini, 1998; among others), there are only two studies that focus on speakers of L1 Chinese (Chen 2007; Liu 2016) and none that include both perception and production. This study examines the relationship between the production and perception of word-initial stop consonants in Spanish (/p, t, k, b, d, g/) by 82 adult Chinese learners studying abroad in Spain, from true beginners to near-native speakers.

Production data was elicited via two reading tasks and six semi-spontaneous tasks, and perception data was elicited through a discrimination task with 120 tokens created from natural stimuli. A total of 16,470 production tokens and 9,840 perception tokens were analyzed. For production, the voice onset time for each stop—measured as the interval between the beginning of the release burst and the onset of the glottal signal reflected in F1 in the following vowel—was analyzed using Praat (Boersma and Weenink 2006). Generalized linear mixed effects models were used to compare the learning trajectories of each modality and to determine the probabilistic weighting of thirteen independent variables on the production and perception of the stop consonants at different levels of proficiency.

The results revealed phonetic learning of stops over the course of the semester following a nonlinear trajectory subject to individual and segment-specific variability. Development in the production of voiced stops only emerged in the intermediate-high and advanced groups once perceptual discrimination had reached a high level of accuracy, which suggests that enhanced perception of the Spanish voicing contrast in beginner and intermediate learners was insufficient to promote the acquisition of voiced consonants. Based on this analysis, I argue that while L2 phonetic category formation can occur at an early stage of development for students exposed to substantial native-like L2 input, gains in perception for new sound categories transfer to production at a later stage. I conclude that a time-lagged model might be more appropriate to understand the perception–production link in L2 Spanish than the synchronous model assumed by previous research.

Theoretically, this research opens a window into the core premises of current models of second language speech, largely based on cross-sectional data, by investigating how perception and production are related over time in a fast-growing population—Chinese learners of Spanish. It proposes an expansion of the existing models of second language phonology that accounts for the interaction between the first language and additional second languages, as well as the role of individual differences, in order to more satisfactorily explain the acquisition of the fine-phonetic detail of stop voicing. Moreover, this work shows the value of incorporating social network analysis—an exploration of the students’ membership in social networks in the target-language community—in the studies of L2 phonology as an external factor that influences the boundary resetting of phonemic categories and the shifting of perceptual boundaries.

References

Boersma, P. and Weenink, D. (2006). *Praat: Doing Phonetics by Computer* (Version 4.45.2.4), 4 November 2010, <http://www.praat.org> (4 November 2010).

Bongiovanni, S., Long, A., Solon, M., & Willis, E.W. (2015). The effect of short-term study abroad on second language Spanish phonetic development. *Studies in Hispanic and Lusophone Linguistics*, 8(2), 243-283.

Casillas, J. V. (2020). Phonetic Category Formation is Perceptually Driven During the Early Stages of Adult L2 Development. *Language and Speech*, 63(3), 002383091986622-581.

Chen, Y.A (2007). A comparison of Spanish produced by Chinese L2 learners and native speakers: An acoustic phonetics approach. Doctoral dissertation.

Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, 26, 249-273.

Liu, Z. (2016). Exploring cross-linguistic influence: Perception and production of L1, L2 and L3 bilabial stops by Mandarin Chinese speakers. Doctoral dissertation.

Zampini, M. (1998). The relationship between the production and perception of L2 Spanish stops. *Texas Papers in Foreign Language Education* 3, 85-100.

**Forgotten allophones and their gradient friends:
complexity in the acquisition of Spanish voiced stops**

Rebeka Campos-Astorkiza, Katriese DeLeon, Kendall LoCascio and Shannon Sullivan
Ohio State University

The acquisition of Spanish voiced stops /b, d, g/ by L1 English learners offers a rich and complex testing ground for theories of second language acquisition (SLA). On the one hand, learners need to acquire a new set of allophones, i.e. the weakened approximants [β, ð̞, ɣ̞] that occur in Spanish after continuant sounds. In addition, they need to learn a new distribution for their L1 voiced stop allophones, namely [p, t, k] and [b, d, g], since voiceless productions are only allophones of /p, t, k/ in Spanish. Exploring both aspects, i.e. the acquisition of new allophones and of a new distribution, allows us to address issues related to the role of L1 and of the (dis)similarity of L1 and L2 sounds in SLA. Previous studies on L2 Spanish /b, d, g/ focus mainly on the acquisition of the approximant allophones and seem to conclude that these sounds present a greater challenge than other types of sounds (Diaz-Campos 2004, cf. Nagle 2017), although learners can improve their production especially in contexts of study abroad and explicit instruction (Lord 2010, Alvord & Christiansen 2012). Most of these studies use a categorical analysis of the learners' data, namely they focus on the type of allophone (i.e. voiced stop vs. approximant; Face & Menke 2009). However, studies on L1 Spanish voiced stop allophones show that approximant/weakening production in the language is gradient and conditioned by linguistic factors. This gradience is captured by measuring the Consonant-to-Vowel (CV)-intensity ratio (e.g. Carrasco et al. 2012), where higher ratios mean more weakening. Some studies have used this acoustic measurement to examine L2 Spanish (Shea and Curtin 2011, Rogers and Alvord 2014, Bongiovanni et al. 2015, Nagle 2017). We argue that a mixed-methods approach, using the two measurements, i.e. the type of allophone and the CV-intensity ratio, allows for a more accurate picture of the learning process (Solon et al. [2018] also use a categorical and continuous analysis but they explore only the acquisition of /d/). In addition, we expand the categorical analysis by also considering voiceless stop productions and including a variety of contexts to better address the question of how a new distribution of allophones is learned. Thus, our research questions are: (i) how do voiced stops categorical realizations change in the acquisition process?, (ii) how does the degree of weakening change in the acquisition process?, and (iii) how do these results complement each other? Our goal is to contribute to our understanding of how different phonological processes in an L2 are acquired.

The data for this study comes from a bigger project that combines pedagogy and research and gives our data some unique features. More precisely, the data comes from a teaching module developed for college-level Spanish Pronunciation courses at a major Midwestern university. As part of this module, students record themselves reading a list of words in isolation via a web-based interface and get instant feedback on their pronunciation via that interface. Students complete the module at the beginning (timepoint 1, T1) and the end of the semester (timepoint 2, T2), so we can compare their production of stops at two time points and analyze any changes as manifestations of the students' acquisition of the sounds. All in all, this integrated project presents a unique opportunity to combine teaching, data collection and research. For this study, we examined production data from 27 L2 learners who received the same Spanish pronunciation curriculum and teaching methodology. More precisely, tokens of /b, d, g/ in different contexts were acoustically analyzed for two dependent variables: type of allophone (i.e. voiced stop, voiceless stop, approximant, tap, etc.) and CV-intensity ratio. Linear and multinomial regression was used to examine the effect of timepoint (T1 vs. T2), stress, position within the word (medial vs. initial) and place of articulation on the two dependent variables.

For the type of allophone variable, the most common productions are voiced stop, approximant and voiceless stop. Our results indicate a significant increase in approximants, a decrease in voiced stops and a small change in voiceless stops in T2 compared to T1. As figure 1 shows, there is an effect of position within the word: most voiceless stops occur word initially, as expected given that this context correlates with utterance-initial position in our data and voiceless allophones of /b, d, g/ in English are very common in utterance initial contexts. This word-initial/utterance-initial effect decreases in T2 in favor of voiced stops and even some approximants (Fig. 1). However, the change in type of production is greater in word-medial contexts than in initial ones; for the former, the rate of approximants changes from 50% to 73% (Fig. 1). The smaller change in type of realization initially than word-medially could suggest that learning a new distribution of L1 allophones, i.e. voiceless productions, observable in initial contexts in our data, is more challenging than learning new allophones, manifested word-medially here. Further evidence for this is the occurrence of approximants word-initially in T2. This difference between learning a new distribution vs. new allophones has implications for theories of SLA. In relation to this, we explore issues of (dis)similarity and new allophones in the Speech Learning Model (e.g. Flege 1988) and of markedness in the Markedness Differential Hypothesis (e.g. Eckman 2008)

As for the CV-intensity ratio, this variable is higher in T2, in unstressed positions, and in word-medial positions, indicating that these are contexts with the highest degree of weakening. Interestingly, the effects of stress and word position mirror those found for native Spanish speakers (Carrasco et al. 2012). Based on these results, we conclude that learners not only become more native-like in their production of approximants but also in the degree of weakening depending on context. Our mixed methods approach allows us to observe patterns of acquisition that would otherwise be overlooked.

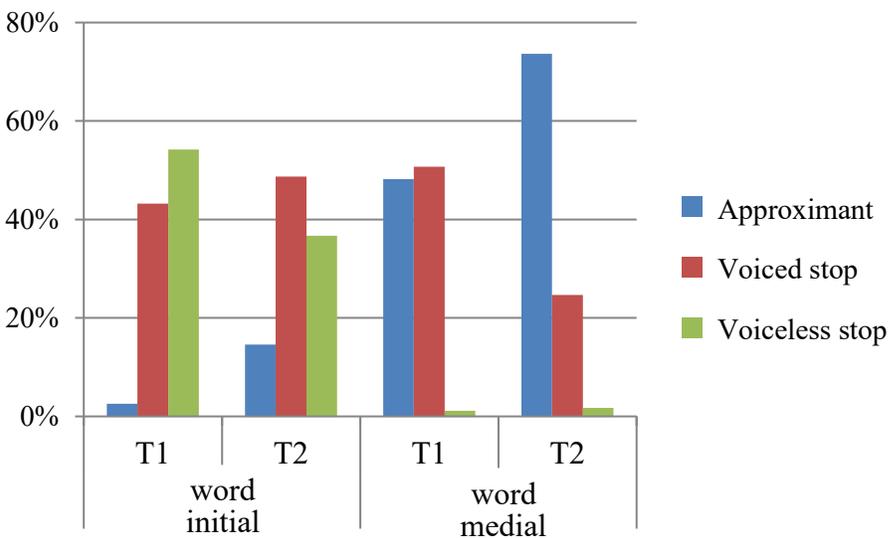


Fig. 1 Percentage of types of realization by timepoint & word position

The role of non-distinctive features in speech perception: voicing assimilation and the perception of contrasts in Spanish

Rebeka Campos-Astorkiza, *Ohio State University*

The role of assimilation in speech production has been explored especially for languages where the assimilated feature is distinctive for the target sound (Gow & Im 2004, Mitterer et al. 2013, Snoeren et al. 2006). In these cases, previous studies have found that perceptual compensation for assimilation allows listeners to retrieve the underlying form of the assimilated sound based on the phonetic context (Mitterer et al. 2013). When explaining these patterns, two approaches have been proposed, a language-specific mechanism that explains the perception of assimilation via language learning (e.g. Gaskell 2003); and a language-independent approach that argues that general phonetic perception mechanisms account for how assimilation plays into perception (e.g. Gow 2003). However, cases where the assimilated feature is not distinctive for the target sound have not been brought to bear on this discussion, and in fact, far less is known about perception in these cases (but see e.g. Meunier 1999). In addition, exploring this type of assimilation allows us to examine how the assimilatory process affects the perception of the triggering sound, rather than the target one, allowing us to contribute to the debate as to whether speech perception deals with invariant cues or rather makes use of variability in speech.

One such case is found in Spanish voicing assimilation of /s/ in pre-consonantal positions, an allophonic process by which /s/, which is not contrastive for voicing in the language, gradually assimilates in voicing to a following consonant (Campos-Astorkiza 2019). While this phenomenon has received recent attention from a production perspective (e.g. Sedó et al. 2020), the role that it might play in perception has not been explored. Understanding this would bring new data on the debates mentioned earlier. More precisely, Spanish voicing assimilation allows us to explore whether the allophonic voicing of /s/ plays a role in the perception of the voicing contrast of a following obstruent (stops are contrastive for voicing in Spanish). In addition, we investigate whether there are perceptual differences depending on the place of articulation of the following obstruent and whether the voiced vs. unvoiced allophones show different effects.

In order to answer these research questions, we conducted a forced-choice identification task where participants listened to a combination of two Spanish words, with the relevant assimilation context, i.e. /s/+obstruent, occurring across the two words, and they had to decide what words they had heard. The stimuli consisted of 4 versions of a word sequence that constitutes a minimal pair in voicing (e.g. /las **g**alas/ vs. /las **k**alas/ ‘the galas, the coves’). More precisely, we included two voicing matching versions: [s]+voiceless obstruent and [z]+voiced obstruent; and two voicing mis-matching versions: [s]+voiced obstruent and [z]+voiceless obstruent. To create the stimuli, a speaker of Spanish was recorded producing the relevant word sequences which were then manipulated, namely voiced and voiceless productions of /s/ where spliced into the 4 contexts/versions mentioned above. There were 40 test items (10 word combinations x 4 versions; see [1]) that included words with the three places of articulations for stops in Spanish (bilabial, dental and velar). The experiment also included 40 fillers that were formed by combinations of words that present other minimal pairs (e.g. /mis **k**anas/ vs. /mis **k**amas/ ‘my white hairs, my beds’). The differing sound in the fillers was also spliced to make sure that all items that participants listened to had been manipulated. The test items and fillers were divided into two balanced lists and each list was presented as an online survey via Qualtrics. In addition, the survey included questions about the sociolinguistic background of participants. Data from 120 participants from central and northern Spain were analyzed, for a total of 2,400 data points. Each

data point was coded for accuracy according to the voicing of the obstruent in the test item. The effect on accuracy of the stimuli structure ([s]+voiceless, [s]+voiced, [z]+voiceless, and [z]+voiced), the place of articulation of the stop, the participants' origin (central vs. north of Spain), and the interaction between stimuli structure and the other factors was explored with logistic regression models and pairwise comparisons in R.

The results show that stimuli structure, place of articulation (POA) and their interaction are significant factors on the degree of accuracy. The general trend is that the mis-matched stimuli show the lowest accuracy, with the lowest one for mis-matched [z]+voiceless, as can be seen in Table 1, with an accuracy rate of only 69%. The dental place of articulation presents the lowest rates of accuracy, compared to bilabial and velar places. The interaction between structure and POA comes from the fact that, while the pattern by stimuli structure is very similar for the three places of articulations, for the bilabial POA, the accuracy rate for the two mis-matched structures are more similar to each other than for the two other POAs. Summarizing, our results indicate that that voicing of /s/ plays a role in the perception of the voicing distinction for a following obstruent. The effect is greatest for mismatched [z]+voiceless. For the other mismatched case, [s]+voiced, the effect is not as robust. This pattern might arise from the fact that [s] sometimes occurs with a voiced obstruent in production but is extremely rare that [z] would occur with a voiceless obstruent (Campos-Astorkiza 2019, Sedó et al. 2020). This explanation of the asymmetry between [z] and [s] relies on the listeners' experience (cf. Mitterer et al. 2013) and aligns with approaches to perception and context-variability that put more weight on language-specific features (e.g. Gaskell 2003). In addition, this study shows that, while voicing is not phonemic for /s/, it plays a role in the perception of voicing for following obstruents, lending support to a view of speech perception that centers the role of variability and the manifestation of distinctive features across sounds.

(1) Test items

/las batas/ ~ /las patas/	‘the robes, the legs’
/las barkas/ ~ /las parkas/	‘the boats, the parkas’
/las bekas/ ~ /las pekas/	‘the scholarships, the freckles’
/los bojos/ ~ /los pojos/	‘the pastries, the chickens’
/las dunas/ ~ /las tunas/	‘the dunes, the music groups’
/las domas/ ~ /las tomas/	‘you tame them, you take them’
/las komas/ ~ /las gomas/	‘you eat them, the erasers’
/las kotas/ ~ /las gotas/	‘the levels, the drops’
/las kalas/ ~ /las galas/	‘the coves, the galas’
/las kasas/ ~ /las gasas/	‘the houses, the gauzes’

Table 1 Accuracy rates by stimuli structure

Stimuli structure		Incorrect	Correct
<i>Matching for voicing</i>	[s]+voiceless obstruent	7.17%	92.83%
	[z]+voiced obstruent	2.17%	97.83%
<i>Mis-matching for voicing</i>	[s]+voiced obstruent	9.67%	90.33%
	[z]+voiceless obstruent	31.00%	69.00%

PRO as the subject of generic null impersonals in Brazilian Portuguese

Janayna Carvalho, *Federal University of Minas Gerais*

Partial pro-drop languages differ from consistent pro-drop languages in at least two ways: (a) their referential null subjects (especially 3rd person singular ones) have a more restricted distribution; and (b) they have non-referential 3rd person singular null subjects. These properties have led Holmberg (2005) and Barbosa (2019), a.o. to propose that partial pro-drop languages have a minimally specified nominal (pro or a null noun) that either gets a referential reading when there is an available antecedent or a default generic reading. However, Brazilian Portuguese (BP), a partial pro-drop language, shows that the generic interpretation cannot be generated by default, as it distinguishes generic from existential null impersonals. The existential impersonal in (1), for instance, describes a habitual action in which the speaker does not take part, thus being odd when followed by *ou pelo menos eu vendo* ('or at least I do'), whereas the generic impersonal in (2) can be followed by *ou pelo menos eu posso* ('or at least I can'), as the speaker is part of the denotation of the null subject (see Carvalho 2019 for additional differences). Focusing on generic null impersonals (in BP), we argue in this paper that the generic reading can be accounted for if the referential subject is actually PRO based on Lebeaux's (1984) analysis of arbitrary PRO.

(1) *They sell candy in that store.*

(2) *Aqui pode vender bala.* (generic null impersonal)
here can sell.3ps candy

'Here one is allowed to sell candy.'

Lebeaux (1984) observes that not only controlled PRO, but also arbitrary PRO is anaphoric. In (3), from Lebeaux (1984:260), for instance, the two PROs (which do not c-command one another) must converge on the same denotation, which would be unexpected if they were referentially independent. Lebeaux argues that the interpretation of the PROs in (3) is captured if they are bound by a universal quantifier (or by a necessity operator introduced by modal verbs in other cases). Assuming that the minimal CP that contains the two PROs is the domain of binding, Lebeaux shows that if two PROs are minimally contained in different CPs, their reference will not be obligatorily linked because they may be bound by different operators, as exemplified in (4).

(3) PRO_i making a large profit (as a slum landlord) requires PRO_i exploiting the tenants.

(4) PRO_i winning the trust of the populace requires that PRO_{i/j} having to serve in the army be abolished.

In BP, generic null impersonals always involve genericity and/or deontic modality, the ingredients Lebeaux argued to be crucial in sentences with arbitrary PRO. We either have cases in which root modal verbs are used as in (2), or law-like statements, like (5) and (6). In generic sentences, a generic operator is assumed to be present (Carlson 1977, a.o.). This operator binds the null subject in generic null impersonals.

(5) *Como vende produtos digitais?*
how sell.3ps products digital.pl

'How is one supposed to digital products?'

(6) *Funciona assim: vende bala, ganha pouco dinheiro.*
works like.this sell.3ps candy make.3ps little money

'It works like this: if one sells candy, one makes little money.'

Consequently, generic null impersonals are banned from episodic embedded clauses. The null subject of the embedded clause in (7) can either have a referential or an existential reading. Generic

null impersonals are only licensed in embedded clauses if the complementizer ‘how’ or a deontic root modal is present, as in (8) and (9).

(7) João disse que vende essa casa naquela imobiliária.
 John said that sells this house in.that real estate.

‘John said that he/someone sells this house in that real state.’

(8) João disse como vende essa casa.

John said how sell this house

‘John said how one is supposed to sell this house.’

‘John_i said how he_i normally sells a house.’

(9) João disse que pode vender essa casa.

John said that can sell this house

‘John said that one is allowed to sell this house’

‘John_i said that he_i will probably sell the house.’

Both (8) and (9) are two-ways ambiguous sentences. Crucially, the non-referential reading arises whenever we have a deontic interpretation of the sentence, even if there is a potential antecedent in the main clause. In (8), if the embedded clause is interpreted as one describing generic norms, the null subject is non-referential. In (9), *poder* ‘can’ must be interpreted as a root modal, conveying permission, for the null subject to be non-referential. If interpreted as an epistemic modal, the null subject can only refer back to the matrix subject. Hence, (9) provides evidence that only permission seeking and permission granting modals select a CP with an operator since the generic reading only arises in the presence of this type of modal verb (see Doliana & Sundaresan to appear). The operator introduced by the modal verb is able to bind the null subject, precluding a coreference reading. These observations indicate that the generic null subjects of partial pro-drop languages are not due to a special type of nominal in these languages, but reflect a broader use of elements available to all languages.

Finally, (10) and (11) below replicate the pattern seen (3)-(4). Like arbitrary PRO, two instances of the generic null will obligatorily have the same reference if they are minimally contained by the same CP, as in (10). If they are minimally included in different CPs, as in (11), they do not coincide. (11) provides further evidence against the view that the putative minimally specified nominal may be interpreted either referentially or nonreferentially: If the null subjects of (11) are interpreted as referential, they obligatorily refer back to João, but under the non-referential interpretation, whoever cleans the house is different from whoever uses the barbecue grill.

(10) João disse como limpa parede e varre garagem.

John said how clean wall and sweep garage

‘John said how one_i should clean walls and Ø_i sweep garages.’

(11) João disse como limpa casa em que usa churrasqueira.

John said how clean house in which use barbecue.grill

‘John said how one_i should clean the house in which one_j uses the barbecue grill.’

While this paper focused in BP, PRO seems to be present in null impersonals in other languages as well. In both Icelandic and Capeverdean, some null impersonal constructions require a deontic modal verb. (cf. Sigurðsson & Egerland (2009) and Pratas (2007)). In light of Lebeaux’s account, such requirement indicates that the null subject must be bound by the operator introduced by the deontic verb. The discussion above leads to the conclusion that a core property of BP and possibly other partial pro-drop languages is that PRO is in principle available in both uninflected and inflected clauses, due to impoverishment of morphology.

Selected references: Barbosa, P. P. (2019). pro as a Minimal nP: Toward a Unified Approach to Pro-Drop. *LI*. • Holmberg, A. (2005). Is there a little pro? Evidence from Finnish. *LI*. • Lebeaux, D. (1984). Anaphoric Binding and the Definition of PRO. In *NELS Proceedings*.

Rethinking Romance vocatives: mapping address inversion in the nominal edge

Alice Corr, *University of Birmingham*

1. Introduction Address inversion (AI) is a typologically widespread phenomenon in which the lexical material used in a vocative address “reverses” the expected pattern indexing the speaker-addressee relationship:

- (1) Maria, du-te să cumperi lapte, **mamă!** (Ro. attested, online)
Maria go.IMP=REFL SBJV buy.2SG milk mummy
'Maria, go and buy some milk, **my dear [child]!**' (mother to daughter)

In (1), the utterance-final vocative does not function as a self-address, but instead expresses affection from mother to daughter, whence the gloss of the lexical material *mamă* (literally, ‘mummy’) as ‘my dear [child]’. This paper addresses the question of how this communicative function—in which the literal meaning of the lexical material is “reversed” in the vocative expression—is produced. Through comparative examination of the empirical patterns of Romance vocatives with and without AI, we propose that the non-literal, affective function of Romance AI is systematically configured by the architecture of grammar; specifically, a ‘topological’ GRAMMAR OF REFERENCE (Longobardi 2005; Sheehan & Hinzen 2011; Hinzen & Sheehan 2013) where an increase in syntactic structure co-varies with an increase in semantic reference. In so doing, we make the case for a conceptual rethinking of how we model vocatives, and, by extension, the grammar-discourse interface in the nominal domain.

2. Observations The fundamental referential function of vocatives with AI is to identify an object and undertake an action—viz. an expression of the speaker’s affective stance—in relation to the extensional referent of that object. In Romance, these vocatives diverge systematically from those without AI in their function, internal build and external syntax (Abbate 2010; Iovino & Rossi 2014). Amongst other diagnostics, AI-vocatives i) involve a lexical “flip” of their descriptive meaning in their vocative function, incurring an affective interpretation (1); ii) are disallowed—unlike imposters—in argumental contexts (2); and, where AI-vocatives co-occur (adjacently or discontinuously) with a ‘regular’ vocative, the latter *must* precede the AI-vocative (3a-b):

- (2) **Mămica** te iubeste mai mult ca orice în lumea asta. (Ro.)
mummy.DEF you=love.3SG more much than anything in world.DEF this
'{**Mummy/I/*your darling child**} loves you more than anything in the world'
- (3)a Forza dai va Olimpia dormi, a **nonna** (reg.It.)
INTJ INTJ INTJ Olimpia sleep.IMP PTC grandma
'C'mon Olimpia, go to sleep, **dear**' (grandmother to grandchild; Corr 2016:10)
- b *Forza dai va a **nonna** dormi, Olimpia (reg.It.)
INTJ INTJ INTJ PTC granny sleep.IMP Olimpia

Furthermore, in regional Italo-Romance, vocatives with and without AI exhibit contrasts in v) the licensing of truncation; and vi) syntactic gemination (*raddoppiamento fonosintattico*, RF):

- (4)a Me lo fai un caffè, (a) *zi'* / (a) *no*? (VOC_{ADDR} / *VOC_{SPKR})
'Can you make me a coffee, aunty (zia) / grandpa (nono)?'
b Me lo fai un caffè, a {*zzia/*zzi*} / a {*nonno/*nno*}? (*VOC_{ADDR} / VOC_{SPKR})
'Can you make me a coffee, **my dear [child]**?' (Iovino & Rossi 2014:220)

3. Framework Under the grammar of reference, words gain reference through insertion into the grammatical structure in line with a basic configurational template composed of a descriptive INTERIOR and a grammatical EDGE ([_{Edge} [INTERIOR]]), where expansion and movement into the edge correlates with an increase of referential strength. A further outcome of this theory is that the larger the grammatical structure, the more reliant that structure is on the grammar itself to produce its meaning, and the less reliant it becomes on the descriptive content of its lexical “core”. In the nominal domain, N-to-D movement (*Cerco [Gianni [Gianni]]* ‘seek.1SG Gianni’) yields a

maximally-specific ‘rigid’ DP, and deictically-anchored expressions are optionally (*Cerco [questa [macchina]]*) ‘seek.1SG this (car/one)’ or, in the case of person reference (*[Io Ø] cerco* ‘I seek’), *exclusively* reliant on interpretation at the phasal edge as a function of their referential strength.

4. Analysis Since on this theory personal names and pronouns respectively involve movement into and direct merge of morpholexical material in the nominal edge, evidence from ‘regular’ vocatives involving particle-N combinations (It. *O tu, O Gianni*, Moro 2003:252ff) requires us to posit extra structure at the phasal edge to accommodate the prenominal particle, yielding the template in (5a), which we generalise for all (Romance) vocatives (5b):

- (5) a [Edge O tu/Gianni [INT Ø/Gianni]] = [Voc O [D tu/Gianni [N Ø/Gianni]]] (It.)
 b [Edge mamă **mamă** [INT **mamă**]] = [Voc mamă [D **mamă** [N **mamă**]]] (Ro.)

A key advantage of the proposed template is that it simultaneously provides a means of modelling the lexical “flip” as a function of the grammatical encoding of the argumental vs. non-argumental distinction within the internal build of the vocative XP. On this view, the lexical content of N has the expected (literal) reading in its argumental position, but further movement leftwards enables it to incur a distinct, non-literal interpretation in vocatives both with (5a) and without (5b) AI:

- (6) a [Edge A mammë_{VOC} **mammë** [INT **mammë**]] = [Voc A [mammë_{VOC} [D **mammë**_Ø [N **mammë**]]]] (Pign.)
 ‘My child!’ (≠ ‘mother!’)
 b [D O [N rapaz/*pá]] vs. [Voc Ó [pá_{VOC} [D **rapaz**_Ø [N **rapaz**]]] (EuPt.)
 ‘the boy’ ‘(o) mate! (< rapaz ‘boy’)’

5. Extensions Furthermore, evidence from RF (4b) suggests that the distinction between vocatives with and without AI corresponds to a *phasal* distinction. Specifically, in some varieties the availability of RF (*A ppapà!* ‘My son!’ vs. *A papà!* ‘Daddy!’) is contingent on a local syntactic configuration—which recent research has shown to be a phasal domain (D’Alessandro & Biberauer 2006; Bošković 2016:34-36; Ledgeway 2018)—between the two constituents (here, Voc_{PTC} and Voc_N) on which the sandhi process acts. Accordingly, we can deduce that, in these varieties, the linearly-adjacent (particle+N) constituents are in a local, phasal configuration in “reversed” vocatives (which permit RF), but not ‘regular’ vocatives (which ban RF). This supports the proposal that the grammar of reference is simultaneously a *phasal* model of grammar, wherein the PHASES of recent syntactic theory instantiate referential-deictic units in the formal ontology of natural language (Arsenijević & Hinzen 2010; Sheehan & Hinzen 2011; Hinzen 2012).

Finally, we present evidence from truncation (4a) which enables us to conclude that ‘regular’ vocatives involve extra structure vis-à-vis vocatives with AI. Drawing on the TMT’s hypothesis that omission of morpholexical material is a function of the heaviness of the edge, we reconceptualize the licensing of truncation as a syntactic condition, viz. the expansion of the phasal template at its leftmost edge. On this view, vocatives permitting truncation are those that have moved even further into the edge, and are, as a consequence, less “reliant” on descriptive content. Since only ‘regular’ vocatives can be subject to this phonological process, they must involve extra structure (or movement thereto) in their internal build vis-à-vis AI-vocatives, the motivation for which, per TMT principles (cf. §3), is the need to gain the appropriate (extensional) reference.

Our findings not only support the hypothesis that interpretative distinctions in Romance vocatives are configurationally codified, but suggest that interlocutor distinctions have correlates in nominal syntactic structure (Ritter & Wiltschko 2019). By demonstrating how referential distinctions are mapped in the nominal left-edge, moreover, we provide *conceptual* motivation for the syntacticization of discourse in the nominal domain.

Selected references: Abbate, 2010. L’“allocuzione inversa” nell’italiano meridionale. *La comunicazione parlata* 3. Iovino & Rossi. 2014. Alcune riflessioni sul “Vocativo Inverso” tra pragmatica e sintassi. *Le mille vite del dialetto*. Longobardi. 2005. Toward a unified grammar of reference. *Zeitschrift für Sprachwissenschaft* 24. Martin & Hinzen. 2014. The Grammar of the Essential Indexical. *Lingua* 148. Ritter & Wiltschko. 2019. Nominal speech act structure: Evidence from the structural deficiency of impersonal pronouns. *The Canadian Journal of Linguistics* 64(4). Sheehan & Hinzen. 2011. Moving towards the Edge. *Linguistic Analysis* 37(3-4).

The active/passive alternation in French : a statistical approach to corpus data
Yanis da Cunha & Anne Abeillé, *Laboratoire de Linguistique Formelle, Université de Paris*

The use of passive varies across languages and genres (Bresnan et al. 2001). For English, Estival (1985) and Weiner & Labov (1983) show that priming and givenness play a role. Some corpus studies exist for spoken French (Blanche-Benveniste, 2000; Hama et al., 2017) and literary French (Attal, 1985), but they don't compare active and passive counterparts. We thus investigated the active/passive alternation with a written corpus, the *French TreeBank* (Abeillé et al., 2003, 2019, hereafter FTB) and three spoken corpora (C-Oral-Rom, CFPP2000, CRFP, hereafter SpkF) (Benzitoun et al., 2016). We first show that passive is sensitive to genre : we found 7.2% of passive verbs in the FTB (Roland et al. 2002 found 9% in the Wall Street Journal) vs 2.4% on average in SpkF (2% in the Switchboard corpus). Among these passives, 19.5% are long passives in FTB and 10.7% in SpkF, as in previous studies (Attal, 1985; Hama et al., 2017).

We selected perfect tense sentences, to avoid adjectival passives, and verbs with nominal or pronominal arguments. We took a random sample of 500 sentences from FTB and 400 sentences from SpkF, with half active (transitive) and half passive sentences in each sample. We took half long passive sentences (1) and half short passives (2), and correspondingly half active sentences with a nominal subject (3a) and half with a clitic subject (3b).

- (1) a. [Un agriculteur]_{ARG2} a été blessé [par une grenade lacrymogène]_{ARG1} (FTB)
b. [Il]_{ARG2} a été recruté [par une équipe d' Alfortville]_{ARG1} (CFPP)
- (2) Spk1 : Les indiens_i étaient absolument eux fascinés de voir que le Pape_j vienne [...] Spk2 : Oui j'avais entendu dire qu'[il]_j_{ARG2} avait été mal reçu [∅]_i_{ARG1} [...] (CFRP)
- (3) a. [Ce monsieur]_{ARG1} lui a refait [un studio tout à fait convenable]_{ARG2}(CFPP)
b. Et puis [ils]_{ARG1} avaient initialement négligé [les phénomènes monétaires]_{ARG2}(FTB)

Using Arg1 for the active subject (the passive *par*-phrase) and Arg2 for the active object (the passive subject), we annotated various factors : person, number, gender, animacy (Zaenen et al., 2004), saliency (Pronoun > Proper Noun > NP), definiteness, length (number of words). We also annotated verb classes, using LVF (Dubois & Dubois-Charlier, 1997), sentence type and the presence of a previous passive (Estival, 1985; Weiner & Labov, 1983). Most of the time, Arg1 could be recovered from the context in short passives, it is the Indians in (2).

We ran logistic mixed-effects models (Baayen, 2008) with Voice as dependent variable, and corpus, speaker and verb lemma as random effects. We found significant effects for 9 factors, including relativization, verb classes and syntactic priming (previous passive) ($p < 0.05$) (Table 1). We found that subjects (active Arg1 or passive Arg2) tend to be shorter and more pronominal, animate and definite than complements (active Arg2 or passive Arg1) (Comrie, 1979 a.o). Furthermore, our models show that spoken and written French don't assign the same weight to each factor. For example, the most important factor is definiteness in the FTB, while it is

length in SpkF. Some factors show significance only in the FTB, like Gender (a feminine Arg1 favoring passive) and communication verb (favoring passive), and others only in SpkF, like Person (a 1st and 2d person Arg1 favoring active and 1st and 2d person Arg2 favoring passive) as in Bresnan et al. (2001).

We conclude that the active/passive alternation obeys general constraints on constituent length (Wasow 2002), information structure (Givón, 1983; Lambrecht, 1994) and harmonic alignment (Bresnan et al., 2001), and that the passive should not be considered as suppressing an argument. Although the passive is less frequent in SpkF, we show that it follows the same tendencies as written French, but with different weights. Our study supports the importance of using quantitative methods to account for multifactorial syntactic phenomena across languages and language varieties.

Factor	Active (FTB+SpkF)	Passive (FTB+SpkF)	Coefficient (FTB+SpkF)	Coeff FTB	Coeff SpkF
Arg1 longer than Arg2	29%	71%	1.06	0.85	1.44
Pronominal Arg2	35%	65%		0.45	0.98
Animate Arg2	30%	70%	0.40	0.53	
Definite Arg1	71%	29%	-0.65	-0.51	-0.70
Relative pronoun Arg2	27%	73%	0.96		
Definite Arg2	42%	58%	0.85	1.05	0.80
Previous passive	34%	66%	0.42		0.40
<i>Hit</i> verb	24%	76%	0.36		0.55
Plural Arg1	42%	58%		0.61	
Plural Arg2	42%	58%	0.42	0.50	

Table 1 : Significant effects in the 900 sentences sample model and in the FTB-SpkF comparative models.

Selected references

- Abeillé, A., Clément, L., & Liégeois, L. (2019). Un corpus annoté pour le français : Le French Treebank. *TAL*, 60, 19 - 43.
- Attal, P. (1985). Le passif : Étude de corpus. *L'Information Grammaticale*, 27(1), 10 - 13.
- Benzitoun, C., Debaisieux, J.-M., & Deulofeu, H.-J. (2016). Le projet ORFÉO : Un corpus d'étude pour le français contemporain. *Corpus*, 15, 91 - 114.
- Bresnan, J., Dingare, S., & Manning, C. D. (2001). Soft constraints mirror hard constraints : Voice and person in English and Lummi. In *LFG Conference* (p. 13–32). CSLI Publications.
- Comrie, B. (1979). Definite and animate direct objects : A natural class. *Linguistica Silesiana*, 3, 13-21.
- Estival, D. (1985). Syntactic priming of the passive in English. *Text*, 5(1-2), 7–22.
- Hamma, B., Tardif, A., & Badin, F. (2017). Le passif à l'oral. In P. Larrivée & F. Lefeuve (Éds.), *Français contemporain vernaculaire (FRACOV)*.

**Translator Style as a Sociolinguistic Variable:
Variation in News Translation from English to Romance**
Justin Davidson and Mairi McLaughlin, *University of California, Berkeley*

The aim of this study is to explore the effects of translator style on the translation of global news from English into French and other Romance languages, building on research carried out in both Romance linguistics and translation studies on news agency dispatches. These dispatches are the building blocks of international news published and broadcast in French, Italian and Spanish. Strikingly, the translation of these dispatches is performed by just a select handful of translators working in the main international news agencies. Combining methods from corpus-based translation studies and variationist sociolinguistics, the study empirically tests for the first time the hypothesis that those responsible for the translation of news from English into Romance do not all translate in the same way, as was once supposed (Hernández Guerrero 2005: 159, Bielsa and Bassnett 2009: 69, Bielsa 2019: 366). The results of this study therefore shed new light on the translation of news into Romance, and critically establish translator style as a sociolinguistic variable that is mediated by both linguistic and social factors, in parallel with canonical production variables in the variationist sociolinguistics tradition (Labov 2001; Tagliamonte 2012).

In the wake of corpus linguistics, corpus-based translation studies has developed into both a central field and a central approach in translation studies. As a field, it is almost synonymous with the search for so-called “translation universals,” which were originally modeled on linguistic universals but are now often defined more loosely as “those recurrent features that make translation different from (or similar to) non-translated language production” (Bernardini and Kelly 2019: 112). Four main universals have historically been recognized: explicitation, simplification, normalization and levelling. Variation between individual translators, commonly referred to as “translator style”, has drawn considerably less attention. Henkel (2020: 2) explains that “while the variation from one author or translator to another is inherent in the very nature of corpus linguistics, this dimension remains absent from the equation in many corpus-based translation analyses”. This is definitely true of the form of specialist translation examined here, namely news translation. Working principally on French, Spanish, and English, scholars have been able to identify key features of news translation. For example, translated texts tend to be shorter than the originals (McLaughlin 2011: 21); those responsible for carrying out the translation consider themselves to be journalists rather than translators (Bielsa 2019: 366); and several linguistic features routinely appear, including omission, addition, replacement, summary, and restructuring (Bielsa and Bassnett 2009: 64). However, there has been no research on style in news translation and scholars are in agreement that there is no variation between individual translators in this domain (Hernández Guerrero 2005: 159, Bielsa and Bassnett 2009: 69, Bielsa 2019: 366). As Bielsa (2019: 366) explains, “the individual style of the journalist or news translator is necessarily sacrificed or neutralized in favour of a collectively defined style, usually embodied in the organization’s style book” and this means that “[n]ews translation must [...] be approached as an institutional process through which texts are collectively produced, with no single journalist bearing responsibility”. As far as we know, there is no robust empirical support for these claims: they tend to be impressionistic, based on scholars’ manual readings of translated news dispatches and their

understanding of the news industry. It is accordingly this gap in the literature that the present study aims to fill.

This study is based on a corpus of news dispatches which were translated from English to French and English to Italian by two of the five leading international news agencies. The corpus was constructed during two periods of ethnographic fieldwork in 2005 and 2006. Sociolinguistic interviews were carried out with twenty of the thirty-two journalists responsible for the translation of general and political news. In total, the corpus (which we shall preliminarily limit to French) contains the original English and translated French version of 989 dispatches which represent in total 558,000 words of which the French make up 238,000 words. We brought together studies of news translation in general (Hernández Guerrero 2005, Bielsa and Bassnett 2009, McLaughlin 2011, Davier 2017) and of translator style in other domains (Kenny 2001, Huang 2015) in order to select a series of linguistic and social variables to consider: dispatch length, number of sentences, ratio of sentences to dispatch length, ratio of quotation to dispatch length, and hapax legomena (as dependent variables), as well as translator gender, translator age, and individual translator (as independent variables).

Data were submitted to a combination of linear regression modeling in R, in order to empirically establish the loci of variation. Preliminary results are consistent with individual-level variation that contradicts the notion of a single, collective, and inherently static translator style. We contextualize the findings in the light of the competition between prescriptive or standardizing pressures and individual variation that exists in parallel across both translation and natural speech settings, which ultimately underscores the value of continued research on translator style as a bonafide sociolinguistic variable.

References

- Bernardini, Silvia and Dorothy Kenny. 2019. Corpora. In M. Baker and G. Saldanha (eds.), *Routledge Encyclopedia of Translation Studies*, 3rd edition, 110-115. New York, NY: Routledge.
- Bielsa, Esperança. 2019. News Translation. In M. Baker and G. Saldanha (eds.), *Routledge Encyclopedia of Translation Studies*, 3rd edition, 365-370. New York, NY: Routledge.
- Bielsa, Esperança and Susan Bassnett. 2009. *Translation in Global News*. New York, NY: Routledge.
- Davier, Lucile. 2017. *Les Enjeux de la traduction dans les agences de presse*. Lille: Presses universitaires du Septentrion.
- Henkel, Daniel. 2020. The Conditional Perfect, a Quantitative Analysis in English-French Comparable-parallel Corpora. In D. Iezzi, D. Mayaffe, and M. Misuraca (eds.), *Text Analytics - Advances and Challenges*, 179-197. Cham: Springer.
- Hernández Guerrero, María José. 2005. La Traducción de los géneros periodísticos. In C. Cortés Zaborras and M. J. Hernández Guerrero (eds.), *La Traducción periodística*, 89-133. Cuenca: Ediciones de la Universidad de Castilla-La Mancha.
- Huang, Libo. 2015. *Style in Translation: A Corpus-based Perspective*. Heidelberg: Springer.
- Kenny, Dorothy. 2001. *Lexis and Creativity in Translation: A Corpus-based Study*. London/New York: Routledge.
- Labov, William. 2001. *Principles of Linguistic Change, Vol. 2: Social Factors*. Oxford: Blackwell.

- McLaughlin, Mairi. 2011. *Syntactic Borrowing in Contemporary French: A Linguistic Analysis of News Translation*. Oxford: Legenda.
- Tagliamonte, Sali. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell.

Sociophonetic variation in the Spanish production of stressed and unstressed /e, o/ by Spanish-Galician bilinguals

Mónica de la Fuente Iglesias, *University of Texas at Arlington*

The long-standing situation of language contact and the typological similarity between Spanish and Galician, the two Romance languages spoken in northwestern Spain, has given rise to variation in the two languages. One of the linguistic features affected by this variation is the mid vowels, which is the focus of this study. Both languages differ in their vowel systems: Spanish has two phonemic mid vowels (/e, o/), whereas Galician has four phonological categories (/e, ε, o, o/), also preserved in other western Romance languages such as Catalan, French, Italian, and Portuguese. Given that previous studies on bilingual research (Flege, 1995, 2007) posit that even early bilinguals are susceptible to cross-linguistic phonetic transfer, the aim of this study is to examine whether highly proficient Spanish-Galician bilinguals transfer the Galician mid vowel contrasts into their Spanish. The few empirical studies that have investigated the mid vowels in Galician Spanish (De la Fuente Iglesias & Pérez Castillejo, 2020; Faginas Souto, 1998, 2004; López Bueno, 2017; Regueira & Fernández Rei, 2020) have exclusively focused on stressed position and the speech of urban speakers, where historically Spanish is more predominant, college speakers, or Galician-dominant speakers in a reading task. In order to obtain a more complete characterization of the variation of mid vowels in Galician Spanish, using a variationist approach (Labov, 2001; Tagliamonte, 2011), this study investigates the use and distribution of mid vowels in stressed and unstressed position produced by speakers with different language profiles and sociodemographic characteristics in spontaneous speech. Additionally, it examines whether variation is conditioned by several linguistic and social factors.

The data come from individual sociolinguistic interviews with 64 Spanish-Galician bilinguals residing in a semi-urban community. A total of 6,771 stressed and unstressed vowels were analyzed using Praat (Boersma & Weenink 2018), and three formant values (F1, F2 and F3) were taken from each vowel's midpoint with a script. Mixed linear models were elaborated in R (R Core Team, 2018) to test the effects of linguistic and social variables on the frequencies of the first two formants of each vowel at the group level. The linguistic factors included underlying phoneme in Galician (/e, ε, o, o/), vowel duration, and whether the word was a cognate in both languages. The social factors included participant's language dominance, age, gender, occupation, and level of education. In order to explore the magnitude of vowel contrasts individually, the effect of the linguistic and social variables on individual Pillai scores was tested through Pearson correlations and fixed linear models in R.

Findings show that the underlying phoneme in Galician influenced participants' vowel production in Spanish. Overall, the use of Galician mid vowels in stressed and unstressed position in Spanish was more likely to happen in the speech of Galician-dominant speakers, female speakers, speakers who have completed primary and secondary education, and speakers who work in manufacturing. Additionally, speakers who possess a Galician vowel system in stressed position demonstrated a robust contrast, manifested by formant height, formant frontness, and vowel duration. However, differences were found between the stressed and unstressed vowel system: in stressed position, the traditional Galician mid vowel system was observed in the majority of participants, whereas, in unstressed position, most participants showed a traditional Spanish mid vowel system. These results are interpreted using Flege's Speech Learning Model (1995, 2007) and by assuming a linguistic repertoires perspective (Recalde, 2012) that connects variation to social-indexical meanings (Coupland, 2007; Eckert, 2000, 2008). Implications of this study for language change in Galician Spanish are also discussed.

References

- Boersma P., & Weenink, D. (2018). Praat: doing phonetics by computer. Version 6.0.42, www.praat.org
- Coupland, N. (2007). The discursive framing of phonological acts of identity: Welshness through English. In C. E. Davies, J. Brutt-Griffler & L. Pickering (Eds.), *English and Ethnicity* (pp. 19–48). London: Palgrave.
- De la Fuente Iglesias, M. & Pérez Castillejo, S. (2020). Spanish mid vowels as sociolinguistic variables in Galicia. *Spanish in Context*, 17(3), 465–488.
- Eckert, P. (2000). *Linguistic variation as social practice*. Oxford, UK: Wiley-Blackwell.
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4): 453–476.
- Faginas Souto, S. (1998). As realización de /e/ no castelán da Coruña. *Cadernos de lingua*, 17, 83–103.
- Faginas Souto, S. (2004): Algúns aspectos do español da Coruña. As vogais [e] [o] tónicas em sílaba inicial. In R. Álvarez, F. Fernández Rei & A. Santamarina (Eds.), *A lingua galega, historia e actualidade, vol. I* (pp. 489–499). Santiago de Compostela: ILG/Consello da Cultura Galega.
- Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues* (pp. 233–277): Timonium, MD: York Press.
- Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. I. Hualde & Jennifer Cole (Eds.), *Laboratory Phonology 9* (pp. 353–382). Berlin: Mouton de Gruyter.
- Labov, W. (2001). The anatomy of style-shifting. En P. Eckert & J. R. Rickford (Eds.), *Style and Sociolinguistic Variation* (pp. 85–108). Cambridge, UK: Cambridge University Press.
- López-Bueno, L. (2017). Galician-Spanish mid vowel production in Southern Galicia: using language mode to explore bilingual speakers' underlying phonological categories (BA thesis, Cardiff Metropolitan University).
- RStudio Team (2018). *RStudio: Integrated Development for R*. RStudio, Inc. (versión: 1.1.463). Boston, MA. Available at <http://www.rstudio.com/>
- Recalde, M. (2012). Aproximación a las representaciones sociales sobre el español de Galicia. In T. Jiménez Juliá, B. López Meirama, V. Vázquez Rozas & A. Veiga (Eds.), *Cum corde et in nova grammatica: estudios ofrecidos a Guillermo Rojo* (pp. 667–680). Santiago de Compostela: Servizo de Publicacións e Intercambio Científico.
- Regueira, X. L., & Fernández Rei, E. (2020). *The Spanish sound system and intonation in contact with Galician*. In R. Rao (Ed.), *Spanish phonetics and phonology in contact: Studies from Africa, the Americas, and Spain* (pp. 328–362). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Tagliamonte, S. A. (2011). *Variationist Sociolinguistics: Change, Observation, Interpretation*. West Sussex, UK: Wiley-Blackwell.

Distinguishing between accounts of the A/ \bar{A} -distinction: the view from Argentinian Spanish Clitic Doubling

Suzana Fong, *MIT*

1 Overview. The A/ \bar{A} -distinction underpins case, agreement, and binding properties of moving DPs and it determines possible movement paths (cf. Ban on Improper Movement; BIM). This distinction has resisted an explanatory account, with the exception of the theories advanced by Van Urk:2015 and Safir:2019. In both, syntactic positions are not inherently A or \bar{A} . Rather, independent and more general properties of the grammar determine, as a byproduct, the nature of the movement that passes through these positions. Both approaches are able to explain the properties that the distinction is based on and they allow for a flexible definition of syntactic positions. I will argue that we can adjudicate between these two theories: Di Tullio et al.’s 2019 analysis of CD in Arg. Spanish will be shown to be compatible only with Van Urk:2015.

2 Clitic Doubling. Di Tullio et al.:2019 investigate CD, which is optional in Arg. Spanish (1).

Santos (CL.3SG.FEM.ACC) look.at.PST.3SG DOM Rosa
‘Santos looked at Rosa.’

[Di Tullio et al.:2019; (2a), adapted]

They observe that the presence or absence of the clitic is correlated with A- and \bar{A} -properties, respectively, of the CD-ed DP. If it moves across a subject containing a pronoun coindexed with it, no WCO effect is induced (2a). In the absence of CD, a WCO violation arises (2b).

(2) a. A MARÍA_k la_k criticó su_k padre.
DOM María CL.3SG.FEM.ACC criticize.PST.3.SG POSS.3SG father

‘Her father criticized MARÍA.’

[Di Tullio et al.:2019; (31/51), adapted]

b. *?A MARÍA_k criticó su_k padre.
DOM María criticize.PST.3.SG POSS.3SG father

A CD-ed DP does not reconstruct (3a). Without CD, reconstruction is possible (3b).

(3) a. *A su HIJO_k lo castigó cada padre_k.
DOM POSS.3SG son CL.3SG.MASC.ACC punish.PST.3SG each father

b. A su HIJO_k castigó cada padre_k.
DOM POSS.3SG son punish.PST.3SG each father

‘Each father punished his (own) SON.’

[Di Tullio et al.:2019; (58), adapted]

Di Tullio et al. propose that CD in Arg. Spanish is triggered by a [PERSON] feature in *v*. They assume that φ -features trigger A-movement (4). Given the properties of A-movement, the absence of WCO effects (2a) and of reconstruction (3a) displayed by CD-ed DPs can thus be modeled as consequences of the fact that this nominal is undergoing A-movement. Di Tullio et al. assume further that [PERSON] may be optional in 3rd person DPs in Arg. Spanish. A [PERSON]-less DP must then \bar{A} -move to Spec-*v*P (5). Consequently, this DP is expected to induce WCO effects (2b) and be able to reconstruct (3b). Di Tullio et al. assume that the clitic in Arg. Spanish CD is a morpheme introduced post-syntactically that expones the [PERSON] feature of an A-moved DP (4).

(4) [_{VP} DO_[PERSON] [_{v'} SUBJ [_{v'} v [_{VP} V t_{DO}]]]] (5) [_{VP} DO [_{v'} SUBJ [_{v'} v [_{VP} V t_{DO}]]]]

It is commonly assumed that a phase edge like Spec-*v*P is inherently an A-position. The proposal in (4), supported by the A-behavior of CD-ed DPs in Arg. Spanish challenges this assumption. Which view of the A/ \bar{A} -distinction could explain it?

3.1 Featureal view. According to Van Urk:2015, syntactic positions are defined in terms of the features that create them: A-positions are created by φ -features, while \bar{A} -positions are created by features like *Wh*, *Foc*, etc. As such, a syntactic position commonly assumed to be intrinsically

of the \bar{A} -type can be an A-position if it is created by φ -features. That Spec-CP is an \bar{A} -position, combined with the BIM, would be why a sentence like (6) is ungrammatical.

(6) * $[_{CP}$ Which students are $[_{TP}$ t_{Wh}] $[_{VP}$ believed $[_{CP}$ t_{Wh}] Terry will give a prize to t_{Wh}]]]?

However, if Spec-CP can be created by φ -features, it can also be an A-position. Fong:2019 shows that a similar movement to that in (6) has the properties of A-movement in Mongolian and yields a grammatical result. This type of flexibility is afforded by Van Urk's theory.

3.2 Free Merge view. The main ingredient in Safir:2019 is the proposal that a countercyclic Merge operation is freely available which combines a moving DP with a null preposition. The newly formed PP shields the DP away from Agree and case operations and alters the DP's binding capabilities. This operation is dubbed 'Insulation'. If Insulation does not apply, the DP can indeed participate in these operations and its binding capabilities remain unchanged. Insulation is a free, costless operation, with its effects regulated by independent principles. In the derivation of a sentence like *Who did Mary praise?*, Insulation may (7a) or may not (7b) apply to the moving *Wh*-phrase. If it does, T can Agree with the subject, allowing for the valuation of φ -features and case. If it does not, the moved *Wh*-phrase intervenes between T and the subject, preventing these Agree and case operations. The derivation in (7b) crashes not because \bar{A} -movement is postulated to be necessarily Insulated but because of independent case and Agree restrictions.

(7) a. $[_{TP}$ T $[_{vP}$ $[_{PP}$ P $[_{DP}$ who]] $[_{v'}$ Mary $[_{v}$ v $[_{VP}$ praise t_{DP}]]]]] \checkmark *T-Mary Agree*
 b. $[_{TP}$ T $[_{vP}$ $[_{DP}$ who] $[_{v'}$ Mary $[_{v}$ v $[_{VP}$ praise t_{DP}]]]]] **T-Mary Agree*

Spec-CP can also behave as an A-position in Safir's theory: it is possible for Insulation not to apply to a DP moving through this position, as long as Agree and case issues do not arise. The author shows that this indeed the case in passivized *wager* constructions.

(8) a. $[_{TP}$ T was $[_{vP}$ v $[_{VP}$ said $[_{CP}$ $[_{PP}$ P $[_{DP}$ the witch]] $[_{C'}$ C $[_{TP}$ t_{DP} $[_{T'}$ to be responsible ...]]]]]] \checkmark
 b. $[_{TP}$ T was $[_{vP}$ v $[_{VP}$ said $[_{CP}$ $[_{DP}$ the witch] $[_{C'}$ C $[_{TP}$ t_{DP} $[_{T'}$ to be responsible ...]]]]]] \checkmark

By assumption, the complement of a *wager* verb is a CP, which prevents ECM of the embedded subject. However, under passivization, this DP would have to escape the CP via Spec-CP, commonly considered to be \bar{A} -movement. If *the witch* is Insulated (8a), the derivation would crash, since this DP would remain caseless throughout the derivation and the matrix T's φ -features could be not be valued. If Insulation does not apply (8b), these requirements can be satisfied. In Safir's theory, movement through a phase edge (e.g. vP) is not intrinsically Insulated. Rather, Insulation is free, but its consequences are evaluated by independent restrictions.

4 Comparing the two approaches. Di Tullio et al.'s analysis can be cast as a representative of Van Urk's featural view of syntactic positions, since Spec- vP is considered to be an A-position created by the valuation of a φ -feature. Safir's Free Merge theory is also flexible in not assuming that syntactic positions are inherently A or \bar{A} . If Insulation applies to the moving CD-ed DP (9a), T can Agree with the subject across it. However, because the CD-ed DP is Insulated, it is expected to exhibit \bar{A} -properties, contrary to fact. If Insulation does not apply (9b), the CD-ed DP displays A-properties. However, this derivation crashes because T cannot Agree with the subject.

(9) a. $[_{TP}$ T $[_{vP}$ $[_{PP}$ P $[_{DP}$ a María]] $[_{v'}$ su padre $[_{v}$ v $[_{VP}$ criticó t_{DP}]]]]] \checkmark *T-su padre Agree*
 b. $[_{TP}$ T $[_{vP}$ $[_{DP}$ a María] $[_{v'}$ su padre $[_{v}$ v $[_{VP}$ criticó t_{DP}]]]]] **T-su padre Agree*

Safir's Free Merge theory provides two possible derivations of the Arg. Spanish CD data in §2. Both of them, however, yield undesirable case, Agree, or binding results.

5 Conclusion. There have been some attempts to derive the otherwise elusive A/ \bar{A} -distinction. For the most part, they are equally able to account for the empirical basis of this distinction. However, a particular analysis of Arg. Spanish CD may help us empirically distinguish between them.

The online comprehension of recomplementation in Spanish: Competing models and individual differences in working memory capacity considered

Josh Frank, *University of Texas at Austin*

Recomplementation is the phenomenon whereby one or more left-dislocated phrases or circumstantial adjuncts intervene between a primary (C1) and secondary (C2) complementizer, e.g., *Me preguntó que₁ esa camisa que₂ cuándo la iba a devolver* ‘S/he asked me when I was going to return that shirt’. Current syntactic-theoretical accounts argue that complementizer doubling is grammatical and that the functions of C1 and C2 differ, where C1 introduces the semantic function of the sentence and merges in the head of ForceP, while C2 marks a left-dislocation and merges in the head of TopicP (e.g., Villa-García, 2019). Further, evidence across paradigms converges on the relationship between C2 expression and intervener length (e.g., Casasanto & Sag, 2008; Echeverría & López Seoane, 2019). Specifically, C2 is a complexity correlate, where it is more likely to be spelled out when intervener length is long. Casasanto and Sag’s (2008) memory-based model adopts a distance locality theory framework. Curiously, it stipulates that overt C2 is the “ungrammatical reiteration” of C1, where complementizer doubling exists as part of a processing/grammar tradeoff when the distance between C1 and its complement is long (Figure 1).

Comprehension model and syntactic-theoretical account are not aligned with regard to grammaticality and function of C2. In order to further scrutinize this discrepancy, the present study attempts to replicate Casasanto and Sag’s (2008) interaction between intervener length (short versus long) and C2 expression (null versus overt) in Spanish, as well as consider how individual differences in working memory capacity (WMC) modulate online comprehension. Despite the clear prediction that a memory-based model makes-- i.e., comprehenders with lower WMC will benefit most from overt C2-- no recomplementation study has investigated the topic.

Methods. In a self-paced reading task (SPR), $n = 25$ native Spanish-speaking adults read recomplementation sentences of varying intervener lengths and C2 expression. Twelve stimuli were created for each of the four conditions with the critical and spillover regions (i.e., *wh*-word and object pronoun) held constant (Table 1). The stimuli from each experimental list were combined with forty-eight distractors of comparable length using a pseudo-randomization design and a counterbalanced yes/no comprehension question preceded each test item. Lastly, participants completed a visual reverse digit span task to assess WMC. **Results.** Linear mixed models specifying intervener length and C2 expression, an interaction of intervener length*C2 expression and one random intercept for subject were run with the LMER function in R. Intervener length was highly significant ($\beta = -97.88$, $t(813.46) = -3.446$, $p = 0.0005$) and the interaction of length*C2 approached significance ($\beta = 77.95$, $t(813.48) = 1.925$, $p = 0.054$). A pairwise comparison with bonferroni adjustment showed that C2 expression significantly facilitated complement integration in the long condition ($\beta = -76.14$, $t(814) = -2.626$, $p = 0.043$) but not in the short one ($\beta = 1.82$, $t(814) = 0.064$, $p = 0.999$) (Figure 2). Lastly, a series of Pearson correlations run to investigate the effect of individual differences in WMC returned a moderately strong and significant negative correlation in the overt/long condition ($r = -0.475$, $p = 0.014$) and a positive correlation in the difference between short and long conditions (Figures 3 and 4, respectively). **Conclusions.** The memory-based model of recomplementation is not supported by our results. Namely, (i) the shape of the interaction predicted by the model was not successfully replicated and (ii) participants with *higher* WMC benefit most from overt C2. We propose an alternative comprehension model that is faithful to the data and compatible with syntactic-theoretical accounts. The expectation-based account puts forth that conventional patterns of C2 expression can be probabilistically predicted along a length-of-intervener

constraint and ease of complement integration depends on the degree to which the input aligns with expectations. It predicts that comprehenders with higher WMC have more resources to allocate to the prediction of the C2 complexity correlate. Importantly, the model need not stipulate that a systematically produced phenomenon is anomalous. Rather, it is compatible with current trends in recomplementation theory (e.g., biclausality and CP resetting), while offering a counterintuitive account of recomplementation convention.

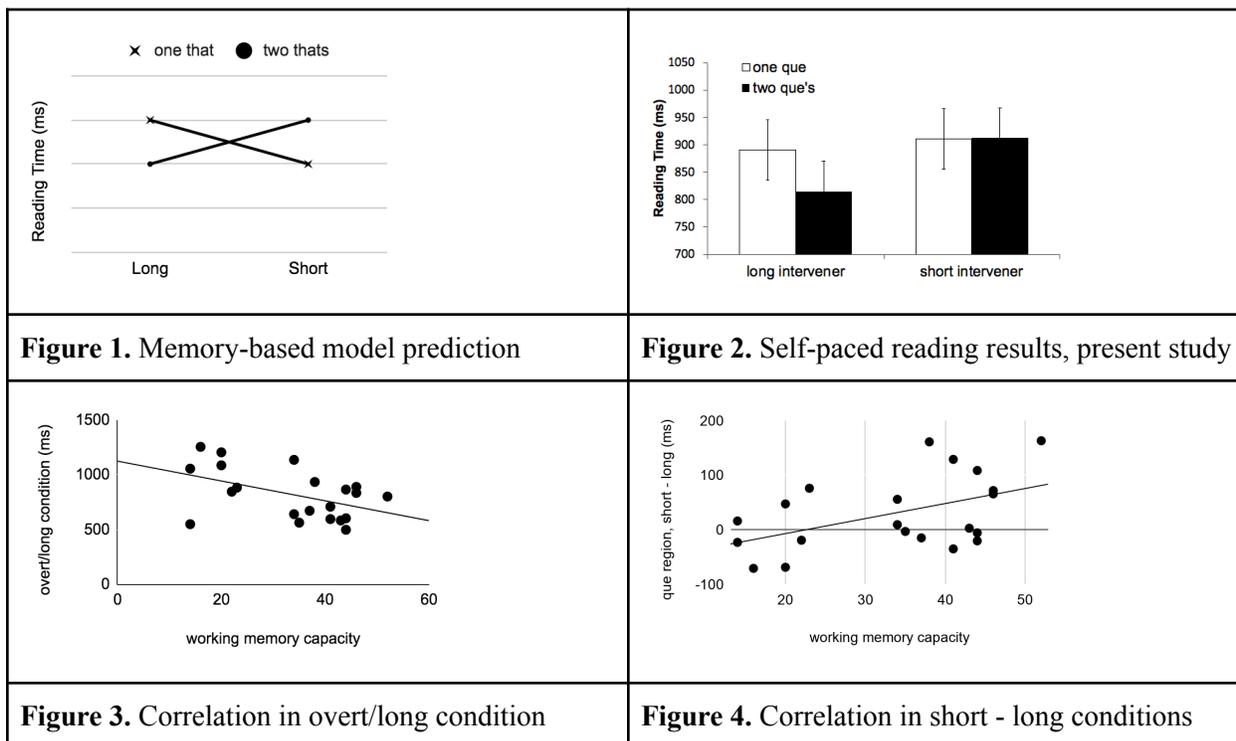


Table 1. Conditions for SPR with moving window paradigm and word-by-word segmentation

Conditions	main clause	IV	IV	DV	wrap-up effects
1. overt/long	/Ella/me/pregunta/que	/ese/ensayo/sobre/mi/historia	/que	/cuándo/lo	/publico/en/una/revista
2. null/long	/Ella/me/pregunta/que	/ese/ensayo/sobre/mi/historia		/cuándo/lo	/publico/en/una/revista
3. overt/short	/Ella/me/pregunta/que	/ese/ensayo	/que	/cuándo/lo	/publico/en/una/revista
4. null/short	/Ella/me/pregunta/que	/ese/ensayo		/cuándo/lo	/publico/en/una/revista

Selected References:

Casasanto, Laura, and Ivan Sag. 2008. The Advantage of the Ungrammatical. In Proceedings of the 30th Annual Conference of the Cognitive Science Society. Edited by B. Love, K. McRae and V. Sloutsky. Austin: Cognitive Science Society, pp. 601–6.

Echeverría, Carlos, and Javier López Seoane. 2019. Recomplementation as constrained CP resetting. Paper presented at the Association of Linguistics and Philology of Latin America (Alfalito), Queens College, City University of New York, New York City, NY, USA, September 20.

Villa-García, Julio and Dennis Ott. 2019. A biclausal analysis of recomplementation. Paper presented at Going Romance 2019, Leiden University.

Villa-García, Julio. 2019. Recomplementation in English and Spanish: Delineating the CP space. *Glossa: A Journal of General Linguistics* 4: 56.

CONSTRAINTS ON V-IO-DO AND V-S-DO IN ROMANCE IOs AS “VP SUBJECTS:” A LABELING THEORY APPROACH

Ángel J. Gallego (Universitat Autònoma de Barcelona)

1. PROPOSAL: We claim that the limited availability on V-IO-DO in Romance (Ormazabal & Romero 2013 and references therein) should be paralleled with that of V-Subj-DO (Gallego 2013, Ordóñez 2007). We argue that the facts follow from a general principle that requires that only one of the elements establishes a AGREE / MOVE (=IM) dependency with v^* . We phrase the solution in terms of Chomsky’s (2013, 2015) Labeling Theory (LT), which we adjust to exclude exocentric {XP,YP} structures. The connection between the constraints on V-Subj-DO and multiple object construction is not new (cf. Alexiadou & Anagnostopoulou 2007): what *is* new is the connection between the behavior of IO and Subj in V-[IO-DO] and V-[Subj-DO] respectively, which we attribute to the possibility to generate V-DO-Subj through object shift (Ordóñez 1998, Gallego 2013).

2. THE PROBLEM: It has been reported that V-Subj-DO sentences are restricted in Romance (Alexiadou & Anagnostopoulou 2001; 2007, and references therein), and it seems that the languages allowing V-Subj-DO are those that generate V-DO-Subj through “object shift” (Ordóñez 1998, Gallego 2013), as binding effects indicate:

- | | |
|---|---|
| <p>(1) Recogió cada coche su propietaria
 picked-up every car its owner
 ‘Its owner picked up each car’</p> | <p>(2) Recogió la alumna su coche (Spanish)
 picked-up the student her car
 ‘The student picked up her car’</p> |
|---|---|

Similarly, both V-DO-IO and V-IO-DO sentences have been shown to be subject to non-trivial constraints, regulated by phenomena like DOM, clitic doubling, NP heaviness, or the strong pronoun status of the relevant arguments (Anagnostopoulou 2003, Lopez 2012, Ormazabal & Romero 2007). Just like V-Subj-DO, V-IO-DO is degraded in Catalan, Italian, and French (only allowed under conditions like NP heaviness; Fournier 2010, Beavers & Nishida 2010, Bleam 2003):

- | | |
|--|------------------|
| <p>(3) */?Vaig recomanar a la Maria els teus estudiants
 AUX-1.sg recommend to the Maria the your students
 ‘I recommended your students to Maria’</p> | <p>(Catalan)</p> |
| <p>(4) ??Ho raccomandato a Giovanni il mio amico
 have-1.sg recommend to Giovanni the your friend
 ‘I recommended my friend to Giovanni’</p> | <p>(Italian)</p> |
| <p>(5) */?J’ai recommandé à Pierre Jean
 I-1.sg recommended to Pierre Jean
 ‘I recommended Jean to Pierre’</p> | <p>(French)</p> |

The same deviance is NOT found in Spanish and Romanian:

- | | |
|--|-------------------|
| <p>(6) He recomendado a María a tus estudiantes
 have-1.sg recommended to Maria the your students
 ‘Have recommended your students to Maria’</p> | <p>(Spanish)</p> |
| <p>(7) I l-am recomandat Mariei pe Ion
 i-cl.DAT I-cl.ACC recommended Mary-DAT ACC John
 ‘I recommended John to Mary’</p> | <p>(Romanian)</p> |

It is well-known that having a ‘too crowded’ VP is problematic, but the specific *connection between IO and Subjects* (would-be “specifiers” in X-bar terms) has not been established so far. Making such correlation makes sense if we consider the similarities between those dependents: IOs and Subjects can be doubled (in the case of subjects, by verb agreement), IO and Subject doubling is compatible with negative quantifiers, IOs and Subjects are typically analyzed as specifiers of sorts, IO and Subjects display island effects, IOs and Subjects fail to affect the lexical aspect of verbs, IOs and Subjects occupy the first position in clitic clusters with DOs, and so on. Given these empirical correlations, trying to establish a correlation between V-[IO-DO] and V-[Subj-DO] makes sense.

3. (SOME) SOLUTIONS SO FAR: We don’t know of accounts tackling the symmetry between V-IO-S and V-Subj-DO, but we do know of proposals that try to explain the problems that V-DO-IO / V-IO-DO pose. Alexiadou & Anagnostopoulou (2007) propose a *Multiple Case Condition* (MCC), which states that VP cannot contain more than one argument with an undeleted Case feature. The MCC is interestingly compatible with Ormazabal & Romero’s (2007) *Object Agreement Constraint* (OAC), which amounts to saying that v^* can only establish one

Clausal Determiner as Inherent Case: Evidence from Spanish

Ángel Gallego & Cristina Ruiz Alonso, *Universitat Autònoma de Barcelona*

1. THE MAIN IDEA: This paper argues that the definite article preceding embedded clauses in Spanish (Leonetti 1999, Picallo 2002, Serrano 2008, among others) is a mark of inherent Case. We take this: (i) to be an argument in favor of the nominal nature of clauses (Plann 1986, Picallo 2002, a.o.), and (ii) to account for the behavior (distribution, island effects, semantic constraints, etc.) of embedded clauses headed by the definite article.

2. THE FACTS: Spanish embedded clauses can *optionally* be introduced by a definite article, which we refer to as “el(que)ismo” (EL(Q), henceforth):

- (1) a. Nos gusta [que digas eso] b. Nos gusta [**el** [que digas eso]]
CL-we like-3.sg that say-2.sg that CL-we like-3.sg the that say-2.sg that
'We like that you say that' 'We like that you say that'

Empirically, the distribution and consequences of EL(Q) are manifold. First, it is by and large preferred in subject positions (cf. Uriagereka 1988):

- (2) a. Nos gusta [el que digas eso] b. *Quiero [**el** [que digas eso]]
CL-we like-3.sg the that say-2.sg that want-1.sg the that say-2.sg that
'We like that you say that' 'I want that you say that'

Third, the embedded verb is inflected in subjunctive: indicative is ruled out.

- (3) a. [El [que sufras]] me preocupa b. * [**El** [que sufres]] me preocupa
the that suffer-sub.2.sg CL-me worry-3.sg the that suffer-ind.2.sg CL-me worry-3.sg
'That you suffer worries me' 'That you suffer worries me'

Fourth, EL(Q) is ruled out in ECM contexts, both with perception and causative verbs.

- (4) a. *Ana vio [**el** salir a Juan] b. *Ana hizo [**el** salir a Juan]
Ana saw-3.sg the go-out-inf DOM Juan Ana made-3.sg the go-out-inf DOM Juan
'Ana saw Juan getting out' 'Ana made Juan get out'

Fifth, control and volitive predicates reject EL(Q) too:

- (5) a. Ana quiere [(***el**) que vengas] b. Ana planea [(***el**) viajar]
Ana want-3.sg the that come-sub-1.sg Ana plan-3.sg the travel-inf
'Ana wants for you to come' 'Ana plans to travel'

Sixth, EL(Q) turns otherwise transparent CPs into islands, blocking A/A-bar movement:

- (6) a. Ana parece [(***el**) <Ana> dormir] b. Qué te molesta [(***el**) que digan <qué>]?
Ana seem-3.sg the sleep.inf what CL-you bother-3.sg the that say-sub.3.pl
'Ana seems to sleep' 'What does it bother you that they say?'

3. PROPOSAL: The evidence above suggests that EL(Q) is the spell-out of Case morphology, not a true determiner. Given the island status and the reluctance to show up in first-Merge (complement) position, we take such Case to be inherent. For concreteness, we assume the following structure, with “el” occupying a K head (Bittner & Hale 1986):

- (7) [K (el) [C [. . .]]]

This proposal makes a series of empirical predictions. Let us consider some of them.

4. PREDICTIONS: If K corresponds to inherent Case, a prediction consistent with the data in (2), (4), and (5), is that EL(Q) should be barred in verbal complexes (periphrases). The data in (8) show that this is borne out:

- (8) a. *Ana puede [**el** [cantar]] b. *María empezó a [**el** [cantar]]
 Ana can the sing-inf María started to the sing
 ‘Ana can sing’ ‘Ana started to sing’

We take these facts to follow from a constrain on EL(Q) appearing in first-Merge (complement) position. This is further supported by the fact that prepositions cannot take clauses introduced by the definite article:

- (9) Alba se alegra de [(***el**) [que tengas suerte]]
 Alba SE be-happy of the that have luck
 ‘Alba is happy that you are lucky’

Finally, notice that EL(Q) is also incompatible with embedded interrogatives, in both object and subject position:

- (11) a. No sabemos [(***el**) [quién vendrá]] b. [(***El**) [quién vendrá]] es una incógnita
 not know the who will-come the who will-come is a mystery
 ‘We don’t know (the) who will come’ ‘(The) who will come is a mystery’

One final empirical prediction made by the previous observations is that if we can force EL(Q) to occupy a non-complement position (a specifier) derivationally, then the problem with first-Merge positions should go away. This is borne out with cleft sentences, as already pointed out in Leonetti (1999):

- (12) a. *Quiero [**el** [que digas eso]] b. [**El** [que digas that]] es lo que quiero
 want the that say that the that say that is it that want
 ‘I want that you say that’ ‘It is that you say that what I want’

5. SOME PENDING QUESTIONS: Our analysis must address various questions. Theoretically, we should at least clarify if the presence of K (and, more generally, the nominal status of clauses) is universal. On the empirical side, we must explain why certain predicates, including factive verbs, can apparently license EL(Q) (Serrano 2008). Our answer is that they deploy a more complex structure, involving a predication (small clause) that turns the clause into a specifier (hence, a non-complement). We should also be able to explain why EL(Q) is, by and large, restricted to Spanish within the Romance family.

6. CONCLUSIONS: This paper has investigated EL(Q) in Spanish, a phenomenon that has not received much attention in the literature. We have argued that the definite article is the spell-out of a Case projection, which we have taken to be inherent. Our proposal makes a series of empirical predictions concerning the incompatibility between EL(Q) and complement (first-Merge) positions.

REFERENCES (selected): Leonetti, M. 1999. “El Artículo”. In I. Bosque and V. Demonte (eds.), *Gramática Descriptiva de la Lengua Española*, 787-890. Madrid: Espasa Calpe.
 Picallo, M.C. 2002 “Abstract Agreement and Clausal Arguments”, *Syntax* 5: 116-14.
 Serrano, S. 2008. Subordinación y determinación: completivas precedidas de artículo definido en español. PhD Dissertation. UAM.
 Uriagereka, J. 1988. On Government. PhD Dissertation, UConn.

**The lexis of Sardinian within the Romance scene:
A minority-language case with valuable insights into historical linguistics**
Marc Gandarillas, *University of Florida*

Abstract

Our study delves into traditional claims of uniqueness regarding Sardinian lexical items that derive from their Latin etyma counterparts. Being an understudied language within the Romance scene (especially since the 1950s), Sardinian shows a strong influence of other Romance languages (especially Spanish, Catalan, and—more recently—Italian), yet it appears to remain largely unaffected by shifts typically observed in neighboring languages. With Sardinian morphology traditionally being regarded as conservative (Meyer-Lübke, 1890), in this paper we explore the domains of lexis and semantics in the light of Wagner (1951) and Porru (1942), who selects a reduced sample of etyma and includes a brief etymological and semantic explanation thereof. In more recent times, Stefenelli (1992: 92-93) identifies a small number of Latin etyma that appear to be found only in Sardinian (e.g., FĒRRE ‘to carry’). The following research questions were established: (1) does Sardinian conservatism apply to lexis?; (2) does the scope of our research have further implications in the study of underrepresented languages?

150 Latin etyma were identified as having survived only in Sardinian. Our revision proved decisive in determining inaccuracies in Porru (1942). Additionally, errors concerning the historical reconstruction of etyma were detected, especially in those cases in which the Sardinian item appeared to be rooted in the Spanish and/or Catalan presence on the island (Sard. *afficare* ‘to build’ < anc. Sp. *aficar*). Each etymon was assigned an individual entry including: (a) its Sardinian outcome(s); (b) its earliest documentation and semantic scope; (c) a possible explanation for morphophonological and semantic shifts. Our results visibly reinforce Wagner’s (1951) focus on lexis. A discussion follows on further challenges that arise from the study of a minority language in contact.

References

- Adams, J. N. (2007). *The Regional Diversification of Latin, 200 BC - AD 600*. Cambridge: Cambridge University Press.
- Bastardas, M. R., & Gargallo, J. E. (2007). *Manual de lingüística románica*. Barcelona: Ariel.
- Blasco Ferrer, E. (2008). "Storia interna del sardo", *Romanische Sprachgeschichte. Histoire linguistique de la Romania* (ed. Ernst, G. et al.) III. Berlin: De Gruyter.
- Blasco Ferrer, E. (1984¹). *Storia linguistica della Sardegna*. Tübingen: Niemeyer.
- Blasco Ferrer, E. (1984²). "Sull'italianità linguistica del gallurese e del sassarese", *Romanian Review of Linguistics* 29, pp. 399-413. București: Editura Academiei Române.
- Clackson, J., & Horrocks, G. (2007). *The Blackwell History of the Latin Language*. Malden: Blackwell Publishing.
- Contini, M. (1987). *Étude de géographie phonétique et de phonétique instrumentale du sarde* (2 vols.). Alessandria: Edizioni dell'Orso.
- Cortelazzo, M., & Marcato, C. (1998). *I dialetti italiani: dizionario etimologico*. Torino: Unione Tipografico-Editrice Torinese.
- Diez, F. (1869-1870). *Etymologisches Wörterbuch der romanischen Sprachen* (5^a ed., 2 vols.). Bonn: Adolph Marcus.
- Ernout, A., & Meillet, A. (2001). *Dictionnaire étymologique de la langue latine* [reimpr. 4^a ed. a cargo de André, J.]. Paris: Klincksieck.
- Guarnerio, P. E. (1892-1893). "I dialetti odierni di Sassari, della Gallura e della Corsica", *Archivio Glottologico Italiano* (vol. 13, pp. 125-140; vol. 14, pp. 131-200/385-422). Firenze: Le Monnier.
- Holtus, G., Metzeltin, M., & Schmitt, C. (eds.) (1988-2005). *Lexikon der romanistischen Linguistik (LRL)*, vol. 4 (italiano, corso y sardo). Tübingen: Niemeyer.
- Lausberg, H. (1976-1982). *Lingüística románica* [2 vols.; trad. a cargo de Pérez Riesco, J. y Pascual Rodríguez, E.]. Madrid: Gredos.
- Meyer-Lübke, W. (1890). *Grammaire des langues romanes (GLR)* [4 vols., trad. Rabiet, E.]. Paris: E. Welter.
- Meyer-Lübke, W. (1861-1936). *Romanisches etymologisches Wörterbuch* [7^a ed., 2009]. Heidelberg: Universitätsverlag Winter.
- Pittau, M. (2000-2003). *Dizionario della lingua sarda: fraseologico ed etimologico* [2 vols.]. Cagliari: Gasperini Editore.
- Pittau, M. (1991). *Grammatica della lingua sarda*. Sassari: Delfino Editore.
- Porru, G. (1942). "Saggio di uno studio sul lessico sardo", *Studi Italiani di Filologia Classica* XIX, pp. 103-124. Firenze.
- Porru, V. (2002). *Nou dizionariu universali sardu-italianu* [ed. a cargo de Lörinczi, M., en 3 vols.: A-C / D-O / P-Z]. Nuoro: Ilisso Edizioni.
- Puddu, M. (2000). *Ditzionàriu de sa limba e de sa cultura sarda (DLCS)*. Cagliari: Condaghes.
- Sanna, M. (2008). *Reimparare il sardo oggi / Re-imparare su sardu oje*. Nuoro: Edizioni Solinas.
- Spano, G. (1873). *Vocabolario sardo geografico, patronimico ed etimologico*. Cagliari: Antonio Alagna.
- Spano, G. (2004). *Vocabolariu sardu-italianu* [ed. a cargo de Paulis, G.]. Nuoro: Ilisso Edizioni.
- Stefenelli, A. (1992). *Das Schicksal des lateinischen Wortschatzes in den romanischen Sprachen*. Passau: Wissenschaftsverlag Richard Rothe.
- Wagner, M. L. (1931). "Das Sardische in der 3. Auflage von Meyer-Lübkes REW (Lieferung 1-3)", *Archiv für das Studium der neueren Sprachen und Literaturen*, 160 pp. 228-39. Braunschweig: Georg Westermann Verlag.
- Wagner, M. L. (1978). *Dizionario etimologico sardo (DES)* [3 vols.]. Cagliari: Edizioni Cagliari.
- Wagner, M. L. (1984). *Fonetica storica del sardo* [ed. y trad. a cargo de Paulis, G.]. Cagliari: Trois.
- Wagner, M. L. (1951). *La lingua sarda: storia, spirito e forma*. Berna: A. Francke S. A.
- Wagner, M. L. (1996). *La vita rustica della Sardegna riflessa nella lingua* [trad. a cargo de Paulis, G.]. Nuoro: Ilisso Edizioni.
- Wagner, M. L. (1947). "Porru, G., «Saggio di uno studio sul lessico sardo»", *Vox Romanica* IX, pp. 299-302.
- Wagner, M. L. (1935-1940). "Rettifiche et aggiunte alla terza edizione del REW del Meyer-Lübke", *Archivum Romanicum* (XIX pp. 1-29 / XX pp. 343-358 / XXIV pp. 11-67). Firenze: Leo S. Olschki.

Target vowel asymmetry in Brazilian Veneto metaphony

Guilherme D. Garcia (*Ball State University*) & Natália B. Guzzo (*McGill University*)

Overview. Metaphony targeting upper mid vowels /e, o/ is a characteristic of Central Veneto (e.g., Zamboni 1974, Belloni 2009, Walker 2010, Perrone 2016), a dialect of Veneto spoken in northeastern Italy. In a closely related understudied dialect spoken in southern Brazil, namely Brazilian Veneto (locally known as Talian), metaphony is also observed (Frosi and Mioranza 1983). Although the phenomenon is reported as variable for both dialects, little is known about how such variation is structured. In this paper, we explore the structural conditioning of metaphony in Talian through a corpus study. We show that metaphony in this dialect is asymmetrical, as it applies at different rates for /e, o/. We formalize this asymmetry using a MaxEnt Grammar (Goldwater & Johnson 2003, Hayes & Wilson 2008).

Background and methods. In Talian, like in Central Veneto, stressed /e, o/ variably raise to [i, u] when followed by unstressed /i/. The phenomenon targets all stress positions (1), and it is morphologically conditioned, since the trigger is a separate morpheme (a plural marker or verbal inflection). In antepenultimate position (1b), the unstressed vowel in penultimate position also raises. In final position (1c), metaphony precedes resyllabification.

- | | | | |
|-----|-----|-------------------|-------------|
| (1) | (a) | óv-i ~ úv-i | ‘egg.PL’ |
| | | bév-i ~ bív-i | ‘drink.2PS’ |
| | (b) | zóven-i ~ zúvin-i | ‘young.PL’ |
| | (c) | fazó-i ~ fazú-i | ‘bean.PL’ |

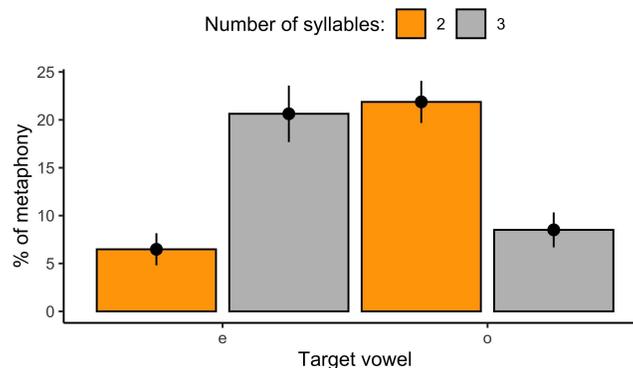
To examine how widespread metaphony is in the language and which linguistic factors condition it, we analyze data from a corpus of written narratives in Talian (the Talian Corpus, *Anonymous*, publicly available). This corpus contains narratives extracted from books and newspapers written by different Talian-speaking authors ($n = 20$). Since Talian has no standardized orthography, orthographic variation may reflect at least in part variation in the authors’ spoken language. Therefore, if a given author writes *uvi* as opposed to *ovi* ‘egg.PL’, we assume that said author may regularly produce metaphony in his/her speech, or that he/she identifies metaphony as part of the Talian grammar.

For the development of the corpus, the narratives were digitized (OCR, optical character recognition) using Google’s Tesseract (Smith 2007), and later coded for a wide range of phonological variables in R (R Core Team 2020), including broad IPA transcription and syllabification. In its current state, the Talian Corpus contains 103,046 words and 8,589 sentences.

For the present study, we initially extracted from the corpus all polysyllabic words ending in /i/ with an underlying /e, o/ in stressed position ($n = 1,462$). Given the low number of tokens with antepenultimate or final stress, we focus on 2- and 3-syllable words with penultimate stress ($n = 992$; n (unique) = 300). Examples include items such as *sentí* ‘feel.2PS’, *amóri* ‘love.PL’, *oví* ‘egg.PL’, *corrí* ‘run.2PS’, and *cagnetí* ‘dog.DIM.PL’, with or without metaphony.

Results and discussion. Figure 1 shows that target vowel and word length interact: while /o/ raises more often in 2-syllable words (22% vs 8.5%), /e/ raises more often in 3-syllable words (20% vs 6.5%). This interaction is statistically confirmed in a mixed-effects logistic model containing a by-author random intercept ($\hat{\beta} = -4.0113$, 95% CI = $[-5.1, -2.9]$, $p < 0.0001$). Additional predictors, such as onset and coda of target syllable and onset of trigger syllable, showed no statistical effect.

Figure 1. Percentage of metaphony by target vowel and word length.



The vowel asymmetry seen in the data is potentially driven by multiple factors. One such factor is lexical statistics. In our corpus, /o/ is more frequent in penultimate position in 2-syllable words than in 3-syllable words (36% vs 23%), the opposite of what we find for /e/ (37% vs 50%). **Constraint-based account.** Following Walker (2010, p. 942), among others, we assume that metaphony can be captured through a conflict between two constraints, namely, **LIC[+hi]** ([+high] in a post-tonic syllable must be associated with a stressed syllable), and **IDENT[hi]**, which penalizes candidates in which a vowel has changed its [high] specification relative to the input. In a MaxEnt grammar, constraints are weighted, and candidates are assigned probabilities. To account for the interaction between lexical statistics and target vowel, which creates the asymmetry in question, we assume that constraint weights are regulated by a scaling factor (e.g., Boersma & Hayes 2001, Coetzee & Kawahara 2013), such that the weight of faithfulness constraints (IDENT[hi]) is scaled down for more frequent words. This reflects the observation that more frequent forms are produced less faithfully across languages (van Oostendorp 1997, Itô and Mester 2001).

In our analysis, the weight of IDENT[hi] is scaled up or down proportionally to the lexical frequency of the group to which an input belongs. For instance, /kont-i/ ‘money.PL’ belongs to 2-syllable words containing the target vowel /o/, which are more frequent than /pes-i/ ‘fish.PL’, which belongs to 2-syllable words containing the target vowel /e/. As a result, the violation of Ident[hi] is less penalized for [kunti] than it is for [pisi]. This, in turn, results in [kunti] being a more probable candidate for /kont-i/ than [pisi] for /pes-i/, thus capturing the observed asymmetry in the corpus.

References

- Belloni, Silvano. 2009. *Grammatica veneta* [A grammar of Veneto]. Padova: Esedra.
- Boersma, Paul, and Bruce Hayes. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45–86.
- Coetzee, A. W., & Kawahara, S. (2013). Frequency biases in phonological variation. *Natural Language & Linguistic Theory*, 31(1), 47-89.
- Frosi, Vitalina M. and Ciro Mioranza. 1983. *Dialetos italianos: Um perfil linguístico dos italo-brasileiros do Nordeste do Rio Grande do Sul*. [Italian dialects: A linguistic profile of the Italian-Brazilians in the Northeast of Rio Grande do Sul]. Caxias do Sul, Brazil: EDUCS.
- Goldwater, S & Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Workshop on Variation within Optimality Theory*. pp. 111–120. <<http://homepages.inf.ed.ac.uk/sgwater/papers/OTvar03.pdf>>
- Hayes, B. and Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39(3):379–440.
- Itô, J., and Mester, A. (2001). Covert generalizations in Optimality Theory: The role of stratal faithfulness constraints. *Studies in Phonetics, Phonology, and Morphology* 7: 273–299.
- Perrone, A. (2016). *Il processo della metaforesi nei dialetti italiani*. Master's thesis, Università degli Studi di Padova.
- R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Smith, R. (2007). An overview of the tesseract OCR engine. *Ninth international conference on document analysis and recognition (ICDAR 2007)*, 2, 629–633.
- van Oostendorp, M. (1997). Style levels in conflict resolution. In *Variation, change and phonological theory*, eds. Frans Hinskens, Roeland van Hout, and Leo Wetzels, 207–229. Amsterdam: John Benjamins.
- Walker, R. (2010). Nonmyopic harmony and the nature of derivations. *Linguistic Inquiry*, 41(1), 169-179.
- Zamboni, Alberto. 1974. *I dialetti del Veneto* [Dialects of Veneto]. Pisa: Pacini.

Adjectival passives with a progressive reading

Elisabeth Gibert-Sotelo (*Universitat Rovira i Virgili*) & Rafael Marín (*CNRS – Université de Lille*)

The problem. According to Rapp (1996), only participles of verbs denoting states or including a stative component can be part of adjectival passives. This is the case of participles of telic verbs (1), which introduce a resultative meaning, as well as participles of certain stative verbs, such as extent verbs (2a), object experiencer psych verbs (2b), or Davidsonian states (2c), all of them combining with *estar* (the stage-level copula) in the Spanish adjectival passive:

- (1) El coche (ya) *está reparado*. ‘The car is (already) repaired’.
- (2) a. La casa *está rodeada* de árboles. ‘The house is surrounded by trees’.
b. Juan *está preocupado*. ‘Juan is worried’.
c. El edificio *está vigilado*. ‘The building is guarded’.

As has been recently observed (Fábregas & Marín 2012, 2017; Bosque 2014), most Davidsonian states (3) show an unexpected behaviour: in *estar* + participle constructions (2c), they give rise to a progressive reading that can be paraphrased by the passive progressive periphrasis *estar* + *siendo* + participle [lit. ‘be + being + participle’]. Thus, *El edificio está vigilado* ‘The building is guarded’ is roughly equivalent to *El edificio está siendo vigilado* ‘The building is being guarded’. As pointed out by Bosque (2014), in the absence of *estar*, the same progressive meaning is observed: *un edificio vigilado* ‘a guarded building’ is not about a building which has been guarded, but about a building which is being guarded.

- (3) *vigilar* ‘guard’, *gobernar* ‘rule’, *controlar* ‘control’, *coordinar* ‘coordinate’, *dirigir* ‘manage’, *presidir* ‘preside over’, *proteger* ‘protect’, *supervisar* ‘supervise’, *asesorar* ‘advise’.

According to Fábregas & Marín (2012, 2017), *vigilar*-verbs (3) denote Davidsonian states (Maienborn 2005), i.e. situations halfway between states and events. According to García-Pardo (2018), *vigilar*-verbs are essentially stative. In both cases, then, it could be maintained that it is their stative component which enables these verbs to be part of adjectival passives; so that Rapp’s generalization is respected. However, we have identified another group of verbs (4) that, like *vigilar*-verbs, also have a progressive reading in adjectival passive constructions (*Está perseguido/ buscado por la policía* ‘He is pursued/ searched by the police’), yet their denotation is properly dynamic.

- (4) *perseguir* ‘follow, chase, pursue’, *buscar* ‘look for, search’, *acosar* ‘harass, pursue relentlessly’, *acompañar* ‘accompany’, *escortar* ‘escort’.

Diagnosing dynamicity. It is not difficult to show that *perseguir*-verbs, unlike *vigilar*-verbs, are fully dynamic, since they pass most tests on eventivity and dynamicity: compatibility with the progressive (5a); availability as infinitival complements of perception verbs (5b); modification by manner adverbs (5c); compatibility with *parar* ‘stop’ (5d); habitual reading in the present tense (5e); unavailability of the universal reading in the perfect tense (5f). Notice that *vigilar*-verbs, as Davidsonian states, do not pass dynamicity tests (5d-f), but only eventivity ones (5a-c).

- (5) a. *Están {buscando/protegiendo}* al jefe. ‘(They) are {looking for/protecting} the boss’.
b. *Vi* a Eva {*perseguir/vigilar*} al jefe. ‘(I) saw Eva {chasing/watching over} the boss’.
c. Eva {*busca/vigila*} al jefe *discretamente*. ‘Eva {looks for/watches over} the boss discretely’.
d. *Ha parado de* {*acosar/#proteger*} al jefe. ‘(S/he) has stopped {harassing/#protecting} the boss’.
e. Eva {*persigue/#controla*} al jefe *una vez por semana*. ‘Eva {pursues/#controls} the boss once a week’.
f. Eva ha {*#buscado/asesorado*} al jefe *desde 2017*. ‘Eva has {#looked for/advise} the boss since 2017’.

Perseguir-verbs, then, pose a serious problem to Rapp’s generalization, according to which activities are explicitly out of adjectival passives.

Analysis. Fábregas & Marín (2012) derive the non-dynamicity of *vigilar*-verbs from the presence of a central coincidence P (Pcc) (cf. Hale 1986) at the complement of Proc, which, by structural homomorphism (Ramchand 2008), prevents the event introduced by Proc to be understood as a dynamic change, being understood instead as the homogeneous maintenance of a stative relation. Accordingly, we analyse the adjectival passives obtained from these verbs

as detailed in (6), where the Asp head merged on top of Init converts the non-dynamic event into a state and provides the participial morphology (Embick 2004). As far as the state encoded by the participle is derived from an event, we do not obtain a ‘pure’ individual-level state, but a stage-level one, which explains the compatibility of these participles with *estar*:

(6) Analysis of *El edificio está vigilado* ‘The building is guarded’.

[_{AspP} Asp -do [_{InitP} Init [_{ProcP} Proc [_{PccP} [_{DP} *el edificio*] [_{Pcc} Pcc [_√VIGIL]]]]]]]

As for *perseguir*-verbs, we assume that their dynamicity lies in the presence of a PathP at the complement of Proc. Evidence on the presence of this Path component is provided by the ability of these verbs to co-occur with adjuncts specifying the route followed in the course of the event:

(7) a. La policía buscó al ladrón *por toda la ciudad*.

‘The police looked for the thief throughout the city’.

b. El director acosa a su secretaria *por toda la oficina*.

‘The principal harasses his secretary all over the office’.

In contrast with other activity verbs (8a), *perseguir*-verbs cannot be telicized by adding a PP expressing an endpoint, and so they always denote atelic events (8b):

(8) a. Ha empujado el carro *hasta la esquina* (en cinco minutos).

‘(S/he) has pushed the cart to the corner (in five minutes)’.

b. La policía ha perseguido al ladrón *hasta la frontera* (*en dos horas).

‘The police have pursued the thief up to the border (*in two hours)’.

Perseguir-verbs encode a dynamic event that lacks a natural culmination and is hence maintained, a reading that emerges if the Path at the complement of Proc is non-bounded and, in addition, cannot be delimited by adding material at its complement. We claim that this effect is obtained if a Pcc is merged at the complement of Path, in a way so that a phrase defining a limit cannot be added directly at the complement of the Path projection to provide it with a boundary (since this position is already occupied by the Pcc). In addition, this Pcc (which is crucially different from a Place projection; cf. Real Puigdollers 2013: §2.2.1) provides the stative layer needed in the adjectival passive, which explains why *perseguir*-verbs can enter these constructions despite their atelicity and dynamicity:

(9) Analysis of *El ladrón está perseguido (por la policía)* ‘The thief is pursued (by the police)’

[_{AspP} Asp -do [_{InitP} Init [_{ProcP} Proc [_{PathP} [_{DP} *el ladrón*] [_{Path} Path [_{PccP} [_{DP} *el ladrón*] [_{Pcc} Pcc [_√PERSEG-]]]]]]]]]

Conclusion. The progressive (or continuous) reading characteristic of the two types of passive constructions analysed here emerges from the Aktionsart of the underlying verb: *vigilar*-verbs encode Davidsonian states and, as such, they involve a relation of maintenance between an event (Proc) and a stative layer (Pcc), rather than a proper causative relation (cf. Neeleman & van de Koot 2012); *perseguir*-verbs are dynamic activities (Proc + Path), but the dynamic event they encode does not have any possible telos, and hence it is understood to be maintained in a central coincidence relation (Pcc) that contributes the stative component needed in the adjectival passive. Given that in both cases the event of the underlying verb lacks a limit, the state codified by the participle cannot be interpreted as starting once the event has culminated (the resultative reading), but as being simultaneous to the event (the progressive reading).

References. Bosque, I. 2014. On resultative past participles in Spanish. *Catalan Journal of Linguistics* 13: 41-77. Embick, D. 2004. On the Structure of Resultative Participles in English. *Linguistic Inquiry* 35(3): 355-92. Fábregas, A. & R. Marín. 2012. Differentiating eventivity from dynamicity: the Aktionsart of Davidsonian state verbs. *LSRL* 42. Southern Utah University. Fábregas, A. & R. Marín. 2017. On non-dynamic eventive verbs in Spanish. *Linguistics* 55(3): 451-488. García-Pardo, A. 2018. *The Morphosyntax of States: Deriving Aspect and Event Roles from Argument Structure*. PhD thesis. USC. Hale, K. 1986. Notes on world view and semantic categories: Some Warlpiri examples. In P. Muysken & H. van Riemsdijk (eds.), *Features and projections*, 233-254. Foris. Maienborn, C. 2005. On the limits of the Davidsonian approach: The case of copula sentences. *Theoretical Linguistics* 31(3): 275-316. Neeleman, A. & H. van de Koot. 2012. The Linguistic Expression of Causation. In M. Everaert, M. Marelj & T. Siloni (eds.), *The Theta System: Argument Structure at the Interface*, 20-51. OUP. Ramchand, G. 2008. *Verb Meaning and the Lexicon. A First Phase Syntax*. CUP. Rapp, I. 1996. «Zustand Passiv? Überlegungen zum sogenannten “Zustandspassiv”». *Zeitschrift für Sprachwissenschaft* 15(2): 231-265. Real Puigdollers, C. 2013. *Lexicalization by Phase: The Role of Prepositions in Argument Structure and its Cross-linguistic Variation*. PhD thesis. UAB.

The interaction between postaspiration and stress in Sevillian Spanish

Madeline Gilbert, *New York University*

Introduction: While many varieties of Spanish aspirate coda /s/ to [h] (/tʃispa/ → [tʃihpa]), an ongoing change in Sevillian Spanish (Southern Spain) variably turns *preaspiration* into *postaspiration* ([tʃihpa]→[tʃip^ha]) (Ruch & Peters 2016; Torreira 2006). Some argue that postaspirated stops are becoming phonemic in Sevillian Spanish (Gylfadottir 2015; O’Neill 2009), but other work treats them as clusters (e.g. Torreira 2012; Parrell 2012). I probe the underlying representation of postaspirated stops with stress. Spanish stress is mostly predictable within a three-syllable, right-aligned window, and is partially weight-sensitive (Fuchs 2018). Penultimate stress is default; antepenultimate and final stress are less frequent. Antepenultimate stress is essentially nonexistent when the penult is heavy (*CV.CV.CVC.CV) (Fuchs 2018).

Under a cluster analysis, postaspiration is represented with an underlying /s/ that makes the preceding syllable heavy. Under a postaspiration analysis, it is represented as an aspirated stop that does not contribute weight to the preceding syllable.

Cluster analysis:	/pas.ta/ → [pa.t ^h a]	HL
Postaspiration analysis:	/pa.t ^h a/ → [pa.t ^h a]	LL

Do listeners treat syllables preceding postaspiration as heavy or light? I conduct a nonce word experiment that relies on the restriction against antepenultimate stress in words with heavy penults. My results show that Sevillian listeners treat words with postaspiration the same as words with surface-heavy penults. For example, antepenultimate stress is strongly dispreferred in both [la'pikinto] and [la'pikit^ho]. Both are treated as LLHL. I analyze this as an opaque interaction between debuccalization/metathesis and stress assignment.

Experiment: 26 Sevillian listeners completed a forced-choice task comparing words differing only by the type of coda in the penultimate syllable. The experiment asks two questions about words with *antepenultimate stress*. (a) In comparison to words with light penults, do listeners disprefer words with postaspiration as much as they disprefer words with other surface-heavy penults? (NoCoda comparisons) (b) In comparison to words with postaspiration, do they disprefer words with light penults (PostAspiration comparisons)? Table 1 shows the conditions. In NoCoda comparisons, listeners chose which word sounded like a better word of Spanish in pairs (a)-(b), (a)-(c), (a)-(d), (a)-(e). PostAspiration comparisons are (e)-(a), (e)-(b), (e)-(c), and (e)-(d).

Table 1 Nonce words

	a. NoCoda	b. Coda	c. CodaS	d. CodaH	e. PostAsp
/p/, /a/	gi'nakapo	gi'nakampo	gi'nakaspo	gi'nakahpo	gi'nakap ^h o
/t/, /i/	la'pikito	la'pikinto	la'pikisto	la'pikihto	la'pikit ^h o

45 test nonce words were recorded by a male speaker of Sevillian Spanish. There were 5 test conditions with 9 words each, differing only in the type of penultimate coda across conditions (Table 1). Nonce words had antepenultimate stress and consisted of light syllables, other than the penults being tested. The nonce words were balanced for onset of final syllable (/ptk/) and medial vowels (/aiu/). There were also fillers and controls, and neighborhood density was controlled for. Nonce words were presented auditorily in pairs, and the order of audio presentation was counterbalanced by participant. Each participant heard 108 trials consisting of both NoCoda and PostAspiration pairs. Results were analyzed in logistic mixed-effects regressions.

Results: In the NoCoda comparisons, listeners preferred the NoCoda form more than any of the forms with a coda (Figure 1, left). The rate of preference for NoCoda is not significantly different

in NoCoda-Coda, NoCoda-CodaS, NoCoda-CodaH, and NoCoda-PostAspiration pairs. This suggests that PostAspiration is representationally similar to forms with an acoustically heavy penult. The PostAspiration comparisons confirm this result: listeners disprefer PostAspiration in Postaspiration-NoCoda forms, but have no strong preference between PostAspiration and other forms with heavy penults (Figure 1, right): Sevillian listeners treat postaspiration as contributing weight to the preceding syllable, just like a surface-present coda [s, h, n].

Analysis: A serial analysis treating postaspiration as a cluster captures the opaque interaction between stress and postaspiration:

UR:	/ka.pi.ta.lis.ta/	LLLHL	'capitalist'
Stress:	[ka.pi.ta.'lis.ta]	LLLHL	(*[LLHL])
Debuccalization:	[ka.pi.ta.'lih.ta]	LLLHL	
Postaspiration:	[ka.pi.ta.'li.t ^h a]	LLLLL	

Stress must be evaluated on the UR, where /s/ is still a coda of the penult, rather than on the surface where the penult is light. The penult is heavy in the UR, so penultimate stress is the only option. Only then does [s] debuccalize to [h] and postaspire. Surface postaspiration cannot be syllabified as [Vt.hV], because [ptk + h] does not occur across syllable boundaries. If stress were evaluated after debuccalization and postaspiration, antepenultimate stress would be more acceptable in words with surface-light penults than in words with surface-heavy penults, regardless of UR ([gi'nakap^ho] > [*gi'nakaspo], [L^hLLL] > *[LHL]). It would also be equally good in words with surface postaspiration and those lacking codas ([gi'nakap^ho] ≈ [gi'nakapo], [L^hLLL]). If postaspirated stops were *underlying*, the same preferences would hold. This is not the case: antepenultimate stress is equally bad in words with penults closed by postaspiration and those closed by [s, h, n]. They share a representation with a heavy penult.

An alternative analysis in which postaspirated stops are single phonemes does not work: it is difficult to ban antepenultimate stress in words whose final onset is postaspirated. One option is to use onset-sensitive stress. Voiceless onset stops can attract stress on to the syllables they are in, e.g. /ginaka'p^ho/ (Gordon 2005; Topintzi 2011; Ryan 2016). Then stress would shift leftwards to result in [gina'kap^ho]. This analysis makes little sense for Sevillian Spanish because there is no surface evidence for underlying onset-sensitive weight or surface stress shift.

Opaque interactions: Postaspiration in Sevillian Spanish is an interesting addition to the body of work on opaque interactions between metrical and segmental structure, since metathesis changes syllable structure in a language whose stress is partially weight-sensitive. This is reminiscent of stress-epenthesis interactions. Epenthetic material is invisible for stress in some languages, but counts for stress in others. Similarly to metathesis, these opaque interactions are easier to analyze in serial frameworks (Elfner 2009). The interaction between stress and metathesis in Sevillian Spanish raises further questions about how stress and metathesis interact typologically.

Selected references: Gylfadottir, Duna. 2015. An investigation of Andalusian stop cluster Post-aspiration in naturalistic speech. Presentation presented at the 39th Penn Linguistics Conference, University of Pennsylvania. • Parrell, Benjamin. 2012. The role of gestural phasing in Western Andalusian Spanish aspiration. *Journal of Phonetics* 40. 37–45. • Ruch, Hanna & Sandra Peters. 2016. On the origin of post-aspirated stops: Production and perception of /s/ + voiceless stop sequences in Andalusian Spanish. *Laboratory Phonology* 7(1). 1–36. • Torreira, Francisco. 2012. Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association* 42(1). 49–63.

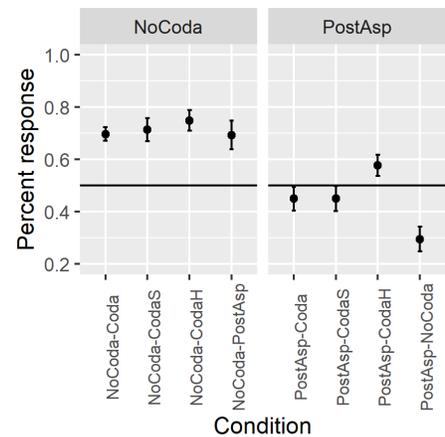


Figure 1 Proportion response of base form (vs. comparison form)

How frequent are these verbs? The role of lexical frequency in children's acquisition of verb morphology at different ages of acquisition

Michele Goldin, *University at Albany, SUNY*

Esther Hur, *Rutgers University*

Julio López-Otero, *Indiana University Bloomington*

In monolingual (L1) acquisition, children produce target-like subject-verb agreement at a remarkably young age in both Spanish (Grinstead, 1998) and English (Guasti, 2002). However, in heritage simultaneous bilinguals (2L1) and child second language acquirers (L2), agreement morphology has been shown to be prone to residual optionality (Goldin, 2020; Herschensohn & Stevenson, 2005) due to age of acquisition (AoA) effects. Lexical frequency is another factor that has been shown to play a key role in modulating L1 (Ambridge et al., 2015; Gathercole et al., 1999, 2002) and heritage acquisition (Giancaspro, 2017; Hur, López-Otero & Sánchez), but little is known about its effect in child L2. Ellis and Collins (2009) categorize frequency effects into token frequency (“how often a particular form appears in the input”, p. 330, hereafter lexical frequency) and type frequency (“the number of distinct lexical items that can be replaced in a given slot in a construction”, p. 330). Monolingual children first acquire more frequent words

(Fenson et al., 1994), as well as high frequency morphological forms (Räsänen et al., 2014), though no type frequency effects have been found (De Villiers, 1985).

Comparisons of L1, 2L1, child L2 and heritage acquisitions demonstrate that these processes are decidedly different in developmental path and ultimate attainment. Thus, this study explores the extent to which verb lexical frequency plays a role in the acquisition of verb morphology for bilingual children with differing AoA. We compare simultaneous heritage children (2L1) with those acquiring Spanish as an L2 from the age of 5 in a dual language school, a naturalistic immersion environment that is also accompanied by instructed learning.

In this study, 42 2L1 heritage children aged 3-7 and 46 L2 Spanish learners aged 5-7 participated in a Spanish fill-in-the-blanks production task in which they saw two images, one of a singular action and one of a plural action. They heard the accompanying sentence for the first image and were asked to complete the sentence with a verb for the second image. The stimuli were designed to elicit third person singular (n=4) and plural (n=4) morphology using common verbs in children's input in an academic setting.

The results of a GLMM indicated no effect of lexical frequency on target-like production of plural verbs. Thus, we focused our analysis on singular *correr* and *comer*, the two verbs featuring the same thematic vowel (Aguirre & Dressler, 2006). Our GLMM with response (target-like or not) as a dependent variable and verb (*correr* or *comer*) and group (2L1 or L2) as independent variables showed that responses to *correr*, the less frequent verb, were less target-like ($p < 0.01$) (see Table 1, Fig. 1 and Fig. 2) and that both groups performed differently ($p < 0.01$). These results are consistent with Putnam and Sanchez's (2013) Activation Approach to heritage language acquisition and maintenance as well as with several studies on lexical

frequency effects in L2 acquisition (e.g., Ellis & Collins, 2009). The results indicate that lexical frequency, as a proxy for language activation, modulates heritage and child L2 acquisition. We discuss these findings as they relate to the nature of the language learning mechanism in children.

Table 1. Proportion of target-like subject-verb agreement in 2L1 and L2 speakers of Spanish

	2L1 speakers				L2 speakers	
	PK 3	PK 4	K	1st	K	1st
Correr	M=0.36, SD=0.51	M=0.6, SD=0.52	M=0.55, SD=0.52	M=0.80, SD=0.42	M=0.42, SD=0.51	M=0.37, SD=0.49
Comer	M=0.55, SD=0.52	M=0.7, SD=0.48	M=0.82, SD=0.41	M=0.9, SD=0.32	M=0.37, SD=0.50	M=0.56, SD=0.51
Pintar	M=0.46, SD=0.52	M=1, SD=0	M=0.73, SD=0.47	M=0.9, SD=0.32	M=0.26, SD=0.45	M=0.67, SD=0.48
Dormir	M=0.55, SD=0.52	M=0.7, SD=0.48	M=0.64, SD=0.51	M=0.8, SD=0.42	M=0.16, SD=0.38	M=0.41, SD=0.50

Fig. 1 Accuracy of singular verb production as a function of verb frequency

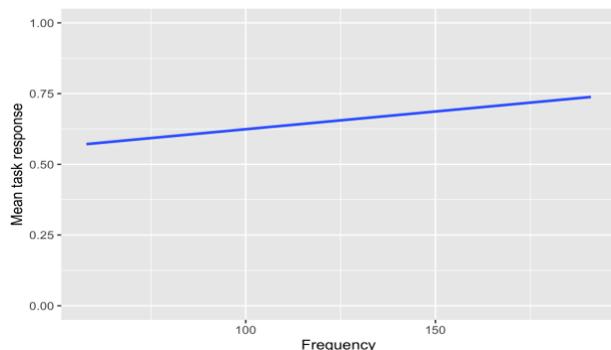
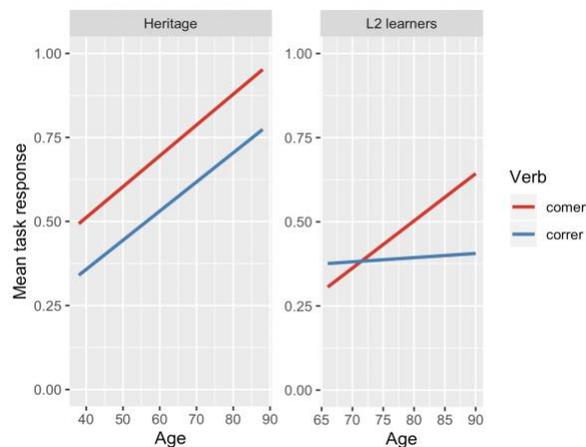


Fig. 2 Production accuracy of *correr* y *comer* as a function of age and group



References

- Aguirre, C., & Dressler, W. (2006). On Spanish verb inflection. *Folia Linguistica: Acta Societatis Linguisticae Europaeae*, 40(1-2), 75-96. // Ambridge, B., Kidd, E., Rowland, C. & Theakston, A. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42, 239-276. // De Villiers, J. G. (1985). Learning how to use verbs: Lexical coding and the influence of the input. *Journal of Child Language*, 12(3), 587-595. // Ellis, N. C. & Collins, L. (2009). Input and Second Language Acquisition: The roles of frequency, form and function. Introduction to the special issue. *Modern Language Journal*, 93, 329-335. // Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... & Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, i-185. // Gathercole, V., Sebastián, E., & Soto, P. (1999). The early acquisition of Spanish verbal morphology: across-the-board or piecemeal knowledge? *The International Journal of Bilingualism*, 3, 133-182. // Gathercole, V., Sebastián, E., & Soto, P. (2002). Negative commands in Spanish-speaking children: No need for recourse to relativized minimality. *Journal of Child Language*, 29, 393-401. // Giancaspro, D. (2017). Heritage speakers' production and comprehension of lexically-and contextually-selected subjunctive mood morphology. Doctoral dissertation, Rutgers University, New Brunswick, NJ. // Goldin, M. (2020). Syntax before morphology? The role of age of acquisition and context of acquisition in the development of subject-verb agreement in bilingual children. Doctoral dissertation. Rutgers University, New Brunswick, NJ. // Grinstead, J. (1998). Subjects, sentential negation and imperatives in child Spanish and Catalan. Doctoral dissertation. University of California, Los Angeles. // Guasti, M.T. (2002). *Language Acquisition: The Growth of Grammar*. Cambridge, MA: MIT Press. // Herschensohn, J. & Stevenson, J. (2005). Children's acquisition of L2 Spanish morphosyntax in an immersion setting. *International Review of Applied Linguistics in Language Teaching*, 43, 3; 193-217. // Hur, E., Lopez Otero, J. C., & Sanchez, L. (2020). Gender Agreement and

Assignment in Spanish Heritage Speakers: Does Frequency Matter?. *Languages*, 5(4), 48. // Putnam, M. T., & Sánchez, L. (2013). What's so incomplete about incomplete acquisition?: A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3(4), 478-508. // Räsänen, S. H., Ambridge, B., & Pine, J. M. (2014). Infinitives or bare stems? Are English-speaking children defaulting to the highest-frequency form. *Journal of child language*, 41(4), 756-779.

Two types of causatives and DOM

Julie Goncharov (*University of Goettingen*) & Monica Alexandrina Irimia (*University of Modena and Reggio Emilia*)

Setting the stage. As is well known, across Romance some objects (generally the human ones, in conjunction with other features such as specificity, etc.) are differentially marked via a (locative) preposition (as in 2b). In both descriptive as well as formal accounts (Comrie 1989, Torrego 1998, a.o.), important insights into the nature of differential object marking (DOM) have been shown to come from its interactions with *inanimate subjects*. In this work we use Romanian to examine a more nuanced interaction between inanimate subjects and DOM, which goes beyond current analyses, as discussed in more detail in Section 3. Our focus are types of experiencer object (EO) causatives, as in (2b), which, surprisingly, are only felicitous with differential marking, and not with unmarked objects (2a).

1. Marked objects and two types of causatives in the Romanian. Romanian shows that physical causatives allow DOM but do not need it obligatorily, as seen in (1). Complications arise with EO predicates, as in (2a) vs. (2b). The puzzle with (2a) is that the inanimate subject of an experiencer predicate is *not* felicitous with an unmarked indefinite object. For some speakers, in fact, these contexts do not even appear to be grammatical. Once differential marking is added, felicity is restored, as in (2a). Note that Romanian DOM on nominals is also accompanied by accusative clitic doubling.

(1) Furtuna (1-)a ucis/lăsat invalid (pe) un bătrân.
 storm.the CL.3M.SG.ACC-has killed/left invalid.M.SG LOC=DOM a old man
 ‘The storm has killed/left invalid an old man.’

(2) a. ?? Reclama amuză un trecător.
 ad.the amuses a passer by
 Intended: ‘The ad amuses a passer by.’

b. ✓ Reclama îl amuză pe un trecător.
 ad.the CL.3SG.ACC.M amuses LOC=SOM a passer by
 Intended: ‘The ad amuses a passer by.’

Romanian

In a pilot study we have conducted, examples like (2a) indeed have received much lower acceptability scores - 4.54 out of 7. They contrast not only with contexts such as (2b), where DOM is seen with an inanimate subject of an EO predicate and which get 6.53 out of 7, but also with contexts with inanimate subjects and physical causatives. The score for examples similar to (1) without DOM is 6 out of 7 and with DOM is 6.5 out of 7. Grammatical fillers get a score of 6.6 out of 7.

As also discussed in Section 3, EO contexts with obligatory DOM as in (2b) are puzzling under most current accounts, which would predict that inanimate subjects should at most trigger optional DOM, and not obligatory DOM. Thus, we are interested in formulating an analysis for these contexts and also evaluating the insight they provide us with when it comes to the nature of DOM.

2. Proposal. Our analysis connects two independently motivated ideas: (i) DOM has a sentience feature (Irimia 2020, a.o.) signalling that the marked object is a sentient entity and (ii) EO verbs, unlike physical causatives, presuppose a perception event in which the object of the asserted event is a perceiver, and thus must be sentient (Lakoff 1995, Oyama 2003, a.o.).

(i) DOM has a sentience feature. There are three strong arguments for the claim that DOM has a sentience feature (Belletti 2019, Irimia 2020a,b, a.o.). First, one telling context comes from ‘pain predicates’ with object experiencers (*durea* ‘ache’, *mânca* ‘scratch, itch’, etc.), as in (3), where DOM is *obligatory* on the animate (in)definite (we have seen above in (1) that animate (in)definites can otherwise be used without DOM in Romanian). Second, oblique DOM triggers PCC effects which cannot be easily derived as a result of competition in terms of a Case feature; a specification more similar to [PERSON] or Sentience gives more adequate results (Irimia 2020a, b). And third, the addition of DOM to non-human animate DPs introduces an antropomorphic feature re-ranking them to the status of *sentient* human entities (e.g., treating pets as family members).

(3) (Îl) doare capul *(pe) (un) copil.
 CL.3SG.M.ACC aches head.the LOC=DOM a.M.SG child

‘A/the child’s head hurts.’

Romanian

To formalize this, we propose that the [Sentience] feature on DOM denotes a partial function of type $\langle e, e \rangle$ defined only for Sentient individuals, as in (4) (building on Heim and Kratzer’s 1998 analysis of gender features). Thus, a simple sentence as in (5a) will be interpreted as in (5b).

(4) $[[\text{Sentience}]]^{g,w} = \lambda x: \text{Sentient}_w(x). x$

(5) a. Maria (îl) vede **pe** un copil. ‘Maria (cl) sees DOM a boy’

b. $[[[\text{un copil}] [1 \text{ Maria vede } [\text{Sentience } t_1]]]]^{g,w} = 1$ iff $\exists x[x \text{ is a boy in } w \wedge \text{Maria sees } x \text{ in } w]$, defined only if $\text{Sentient}_w(x)$

(ii) **EO verbs, but not physical causatives, involve a perceiver.** It has been previously observed that causation in EO verbs, such as *annoy* and *amuse*, is different from physical causation in predicates such as *hit* and *break*, in that the former but not the latter have an experiencer centred construal (e.g., McCawley 1973, Lakoff 1995). In particular, EO verbs presuppose that there is a perception event in which the EO is also the perceiver of the causation event (e.g., Oyama 2003). Arguments for this presupposition come from examples like (6) and (7). The oddness of the EO construction in (6b) is explained by the fact that the speaker is asserted as not participating in the perception event. The same additional statement in the physical causation event in (6a) does not result in infelicity. Similarly, (7b) is odd because a blind man cannot perceive the lamp’s colours and thus, be amused by them, but he can perceive the shape of the lamp, see (7a). These contrasts hold for Romanian as well as other languages.

(6) a. The car that I couldn’t see knocked me over.

b. # The view that I couldn’t see knocked me over.

(7) a. The lamp amused the blind man because of its shape.

b. # The lamp amused the blind man because of its colour.

We can capture this proposal by saying that EO verbs have the denotation as in (8a), which unlike that of physical causatives, see (8b), has presuppositional content.

(8) a. $[[\text{amuse/amuza}]]^{g,w} = \lambda x \lambda y: x \text{ perceives } y \text{ in } w. y \text{ causes } x \text{ to become amused in } w$

b. $[[\text{injure/răni}]]^{g,w} = \lambda x \lambda y. y \text{ causes } x \text{ to become injured}$

Therefore, when the EO verb appears with a marked object, the presuppositions of both are satisfied. We explain the contrast observed in (2) by invoking the Maximize Presupposition (MP) principle (Heim 1991) formulated to explain the deviance of sentences like #*A sun is rising*. According to MP, a sentence p is infelicitous in the context c if there is an alternative sentence q that has the same assertive component as p but is presuppositionally stronger in cases where the presuppositions of both p and q are entailed in c . Therefore, with EO verbs which presuppose that the EO is a perceiver and thus is a sentient individual, DOM-less objects give rise to infelicity. Physical causation does not presuppose sentience of the object, therefore DOM is optional.

3. Discussion. For Comrie (1989) and subsequent functionalist works, DOM is seen as a disambiguation strategy in those configurations in which objects and subjects are (semantically) too similar. More precisely, it signals those objects that have characteristics more similar to subjects (animacy, etc.) so that the two classes don’t get confused. Similarly, the possibility of DOM with inanimate subjects under causatives is explained by Torrego (1998) under the assumption that marked objects can only be licensed after scrambling to a [Spec, v] position (more accurately, a second specifier of v). The problem with accounts in this direction is that they cannot predict why DOM is optional with physical causatives as in (1), but obligatory with EO causatives as in (2). Both these configurations contain a causer and an object with properties which are more similar to subjects, and thus we would expect them to behave in the same way, contrary to fact. Our account shows that there is variation in the class of causative predicates and this has consequences for DOM beyond animacy and the presence of a causer. We will also discuss exceptions to the DOM obligatoriness such as EO verbs with animate subjects and EO verbs with inanimate subjects and unmarked generic definite plurals. In summary, our proposal confirms instead that DOM is a semantically complex marker, interacting not only with specificity or definiteness, but also with less studied *sentience*.

Selected references. **Comrie**, B. 1989. *Language universals and linguistic typology*. Chicago. **Irimia**, MA. 2020a. Differential objects and other structural objects. In *Case, agreement and their interactions*, 77-126, De Gruyter. **Oyama**, K. 2003. Experiencer-centered event construal of psychological causation. *Bulletin of Tokyo Institute* 9, 73-78. **Torrego**, E. 1998. *The dependencies of objects*. MIT Press.

In a pioneering discussion, Ormazabal & Romero (O & R 2007) have addressed PCC-effects that are closely linked to animacy, with data from clitic clusters in *leísta* varieties (6). It immediately became clear that such effects go well beyond clitic clusters and affect even differential object marking (DOM) on full nominals, as in (1). In this investigation, we examine DOM co-occurrence restrictions on full nominals, focusing on their repair strategies. More specifically, we address a (generally ignored) split between languages like Spanish, on the one hand, and languages like Galician and Romanian, on the other hand. As opposed to Spanish, the latter two languages use accusative clitic doubling on DOM in order to repair ungrammatical configurations containing a marked full nominal D(irect) O(bject) and a clitic-doubled I(ndirect) O(bject), as in (2) and (3). Our claim is that movement of the clitic-doubled DOM-ed DP to a ν P-peripheral position permits its licensing, leaving the initial licenser in a position below ν P available for the licensing of the clitic-doubled dative.

1. DOM and co-occurrence restrictions. The Spanish example in (1) shows the inability of the DO to carry differential marking in the presence of an IO that has been clitic doubled. O & R (2013a, b, 2019, et subseq.) have explicitly connected these effects to principles behind the better known P(erson)C(ase)C(onstraint). More specifically, the animacy feature on the marked object introduces a Case feature competing for licensing with the Case feature of the clitic-doubled IO.

CL_{DAT.3SG} send.PST.3PL DOM all the sick.people DAT the doctor

Intended: ‘They sent all of the sick people to the doctor.’

Another DOM prominent language, Romanian, shows the same restriction; as discussed by Cornilescu (2020: ex. 4), a structure like (2a), where we try binding from a marked DO into a clitic-doubled IO, is ungrammatical. Romanian, however, shows a repair strategy: clitic doubling of the DOM-ed nominal restores grammaticality, as in (2b) - Cornilescu’s (2020) example 6.

(2) a. *Comisia le-a repartizat **pe** mai mulți medici
board CL_{DAT.3PL}-hasassigned DOM more many medical
rezidenți unor foști profesori de-ai lor
residents some former professors of theirs

Intended: ‘The committee assigned several residents to some former professors of theirs.’

b. Comisia *i* *i*-a repartizat **pe** fiecare resident
board CL_{DAT.3SG}CL_{ACC.M.SG}-has assigned DOM each resident
unei foste profesoare a lui.
some former professor of his

Intended: ‘The committee assigned every medical resident to some former professor of his.’

The Romanian picture is mirrored by Galician. Although Galician does not have accusative clitic doubling of full DPs in argument position, differentially marked strong pronouns may be doubled in the same way as in Romanian as a repair strategy, as illustrated by the contrast in (3).

(3) a. *Mandáronlle a el ó medico [ó=a_{DAT}+O_{DET.M.SG}]
send.PST.3PL-CL_{DAT.3SG} DOM he to-the doctor

b. Mandáronllo a el ó médico [llo=l_{DAT}+O_{ACC}]
send.PST.3PL-CL_{DAT.3SG}-CL_{ACC.M.SG} DOM he to-the doctor

Intended: ‘They sent him to the doctor.’ *Galician*

These languages contrast with (Peninsular) Spanish which, like Galician, does not generally exhibit clitic doubling of full DPs, but does double strong pronouns, as in (4a). However, doubling both the DO and the IO yields ungrammaticality in (4b), unlike in Romanian (2b) and Galician (3b).

(4) a. Lo envié a él al médico
CL_{ACC.3SG} send.PST.1SG DOM he to-the doctor

b. *Se lo envié a él al medico
SE CL_{ACC.3SG} send.PST.1SG DOM he to-the doctor

Intended: ‘I sent him to the doctor.’ *Standard Peninsular Spanish*

Structure preservation and contact effects: Subjects in Brazilian Veneto

Natália Brambatti Guzzo & Avery Franken, *McGill University*

Overview: In contact situations, certain constructions in one of the languages may be affected by the way in which they are structured in the other language. In this paper, we examine the manifestation of subjects in Brazilian Veneto (often referred to as Talian), which is in contact with Brazilian Portuguese (BP). While Central Veneto (CV), a closely related dialect spoken in northeastern Italy, is a null subject language that exhibits subject clitics as agreement markers (Cardinaletti, 2004), BP is a partial null subject language that has no subject clitics (Barbosa et al., 2005). It has been observed that, in spoken Talian, the distribution of subject clitics mirrors the patterns for subject pronouns in BP, which seems to suggest that the language has undergone contact-induced change (Frasson, 2020). Based on an analysis of written narratives, we argue that, while some contact effects with BP can be observed in the data, Talian speakers still have partial access to the CV structure for subjects, since their use of null subjects and subject clitics is mostly compatible to what is observed in CV.

Subjects in Veneto and Portuguese: In CV, subject doubling is observed – subject clitics may follow both strong subject pronouns (2sg and 3sg/pl) and phrasal subjects, as exemplified in (1) (the examples are adapted from Belloni, 2009:78). There are no subject clitics for first person (singular/plural) and second person plural.

- (1) Ela/La tosa la sa tuto.
she.STRONG/the girl she.CL knows everything

It has been argued that subject clitics in CV are agreement markers realized in T – in negative sentences and sentences with preverbal adverbs, the clitic immediately precedes the verb, while strong pronouns and phrasal subjects precede the negator/adverb (2) (Cardinaletti, 2004; see also Alexiadou and Anagnostopoulou, 1998).

- (2) Ela no la sa tuto.
she.STRONG NEG she.CL knows everything

In CV, discourse subjects can be retrieved from verbal morphology or the use of clitics. For this reason, the language has been classified as a null subject language (see Cardinaletti, 2004). BP, on the other hand, has been argued to be a partial null subject language, given its shift to a more simplified verbal morphology and the overuse of overt pronouns (Barbosa et al., 2005; Duarte and Varejão, 2013). Still, in structures with no topic shift, such as coordinate constructions, null subjects are preferred. Unlike Veneto, however, BP does not allow subject doubling, except in cases where a phrasal subject is topicalized (in this case, the phrasal subject is followed by a subject pronoun). In its spoken form, Talian seems to be similar to BP, suggesting crosslinguistic influence – subject clitics are produced before the negator, and they are omitted in coordination (Frasson, 2020).

Methodology: The question that arises is whether Talian has completely incorporated the features of BP subjects, or whether its speakers' grammars have access to the structure of CV. To probe this issue, we analyzed 996 clauses extracted from short narratives written in Talian by seven different authors and published in newspapers from southern Brazil between 2010 and 2019. These narratives are part of the Talian Corpus (Garcia and Guzzo, 2020). The choice to examine written materials stems from the idea that the written version of a language, being more conservative than its spoken form, may contribute to a more comprehensive view of speakers' competence. Since

Talian has no official orthography or standardized grammar, writers' use of subjects may reflect their intuitions about linguistic structures.

All clauses had an inflected verb. The clauses were coded for *type of subject* (null vs. overt), *type of overt subject* (phrasal, strong pronoun, relative pronoun, quantified expression, only clitic), *person, position of the clause in the sentence* (first vs. non-first), whether they exhibited *topic shift*, and whether or not they had a clitic. In *type of overt subject*, *phrasal* comprises both phrasal subjects such as *la tosa* 'the girl' and proper nouns. *Only clitic* (i.e., the clause has no subject other than a clitic form) was included as a level in this variable under the assumption that, if Talian subjects are structured similarly to BP subjects, Talian clitics should behave like pronouns in BP. *Topic shift* was coded as *no* in cases where the subject of the target clause and the subject of the preceding clause had the same antecedent. Additionally, negative clauses were coded for position of the clitic relative to the negator. The data were modelled with logistic regressions in R (R Core Team, 2020).

Results: We report and discuss a subset of our results here. Regarding the use of null vs. overt subjects by grammatical person, 2pl subjects were excluded from the analysis due to a negligible number of tokens. 1sg and 1pl both favored null subjects (82% of null subjects each; 1sg: $\hat{\beta} = -4.02$, $p < 0.000$; 1pl: $\hat{\beta} = -4.01$, $p < 0.000$) relative to 3sg (the baseline). 3sg, 3pl and 2sg all displayed a high number of overt subjects (87%, 99% and 97%, respectively), although many tokens of 3sg, 3pl and 2sg exhibited only a clitic as the subject (45%, 39%, 95%).

We examined the realization of subject clitics in contexts of [\pm topic shift] as well as relative to the position of the clause in the sentence, as a way to shed light on their role in the Talian grammar (i.e., as agreement markers like in CV, or as subject pronouns like in BP). Regarding clauses that had only a clitic in subject position, 37% of them were [+topic shift] and 46% of them were sentence-initial clauses. In this case, 2sg and 3pl clitics appear significantly more in [+topic shift] relative to 3sg ($\hat{\beta} = 1.14$, $p = 0.01$ and $\hat{\beta} = 1.00$, $p = 0.01$, respectively). There are no significant differences between grammatical persons for *position of the clause in the sentence*.

To further explore contexts where clitics are used, we probed the distribution of 3sg and 3pl clitics according to *type of overt subject*. There are significantly more clitics following a strong pronoun than a phrasal subject ($\hat{\beta} = 1.04$, $p = 0.04$). There are no significant differences between phrasal subjects and the other types of overt subject included (quantified expression, relative pronoun). With respect to negative clauses, out of the 37 negative clauses with subject clitics in the data, only one had the clitic before the negator.

Discussion: The results for 1sg and 1pl are consistent with what is expected of a null subject language – since these grammatical persons are signaled by verbal morphology, their corresponding pronouns are not required. This contrasts with the patterns observed in (spoken) BP, where first person displays *overt* subjects over 70% of the time (Duarte, 2000). In addition, the fact that clitics are favored by a preceding strong pronoun, as well as the observation that clitics may follow other types of subjects (phrasal, quantified expression, relative pronoun) seems to indicate the preservation of subject doubling in Talian. Another result that further indicates a match between the Talian grammar and the CV grammar is the prevalence of subject clitics following negators. On the other hand, the results for topic shift and position of the clause in the sentence support the idea that subject clitics are accepted as possible sentence subjects in Talian, resembling what is observed in the use of subject pronouns in BP. The observation that CV-like structures are found in written narratives but dispreferred in oral communication – in favor of structures that match those of BP (as per Frasson, 2020) – seems to further reveal the influence of BP on the current status of Talian.

The Role of C and T in the *That*-trace Effect: Evidence from Code-switching

Bradley Hoot (*DePaul University*) & Shane Ebert (*University of Illinois Chicago*)

Background: Several Romance languages (e.g. Spanish, Italian, and Catalan) contrast with other languages (e.g. English, Russian, and Wolof) in their treatment of subject extraction from a subordinate clause. Languages like English ban extraction from subject position, but not from object position, in subordinate clauses in the presence of an overt complementizer like *that*. This is known as the *that*-trace effect; it remains essentially a mystery (see Pesetsky, 2017 for an overview).

(1) *That*-trace effect

- | | |
|---|---|
| a. ✓ Who do you think that Sue met ___? | c. * Who do you think that ___ met Sue? |
| b. ✓ Who do you think Sue met ___? | d. ✓ Who do you think ___ met Sue? |

Spanish, however, has obligatory complementizers and extracting subjects over them is permitted.

(2) Spanish extraction over obligatory *que*

- | |
|---|
| a. ✓ ¿Qué crees que compró Susana ___? 'What do you think that Susana bought?' |
| b. ✓ ¿Quién crees que ___ compró el libro? 'Who do you think that bought the book?' |

This difference has been tied to another contrasting property between these sets of languages, viz. the availability of post-verbal subjects in a wide variety of contexts. Essentially, if the *that*-trace effect stems from an independent restriction on subject extraction from the specifier of TP (the commonly assumed pre-verbal subject position), then languages that allow post-verbal subjects can obviate the *that*-trace effect by instead extracting the subject from a post-verbal position (Rizzi, 1982).

We thus observe two differences between English and Spanish: one tied to the complementizer and one tied to subject position, which is usually attributed to the EPP, generally understood as a property of TP. Under a view of syntax which attributes cross-linguistic variation to the features of lexical items (Chomsky, 1995), variation can be traced to different feature specifications on the relevant functional heads. In this case, the most plausible sites for feature differences are the C and T heads.

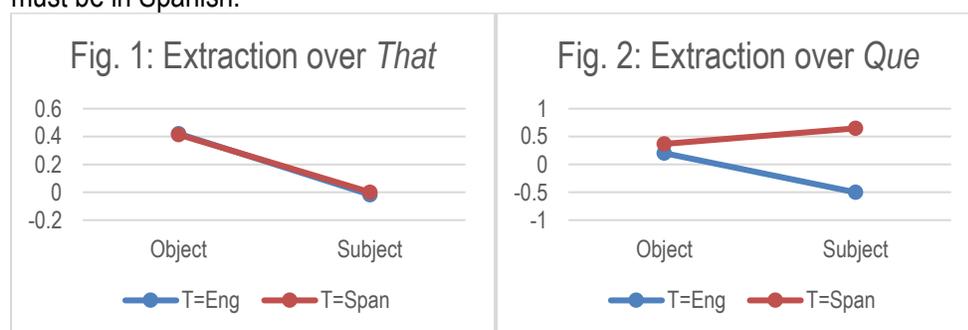
Because English and Spanish differ, it may be possible to gain additional insight into these puzzles by combining lexical items from both languages in a single derivation, as in code-switching (MacSwan, 1999). By manipulating which lexical items come from which language, it is possible to isolate specific syntactic features, providing novel evidence unavailable in monolingual data (González-Vilbazo & López, 2012). Indeed, previous results from Spanish-English code-switching suggest that two properties related to *that*-trace—null subjects and post-verbal subjects—are only licensed when *both* C and T are in Spanish (Ebert & Hoot, 2018; Sande, 2018). Assuming the language of the C and T heads indicates the features they possess, subjects in code-switched sentences behave like Spanish only when C and T both share the relevant features, suggesting that properties of subjects rely not on either head alone but on both.

Method: We expand this line of research by examining the contrast between subject and object extraction from embedded clauses in sentences with Spanish-English code-switching, like (3) and (4). Following previous findings, we predicted that subject extraction over a complementizer will be allowed only when C and T are both in Spanish.

- (3) [Quién/Qué] *asumieron los maestros que* [the students/___] had read [___/the book]?
who what assumed the teachers that
- (4) [Who/What] did the teachers assume that [*los alumnos/___*] *habían leído* [___/el libro]?
the students had read the book

Spanish-English bilinguals ($n = 36$) completed a written acceptability judgment experiment. We tested subject extraction over an overt complementizer via a $2 \times 2 \times 2$ design: Language of C (Spanish or English), Language of T (Spanish or English) and Wh-Type (subject or object). We z-score transformed the ratings, then fit a linear mixed model to the data for each combination of C and T. Results mentioned are significant ($p < .05$) tests of fixed effects.

Results: Fig. 1 shows no interaction between extraction type and Language of T when C is English *that*; instead, extracting subjects is always worse. This result shows that Spanish T alone is not enough to license subject extraction. Fig. 2 shows a significant interaction between extraction type and Language of T when C is Spanish *que*, such that subject extraction is acceptable only when both C and T are in Spanish. Overall, then, this experiment finds support for the hypothesis that both C and T must be in Spanish to obviate the *that*-trace effect. This provides evidence against syntactic accounts appealing to properties of C or T individually (e.g. González-Vilbazo & López, 2013; Rizzi & Shlonsky, 2007), instead supporting work closely tying together C and T (e.g. Broekhuis, 2016; Chomsky, 2008; Pesetsky & Torrego, 2001). These results also offer indirect support for the hypothesis that the availability of post-verbal subjects obviates *that*-trace (Rizzi, 1982), since previous work on post-verbal and null subjects also suggested that C and T must be in Spanish.



References

- Broekhuis, H. (2016). Feature inheritance versus extended projections. In G. Legendre, M. T. Putnam, H. de Swart, & E. Zaroukian (Eds.), *Optimality-Theoretic Syntax, Semantics, and Pragmatics* (pp. 136–157). Oxford University Press.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press.
- Chomsky, N. (2008). On phases. In R. Freidin, C. Otero, & M. L. Zubizarreta (Eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud* (pp. 134–166). MIT Press.
- Ebert, S., & Hoot, B. (2018). That-trace effects in Spanish-English code-switching. In L. López (Ed.), *Code-switching – Experimental Answers to Theoretical Questions: In honor of Kay González-Vilbazo* (pp. 101–145). John Benjamins. <https://doi.org/10.1075/ihtl.19.05ebe>
- González-Vilbazo, K., & López, L. (2012). Little v and parametric variation. *Natural Language & Linguistic Theory*, 30(1), 33–77. <https://doi.org/10.1007/s11049-011-9141-5>
- González-Vilbazo, K., & López, L. (2013, April). *Phase switching*. Code-switching in the bilingual child: Within and across the clause, Center for General Linguistics (ZAS), Berlin.
- MacSwan, J. (1999). *A Minimalist Approach to Intra-sentential Code-switching*. Garland.
- Pesetsky, D. (2017). Complementizer-trace effects. In M. Everaert & H. van Riemsdijk (Eds.), *The Wiley Blackwell Companion to Syntax* (2nd ed.). Wiley Blackwell.
- Pesetsky, D., & Torrego, E. (2001). T-to-C movement: Causes and consequences. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 355–426). MIT Press.
- Rizzi, L. (1982). *Issues in Italian syntax*. Foris.
- Rizzi, L., & Shlonsky, U. (2007). Strategies of subject extraction. In U. Sauerland & H.-M. Gärtner (Eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics* (pp. 115–160). Mouton de Gruyter.
- Sande, A. (2018). *C plus T as a necessary condition for pro-drop: Evidence from code-switching* [Doctoral dissertation]. University of Illinois at Chicago.

On the development of negative *en DET vida* in Spanish

Chad Howe & Javier Cabezas Zapata, *University of Georgia*

The literature on linguistic negation is characterized by a number of open questions, not the least of which are those regarding the grammaticalization of NEG-words (Laka, 1990). In this paper, we explain the diachronic development of the lexicalized PP structure in Spanish, *en DET vida* ‘in DET life’, exemplified in (1) and (2), extracted from the *Corpus del Español* (Davies, 2016, CdE). Contrary to Vallduví (1994), Espinal (2000) argues that n-words, such as *ningún*, *nadie*, and *en la vida*, are not lexically ambiguous, contrasting these expressions with minimizers (e.g., *ni* indefinite NP) and polarity items. Our analysis reconsiders this proposal in the light of diachronic corpus evidence and maintains that, although the contemporary use of *en DET vida* parallels that of other NEG-words, the development of the negative use of this structure requires contextual ambiguity. As part of our analysis, we propose a grammaticalization pathway that accounts for the current meaning of this structure.

In modern Spanish, the *en DET vida* patterns like other NEG-words, including the ability (i) to occur in isolation (as in 4 a) and (ii) to occur in pre-verbal position (shown in 4 b). Unlike many of its counterparts, however, *en DET vida* allows for a non-negative interpretation in post-verbal position, as illustrated in example (5 a). In pre-verbal position, we argue that an ambiguity arises which allows for either a positive or a negative interpretation, both illustrated in example (5 b). It is this ambiguity which serves as the source of the semantic change resulting in the negative meaning.

To determine the development of this structure, we extracted cases of *en DET vida*, where “DET” stands for either definite and possessive determiners, from the *Corpus Diacrónico del Español* (CORDE). Based on these data, there were no attested cases of negative *en DET vida* prior to the 15th century, with the majority of the tokens expressing meanings similar to that in example (6). By the 16th century, *en DET vida* had acquired an emphatic, negative function, precipitated, we argue, by its co-occurrence with other NEG-words, such as *nunca* (see example 7). Our analysis of the diachronic data thus reveals that, like French *pas*, *en DET vida* participates in a Jespersen cycle, eventually being compatible with a negative meaning in pre-verbal position (see Amaral, 2020; Hopper and Traugott, 2003; Jespersen, 1917), though it does not supplant the original negative adverb (i.e. *no*).

In addition to our proposed pathway, we observed other collocational patterns in the data following Ruppenhofer and Michaelis (2016) including the aforementioned co-occurrence with universal NEG adverbials (*nunca* / *jamás* ‘never’) and the increasingly fixed use of the definite determiner, which increases from 35% in the 15th century to 66% in the 19th. We maintain that the shift from lexical PP to universal negation arises via a contextual ambiguity facilitated by the co-occurrence with items like *nunca*. Our analysis provides critical insight regarding the development of lexical negative structures and their status vis-à-vis other items of this class.

Examples

- (1) *En la vida he visto gente que conduzca peor, a tanta velocidad y de forma tan sumamente agresiva.*
- (2) ‘**Never in my life** have I seen people who drive so poorly, so fast and in such an aggressive manner.’ (CdE)
- (3) *En mi vida he visto a nadie consumir cerveza después de ejercicio*
‘I have **never** seen anyone drink beer after exercising.’ (CdE)
- (4) (a) A: *¿Has ido a Brasil?* ‘Have you been to Brazil?’
B: *¡En la vida!* ‘**Never!**’
(b) *En la vida he visto tanta violencia.*
‘**Never in my life** have I seen such violence.’
- (5) (a) *He ido a Egipto tres veces en mi vida.*
‘I have been to Egypt three times **in my life**.’
(b) *En mi vida he ido a Egipto tres veces.*
‘**In my life**, I have been to Egypt three times.’ → **Positive**
‘I have **never** been to Egypt three times [i.e. in a year].’ → **Negative**
- (6) *Onde conuiene mucho al pueblo assi como en la vida son tenudos de onrrar a su rey que asi lo fagan a su finamiento.*
‘Now it is appropriate that the people **in life** must honor their king, so they do until the end’ (1481, Anónimo, *Siete partidas de Alfonso X*, CORDE)
- (7) *¡Quedaos para loco, / que nunca en mi vida tu amigo seré!*
‘Be as crazy as you want, / because **never in my life** I will be your friend!’
(1524, Hernán López de Yanguas, *Farsa del Mundo y Moral*, CORDE)

References

- Amaral, P. (2020). Bocado: Scalar semantics and polarity sensitivity. *Zeitschrift für romanische Philologie*, 136, 1114–1136.
- Davies, M. (2016). *Corpus del Español: Web/Dialects*. <http://www.corpusdelespanol.org/web-dial/>
- Espinal, M. T. (2000). On the semantic status of n-words in Catalan and Spanish. *Lingua*, 110(8), 557–580.
- Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization* (2nd ed.). Cambridge University Press.
- Jespersen, O. (1917). *Negation in english and other languages*. Høst.
- Laka, I. (1990). *Negation in syntax: On the nature of functional categories and projections* (Doctoral Dissertation). MIT. Cambridge, MA.
- Real Academia Española. (n.d.). *CORDE: Corpus Diacrónico del Español*. Retrieved December 22, 2020, from <http://corpus.rae.es/cordenet.html>
- Ruppenhofer, J., & Michaelis, L. A. (2016). Frames, polarity and causation. *Corpora*, 11, 259–290.
- Vallduví, E. (1994). Polarity items, n-words and minimizers in Catalan and Spanish. *Probus*, 6(2-3), 263–294.

Lenition and contrastiveness: word-final stops in Catalan

José Ignacio Hualde & Jennifer Zhang, *University of Illinois at Urbana-Champaign*

In this study we examine the effects of word boundaries on the lenition of intervocalic voiceless plosives in Catalan in order to test the role of phonological contrastiveness in phonetic processes. In Catalan, like in all other Romance languages, there is a contrast between voiced and voiceless plosives in syllable-initial position (e.g. *pala* /palə/ ‘spade’ vs *bala* /balə/ ‘bullet’; *capa* /kapə/ ‘cape’ vs *cava* /kabə/ ‘cave’). This contrast, however, is neutralized word-finally in favor of the voiceless series, giving rise to numerous morphophonological alternations (e.g. *sap* ‘s/he knows’, *sabem* ‘we know’). Before a pause or a vowel-initial word, only voiceless plosives are found [1,10]. Here we test the **hypothesis** that word-final intervocalic voiceless plosives (VC#V) will show greater lenition than word-internal and word-initial intervocalic tokens (VCV, V#CV), since in word-final position the voice contrast is neutralized. Lenition should be manifested acoustically as greater intensity, shorter duration and greater voicing. **Methodology.** We have chosen the Catalan News subcorpus of the bilingual Spanish/Catalan corpus *Glissando* [7], which includes data from 8 professional speakers (4f, 4m) reading the news. The style that we are analysing is thus the fluent read speech of professional news reporters. In its degree of reduction, this type of speech is expected to be intermediate between laboratory speech and casual conversation. From this corpus we have extracted 1909 tokens of intervocalic /ptk/ in different positions in the word (V#CV = 772, VCV = 990, VC#V = 147), plus 281 tokens in word-final preconsonantal position (VC#C), distributed as in Table 1. For each target consonant we have segmented a window in PRAAT containing the consonant and the following vowel and calculated the difference in intensity (*IntDiff*), where a greater difference between minimum and maximum intensity indicates a more constricted consonant [2,3,4,5,9]. This methodology does not require the accurate placement of segment boundaries, which in the case of highly reduced consonants can be difficult to do in an objective manner. On the other hand, it does not allow us to report duration values. We have calculated the difference in intensity within the window in two different ways: without filtering and applying a high-pass filter at 250 Hz to exclude the effects of voicing. In addition, we have also taken another energy measurement, *Delta_i*, using the methodology proposed in [6]. We have also automatically extracted the amount of voicing within the segmented window.

Results. A mixed-effects regression analysis with unfiltered *IntDiff* as dependent variable and context, phoneme and sex as fixed factors (using the function *lmer* in the R package *lme4*) shows that word-final prevocalic stops are more lenited (i.e. they show a lower value in the *IntDiff* measurement) than word-initial ($\beta = 2.34$, $t = 4.2$, $p < 0.0001$) and word-medial ($\beta = 1.1$, $t = 2.1$, $p = 0.03$) intervocalic tokens. (The effects are essentially the same with high-pass filtered signals, but weaker and significant for only one comparison for *Delta_i*). Regarding voicing, final consonants in VC#V are significantly more voiced than initial consonants in V#CV ($\beta = -4.3$, $t = -2.9$, $p = 0.004$), but not than word-medial intervocalic consonants. However, only 3 VC#V tokens were found to be fully voiced (vs 44 in VCV and 52 VC#C tokens preceding a voiced consonant). **Conclusion.** Overall, the predicted effect of context was found. Compared with the results for Basque in [8], however, the effects of word context appear to be much weaker in Catalan, with smaller differences in reduction between VC#V and other intervocalic contexts. This suggests the existence of language-specific patterns of reduction, even when the same phonological condition (word-final devoicing) obtains.

Table 1: number of tokens by phoneme and context

	VC#C (fc)	VC#V (fv)	V#CV (i)	VCV (m)
k	29	12	423	433
p	25	5	265	191
t	227	130	84	366
TOTAL	281	147	772	990

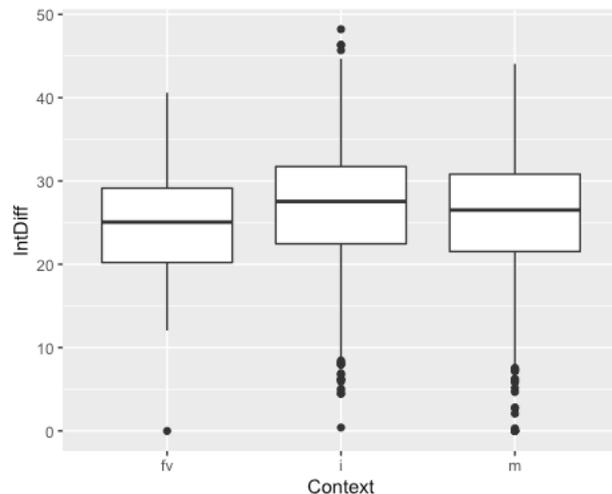


Figure 1. Intensity differences in CV window for VC#V (fv), V#CV (i) and VCV (m). Greater IntDiff indicates more constriction.

References

- [1] Bonet, E. & Lloret, M.R. 1998. *Fonologia catalana*. Barcelona: Ariel.
- [2] Carrasco, P., Hualde, J. I., Simonet, M. 2012. Dialectal differences in Spanish voiced obstruent allophony: Costa Rican vs. Iberian Spanish. *Phonetica* 69: 149–179.
- [3] Colantoni, L., Marinescu, I. 2010. The scope of stop weakening in Argentine Spanish. In: Ortega-Llebaria, M. (ed.), *Sel. Proc. 4th Conference on Laboratory Approaches to Spanish Phonology*, 100–114.
- [4] Cole, J., Hualde, J. I., Iskarous, K. 1999. In: Fujimura, O., Joseph, B. D. and Palek, B. (eds.), *Effects of prosodic and segmental context on /g/-lenition in Spanish. Proc 4th International Linguistics and Phonetics Conference*. Prague: The Karolinum Press 2: 575–589.
- [5] Eager, C. 2017. Contrast preservation and constraints on individual phonetic variation. PhD diss. Urbana-Champaign, IL: UIUC.
- [6] Ennever, T., Meakins, F., Round, E. 2017. A replicable acoustic measure of lenition and the nature of variability in Gurindji stops. *Laboratory Phonology*: 8(1), 20. DOI: <http://doi.org/10.5334/labphon.18>
- [7] Garrido, J.M. et al. 2013. Glissando: A corpus for multidisciplinary prosodic studies in Spanish and Catalan. *Language Resources and Evaluation* 47(4), 945–971.
- [8] Hualde, J.I., Beristain, A., Icardo Isasa, A. & Zhang, J. 2019. Lenition of word-final plosives in Basque. In Calhoun, S., Escudero, P., Tabain, M. & Warren, P. (eds.) *Proc of the 19th ICPhS, Melbourne, Australia*, 642–646.
- [9] Recasens, D. 2017R. The effect of contextual consonants on voiced stop lenition: Evidence from Catalan. *Language and Speech* 59(1): 139–161
- [10] Wheeler, M. 2005. *The phonology of Catalan*. Oxford: OUP.

Los Pedros y los Picapiedra: Resolving the semantic ambiguity of family names

Samuel Jambrović, *University of Toronto*

1. Introduction. Bare family names are ambiguous between reference to individuals and reference to families. For example, the family names in (1) could refer to individuals who have checked in or families who have checked in.

- (1) “Who has checked in so far?”
“García, Silva, and Smith.”

This ambiguity supports the view that proper names are predicates rather than rigid designators (Geurts 1997; Thomsen 1997). While a given name like *Ana* denotes a homogeneous set containing only individuals (2a), a family name like *García* denotes a heterogeneous set containing both individuals and families (2b).

- (2) a. $[[\mathbf{Ana}]] = \{\text{Ana (Fernández), Ana (García), Ana (López), Ana (Martín)}\dots\}$
b. $[[\mathbf{García}]] = \{(\text{Ana}) \text{García, (Pedro) García, (family}_1\text{) García, (family}_2\text{) García}\dots\}$

Since the most salient interpretation of a bare name is individual reference, languages often resort to different means to specify family reference. One possibility is using the word *family* itself: *la familia García, a familia Silva, the Smith family*. Another possibility is indicating family reference morphosyntactically: *los García, os Silva, the Smiths*. In Romance, family reference results in an exceptional construction—the combination of a plural determiner and a singular nominal element—that raises questions at all levels of the noun phrase.

2. Data. Non-lexicalist accounts of *los García* must contend with the following phenomena, which are presented using Spanish data but also occur in Portuguese and Catalan.

First, *los García* is syntactically and semantically plural (Camacho 2021).

- (3) Los García siempre piensan/*piensa en sí mismos/*mismo.
the.PL García always think.3PL/think.3SG in 3.REFL self.PL/self.SG
‘The García always think about themselves.’

The verb *piensan* ‘think’ shows plural syntactic agreement, and the reflexive anaphor *sí mismos* ‘themselves’ is a distributive predicate that indicates semantic plurality.

Second, *los García* is not a unique case of a gender and class marker “mismatch”. The same phenomenon is observed with *el atleta* ‘the.M athlete’, which requires an overt plural marker: *los atletas* ‘the.M.PL athletes’. Similarly, referring to the female members of a family, as in *las Serrano*, shows different behavior than *las testigos* ‘the.F.PL witnesses’.

Third, given names behave differently than family names. If a family has two members named “Ana”, they can only be referred to as *las Anas*, not **las Ana*.

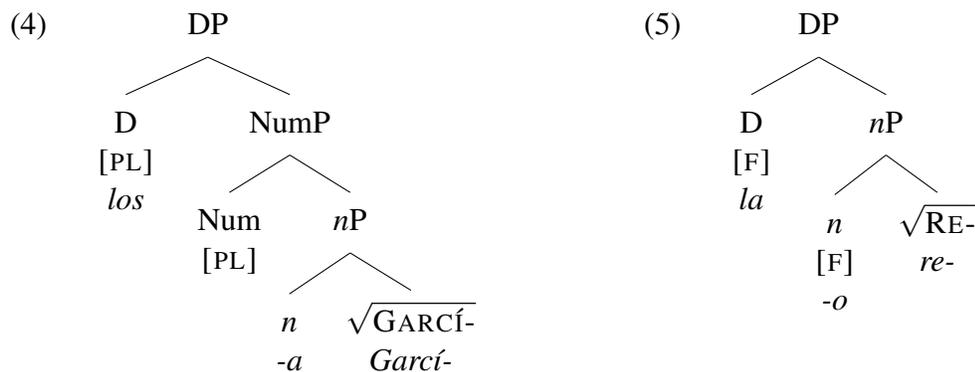
These data challenge a narrow syntactic account of *los García*, particularly the final point, which would require the syntax to differentiate between given and family names. Instead, I argue that *los García* is the result of disambiguation at the syntax-semantics interface.

3. Analysis. The existence of constructions like *los García* supports the notion that family reference is distinct from additive plurality. For instance, when *los García del vecindario* and *los Garcías del vecindario* (*del vecindario* ‘of the neighborhood’) are used contrastively, the former refers to a particular family (a collective interpretation), while the latter refers to all of the individuals named “García” (an additive plural). Given its compositional semantics, there is

no reason to suggest that *los Garcías* is generated differently in the narrow syntax than *los Pedros*, *las Anas*, and *los niños* ‘the children’. On the other hand, *los García* cannot be explained independently of the conceptual interface.

Building on Kučerová’s (2018) proposal regarding the contextual assignment of gender, I argue that number may likewise be contextually valued on D (assuming that D is a phase head accessible to the interfaces). Recall that, given the appropriate context, a bare family name is sufficient to denote family reference. However, as with other proper names, the most salient interpretation of bare family names is individual reference, leaving plural morphology as the principal means of resolving the ambiguity. In fact, the Romance languages are not the only ones in which family reference corresponds to a unique construction: Greek has a special affix for this purpose (oi Papađopoul-aí-oi ‘the Papadopouloses’), and Hungarian uses an associative marker (Nagy-èk ‘the Nagys’). In all three cases, some notion of plurality interacts with other elements of the noun phrase to disambiguate individual and family reference.

In Romance, I argue that D may be contextually valued with number features in addition to gender features, meaning that just as *n* may be valued as [F] via Agree with D, so may Num be valued as [PL]. Furthermore, in a refinement to Kučerová’s (2018) account, I propose that agreement with D may be covert, capturing the lack of *-s* on *García* in (4) and the lack of *-a* on *reo.F* ‘defendant, prisoner’ in (5); crucially, *Garcías* and *rea* are both possible forms. Since Spanish nouns require a class marker, an underspecified morpheme or epenthetic vowel is inserted at spell-out (e.g., *-o* in *reo.F* and *-e* in *jefe.F* ‘boss’).



4. Conclusion. The semantic ambiguity of family names results from their sociocultural function; it is not inherent to the grammar. Given the usefulness of being able to refer to a family as a collective entity, many languages have developed morphosyntactic strategies for specifying family reference, including “specialized” morphemes and unique patterns of agreement.

References

- Camacho, José. 2021. Last names, agreement, and the structure of DPs. Paper presented at the 39th Meeting of the West Coast Conference on Formal Linguistics, Virtual.
- Geurts, Bart. 1997. Good news about the description theory of names. *Journal of Semantics* 14(4). 319–348.
- Kučerová, Ivona. 2018. ϕ -features at the syntax-semantics interface: Evidence from nominal inflection. *Linguistic Inquiry* 49(4). 813–845.
- Thomsen, Hanne Erdman. 1997. On the proper treatment of proper names. *Nordic Journal of Linguistics* 20(1). 91–110.

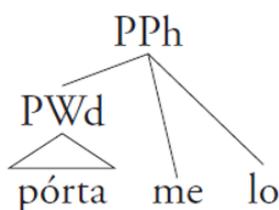
*Simple clitics in Italo-Romance: dialectal variation and
phrasal phonology*

The data that Peperkamp (1997) presents from three Italian dialects – Standard Italian, Lucanian, and Neapolitan, see (1) – have been used to argue for multiple, dialect-specific phonological clitic representations.

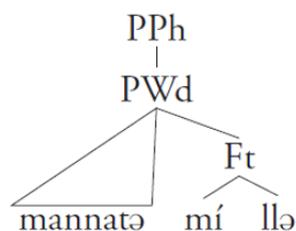
(1)	Standard Italian:	<i>Pórta</i> 'bring'	<i>Pórtami</i> 'bring me'	<i>Pórtamelo</i> 'bring me it'
	Lucanian:	<i>Vínnə</i> 'sell'	<i>Vənníllə</i> 'sell it'	<i>Vənnəmíllə</i> 'sell me it'
	Neapolitan:	<i>fá</i> 'do'	<i>fállə</i> 'do it'	<i>fattíllə</i> 'do you it'
		<i>Cónta</i> 'tell'	<i>Cóntalə</i> 'tell it'	<i>Cóntatíllə</i> 'tell it to yourself'

The prosodic representations assumed by Peperkamp are shown below in (2–4). These exemplify, respectively, adjunction of stray material to the phonological phrase, incorporation of stray material into the P-word, and the creation of a recursive P-word.

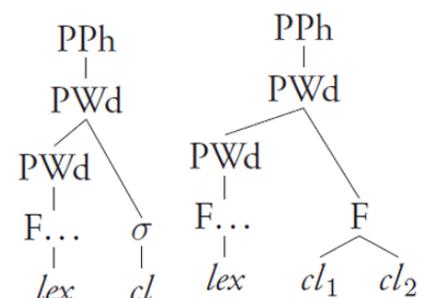
(2) Standard Italian
Adjunction



(3) Lucanian
Incorporation



(4) Neapolitan
Recursion of P-words



There are two traditions of analysis in computing these representations in OT: one assumes an interface theory between morphosyntax and phonology, while the other argues that computation is purely phonological. These are the positions that Peperkamp, building on Selkirk (1995), and Anderson (2011) take, respectively. My analysis also takes the latter view: i.e. that the computation of these representations is phonological; but crucially, this requires the adoption of a stratal grammatical architecture (Kiparsky 2010, 2015; Bermúdez-Otero 2017). This contrasts to the previously mentioned analyses, both of which assume single-level, parallel derivation in OT. I argue this is untenable for the Italian dialectal data.

We see in the data in (1) that the primary stress in Standard Italian is fixed on the base, while in Lucanian and for monosyllabic bases in Neapolitan, the primary stress shifts to the penult. For disyllabic bases in Neapolitan, an additional case of primary stress is created on the clitics: e.g. in *cóntatíllə*.

OT analysis of prosodic structures typically adheres, to some extent, to constraints that Selkirk (1995) first introduced. These can be defined as follows (own formulation):

(5) Prosodic constraints based on Selkirk (1995)

- **LAYEREDNESS:** No lower unit in the prosodic hierarchy may dominate a higher one.
- **HEADEDNESS:** Each unit must immediately dominate at least one element that comes directly below it in the prosodic hierarchy.
- **NON-RECURSIVITY:** No unit may repeat. E.g., No p-word can dominate another p-word.
- **EXHAUSTIVITY:** Every unit must exhaustively dominate the unit it immediately dominates.

These constraints are motivated from previous research into prosodic phonology, notably in connection to the Strict Layering Hypothesis (Nespor & Vogel, 1986). *LAYEREDNESS* and *HEADEDNESS* are generally considered to be undominated in all OT tableaux .

Anderson argues for a single-levelled (i.e. parallel) analysis of the data. He assumes that the lexical P-word is present underlyingly, including its primary stress, and that only the stray material is computed. He posits **CLASH* to prohibit adjacent occurrences of primary stress, and *PROSFAITH* and **STRUCT* as additional faithfulness constraints, the latter presumably undominated, and lists the constraint rankings for each dialect. However, tableaux are not provided to illustrate the derivations. Furthermore, **STRUCT* is not logically independent from *NON-RECURSIVITY*, in fact it subsumes it. It also cannot be undominated across the board because Anderson assumes a recursive structure in Neapolitan.

By contrast, Selkirk's alignment theory, applied to this data by Peperkamp, does make the correct predictions for some of the data; but unfortunately, the analysis is incomplete. It does not address the complex prosodic behaviour of clitics in Neapolitan, depending on the syllable count of the base. Additionally, the representation of Neapolitan with recursive structure does not account for Peperkamp's position that P-words have only one case of primary stress. This analysis also fails to account for the fact that monosyllabic bases in Neapolitan show identical prosodic behaviour to the Lucanian data.

In my analysis, I seek to remedy these problems by introducing new constraints, namely *PWDSTR*, *IDENT(S)* and *DEP(P-WORD)*. These formalise the requirements that P-words must have one case of primary stress, there can be no modification of the structures present in the input, and no new P-words can be created in the output. The tableau in (6) exemplifies the derivation of *cóntatilla* in Neapolitan. Here the winner is candidate (d), in which an independent P-word is formed over the enclitic sequence. Consequently, I dispute that claim that Neapolitan has recursive structure. My approach incurs no violation of the stipulation that P-words should have only one incidence of primary stress. Moreover, the stratal derivation I assume depends neither on underlying prosodic structure nor underlying stress.

(6) Sample tableau (phrase-level) for generation of Neapolitan *cóntatilla*

[Cón _{pwd} tə lə]	*CLASH	EXH _{p-ph}	PWDSTR	IDENT(S)	NONREC	DEP(P-WORD)	*GEM
a. [[Cón _{pwd} tə=lə _{p-ph}]		! **					
b. [[[Cón _{pwd} tí=llə _{pwd}] _{p-ph}]			! *		*		*
c. [[Cùn _{ta} =tí=llə _{pwd}] _{p-ph}]				! **			*
d. [[Cón _{pwd}][tí=llə _{pwd}] _{p-ph}]						*	*

Do we really need context? The role of acoustic cues in idiomatic disambiguation

Statement of the problem

The sentences in (1) exhibit an ambiguity that stems from the fact that, in addition to a compositional interpretation (1a), there is also a felicitous idiomatic interpretation (1b).

- (1) a. Tomás se apretó el cinturón b. Tomás se apretó el cinturón
 'Tomás tightened his belt' 'Tomás reduced costs'

The generative literature claims that this ambiguity is resolved in the semantics, with phonology playing little to no role in helping resolve this ambiguity (Fraser 1970; Katz & Postal 1963). Given this view, there should be no systematic differences in how the sentences in (1) are produced, and, consequently, in the absence of context, speakers should not be able to decide the correct interpretation. However, this assumption has never been tested experimentally.

In addition to the experimental motivation, recent work in Spanish phonology has found that intonation serves pragmatic purposes that move beyond marking semantic distinctions at the propositional level, as it is the case of declarative vs. interrogative statements (O'Rourke 2012). For instance, Nibert (1999) shows that by using intermediate prosodic phrases, speakers reliably signal their intended reading of otherwise syntactically ambiguous sentences. (Face 2001; Face 2002) finds that speakers manipulate the intonational structure of a sentence to signal changes in broad vs. narrow focus with early peak alignment. Rao (2013) reports that changes in F0, intensity, duration, and speech rate can be used to signal sarcasm. All three studies based their analyses on identical strings of words which give rise to different interpretations by means of intonational differences.

With this backdrop, I conducted an experiment in which I found that a dozen native speakers working under the constraints of a discourse completion task (Vanrell et al. 2018) produced literal vs. idiomatic sentences that were significantly different. Specifically, the idiomatic sentences were shorter than their literal counterparts ($p = .001$) and their F0 excursions were narrower as well ($p = .05$). I found these differences both at the phrase level and within the subject and verb constituents of the utterances. It is known that F0, intensity and duration are acoustic cues used by Spanish speakers to distinguish lexical meaning (Ortega-Llebaria & Prieto 2010). Therefore, it isn't surprising that if differences were to be found, they would involve the presence of these acoustic correlates. This production data serves as the starting input for the next phase of the experiment.

Experiment

The present experiment continues probing these differences by taking a look at the perception side of the equation. To do so, I took an output from the previously described production experiment consisting of 4 different sets of minimal pairs (literal/idiomatic) such as (1). These pairs were produced by 12 native speakers for a grand total of 96 sentences, or 48 minimal pairs. These four sets were minimally different: members of each pair were identical to one another; the only distinction was the intended meaning the speaker had in mind at the time of production. For clarity, people involved in the first experiment are called speakers. People involved in the 2nd experiment are called participants. The goal of the 2nd experiment is to find out if participants are capable of reliably picking out the intended meaning of a sentence when said sentence is heard in isolation.

Using Qualtrics as a survey platform, participants were paired up with 2 randomly assigned speakers and were asked to listen to their 4 sets of sentences wholly stripped of context. Participants were then asked to estimate the intent of the speaker by means of a forced-choice question. Following work done in English (Van Lancker & Canter 1981), I presented the stimuli in two different ways: either (a) participants saw the two members of the pair first, then were asked to guess the speaker's intent for the individual sentences; or (b) participants saw a member of the pair without the benefit of having listened to the sentences contrasted first, then were asked to estimate the speaker's intent.

Experimental evidence of palatal approximant strengthening in Medellín Spanish

Spanish /j/ strengthening involves an increase in degree of stricture and frication of the phone, yielding a number of possible phonetic outcomes including a palatal stop [tʃ], a palatal or postalveolar affricate ([tʃ̟], [dʒ̟]), and a palatal or alveopalatal fricative ([ç], [ç̟], [ʝ]) [3], [10]. Despite being attested in multiple Latin American and Peninsular varieties, to date there have been no acoustic studies dealing exclusively with the sociophonetics of Spanish /j/ strengthening (cf. [7], [8]). The variety of Spanish spoken in Medellín, Colombia serves as an ideal starting point to address this, as transcription-based evidence suggests that palatal approximant strengthening in this variety is increasingly frequent, phonetically variable, and phonologically diffuse [5]. Strengthened /j/ in Medellín encompasses almost all the possible outcomes listed above with an apparent preference for affricate-like realizations. Moreover, it occurs not only in environments understood to favor phonetic strengthening (e.g., word-initially in a stressed syllable before a high vowel, as in *yuca*, ‘yuca’) but also in those where strengthening would be less expected (e.g., word-internally in an unstressed syllable between two low central vowels, as in *malla*, ‘screen’). The reported study examines five social and linguistic factors predicted to condition /j/ strengthening and assesses the validity of two proposed acoustic correlates: **CV intensity difference** (IntDiff (dB) = Max intensity of V – Min intensity of /j/) and **harmonics-to-noise ratio** (HNR (dB) = average harmonics-to-noise ratio for the constriction interval spanning Min intensity of /j/ to Onset of V).

Data were elicited through a two-part sequential production experiment involving a sentence reading task and a sentence creation task, the former approximating careful speech and the latter approximating semi-spontaneous speech. A sample of ten college educated speakers of Medellín Spanish balanced by sex and ranging in age from 21-36 produced 1,345 analyzable tokens of intervocalic /j/. The tokens were initially segmented in *faseAlign* [9]. Interval boundaries were then visually inspected and manually adjusted in *Praat* [2] to ensure valid measurements of the two proposed acoustic correlates. Each correlate was modeled as a function of two social factors (Sex, Style) and three linguistic factors (Lexical Stress, Word Domain, and Frontness of the Following Vowel) with linear mixed effects regression using the *lme4* package [1] in *RStudio* [6]. Models included all main effects and an interaction term for Stress and Word Domain, with random intercepts added by Speaker and Word. Model evaluation was conducted using the *lmerTest* package [4]. Based on previous research as well as theoretical considerations, degree of intervocalic /j/ strengthening was predicted to be greater in female speech, in read speech, in stressed syllables, in word-initial position, and before a non-back vowel. The results of the mixed-effects linear models offer largely divergent findings that confirm only some of the predictions.

For the model with HNR as the dependent variable, significantly lower estimates indicating /j/ strengthening were found in semi-spontaneous speech ($SE=.16, t=-2.42$), in stressed syllables ($SE=.23, t=-2.46$), in word-medial position ($SE=.44, t=$ and where the following vowel was central /a/ ($SE=0.28, t=-3.55$). A trend toward /j/ strengthening was also found in the speech of female speakers ($SE=0.76, t=-1.93$). For the model with IntDiff as the dependent variable, significantly higher estimates indicating /j/ strengthening were found in read speech ($SE = 0.19, t = 11.95$), in stressed syllables ($SE=.47, t=6.53$), in word-initial position ($SE=.66, t=3.57$), and in contexts preceding a central or front vowel ($SE=0.57; 0.49, t=4.92; 2.99$). A significant interaction was also found between *stress* and *word domain* ($SE=1.01, t=-2.14$), as unstressed, word-initial tokens of /j/ showed higher average estimated IntDiff values than unstressed, word-internal tokens.

The trend among female speakers to exhibit overall stronger tokens of /j/, at least in terms of HNR, lends some tentative support to prior findings showing /j/ strengthening to be an incipient change in progress in Medellín Spanish [5]. In addition, model agreement regarding the effect of non-back vowels and lexical stress points to the complementary roles of coarticulation and prosodic structure in facilitating /j/ strengthening. That the models disagreed concerning the effect of the remaining predictors is attributed to the small sample size and limitations inherent to the two proposed acoustic correlates. Specifically, while IntDiff adequately captures variation in degree of stricture, it might not be as robust a measure of frication noise, a key acoustic property of /j/ strengthening in Medellín Spanish [7]. HNR might also fail to fully capture variation related to frication noise due to a tendency among female speakers to devoice the frication interval. This tendency degrades the harmonic component of the signal to such an extent that there are no measurable glottal pulses even after voice optimization, leading to underestimation or omission of the frication noise present in the interval. Further analysis using a third, potentially more robust acoustic measurement—zero-crossings rate (ZCR)—will be carried out in the interim.

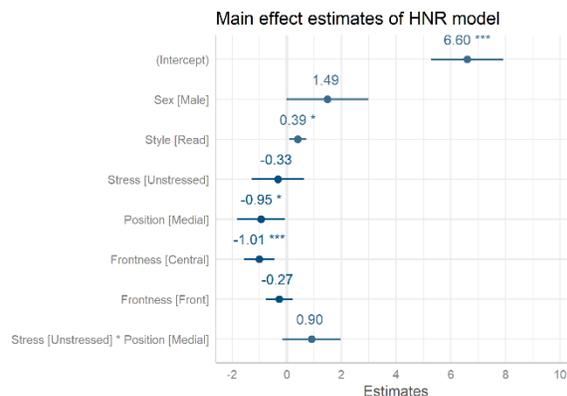


Figure 1. Forest plot of HNR linear regression estimates. Intercept ($\beta=8.11$): Sex [Female], Style [Original], Stress [Stressed], Position [Word-initial], Frontness [Back vowel].

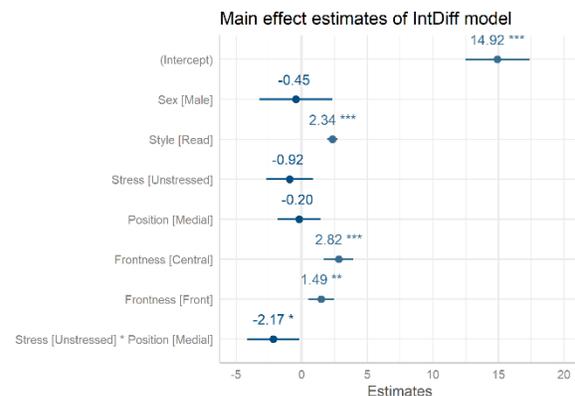


Figure 2. Forest plot of IntDiff linear regression estimates. Intercept ($\beta=11.18$): Sex [Female], Style [Original], Stress [Stressed], Position [Word-initial], Frontness [Back vowel].

References

- [1] Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- [2] Boersma, P. & Weenink, D. (2016). *PRAAT: Doing phonetics by computer*. Version 6.1.16
- [3] Hualde, J. (2005). *The sounds of Spanish*. Cambridge, UK: Cambridge University Press.
- [4] Kuznetsova A., Brockhoff, P.B., Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1-26.
- [5] Leslie, T. (2017). Social and linguistic predictors of /j/ in Medellín, Colombia. Unpublished manuscript.
- [6] RStudio Team (2020). *RStudio: Integrated Development for R*. RStudio, PBC, Boston, MA.
- [7] Scarpace, D., Beery, D., & Hualde, J. I. (2015). Allophony of /j/ in Peninsular Spanish. *Phonetica*, 72, 76–97.
- [8] von Essen, M.C. (2016). Variedades del español en contacto: acomodación sociolingüística de una comunidad de inmigrantes argentinos en la ciudad de Málaga. Análisis acústico de las variantes alofónicas de /j/. *Lengua y migración*, 8(2), 7-43.
- [9] Wilbanks, E. (2020). faseAlign (Version 1.1.10) [Computer software]. Retrieved Oct 04, 2020 from <https://github.com/EricWilbanks/faseAlign>.
- [10] Zampaulo, A. (2019). *Palatal Sound Change in the Romance Languages: Diachronic and Synchronic Perspectives*. Oxford University Press.

Inhibition vs. Lexicon in Implicature Generation among Bilingual Speakers

In monolingual English-speaking children and monolingual Spanish-speaking children, distributive and collective interpretations are not adult-like until roughly 10 years-old (e.g. Hanlon 1986, Brooks & Braine 1996, Musolino 2009; Roeper et al. 2011, Pagliarini et al. 2012, Syrett & Musolino 2013; Padilla-Reyes et al. 2016, de Koster et al. 2017, de Koster et al. 2018). Dotlačil (2010) claims that collective interpretations of generalized quantifiers such as “some” and “the”, which, though inherently ambiguous between distributive and collective readings, derive their collective meaning from their position on a distributive-collective pragmatic scale. This scale is anchored by the unambiguous distributive entailment of “each”, which confers upon “some” comes a collective meaning by informativeness implicature. Monolingual Spanish and English-speaking children’s collective and distributive interpretations are predictable from their lexical development and their collective, but not their distributive, interpretations are predictable from their inhibitory executive function abilities (Padilla-Reyes et al. 2016; Grinstead et al. 2018).

Bilingual children have been claimed, controversially, to possess greater inhibitory abilities than monolingual children (e.g. Bialystok et al. 2012; Paap et al. 2015). Further, it has been claimed, again controversially, that bilingual children pass through a slower lexical developmental trajectory in each of their languages than do monolingual children (Ben-Zeev, 1977; Pearson et al., 1993; Bialystok et al., 2010; Bialystok and Luk, 2012; De Houwer, et al., 2014). Given this state of affairs, we might expect bilingual Spanish-English-speaking children to develop collective implicature-generating capacities **more** quickly than monolinguals, if their executive function abilities are more advanced. To the contrary, we might expect that their implicature-generating capacities would develop **less** quickly than monolinguals if their lexical development is less advanced. Determining which factor is (more) predictive could tell us whether lexicon or executive function plays a more important role in collective implicature generation.

To investigate this question, we tested a monolingual, typically-developing sample of child Spanish-speakers in Mexico City (n=11, mean age=96 months, age range – 84-98 months, SD=5.3 months) and an age-matched, bilingual, typically-developing sample of child English-Spanish-speakers in the US Midwest (n=11, mean age=96 months, age range=84-112 months, SD=8.9 months). Children were presented with a video-recorded Truth-Value Judgement Task (Crain & McKee 1986), including 12 experimental items with the Spanish plural determiner “unos” presented half in collective contexts and half in distributive contexts and another 12 experimental items with the Spanish distributive universal quantifier “cada”, again half in collective and half in distributive contexts. Children were also given the Spanish Peabody lexical measure (Dunn et al. 1986) and the Flanker test of inhibition.

Results showed that bilingual children had greater inhibition scores than did monolinguals (Figure 1) and that monolingual children had greater lexical scores than did bilinguals (Figure 2). To answer our core questions, monolinguals generated more collective implicatures (less acceptance of “unos” in distributive contexts) than did bilinguals (Figure 3). Lexicon, but not executive function, was predictive of collective judgments (Figure 4). These findings lend greater clarity to the relative roles played by lexicon and executive function in scalar implicature generation. **493 Words**

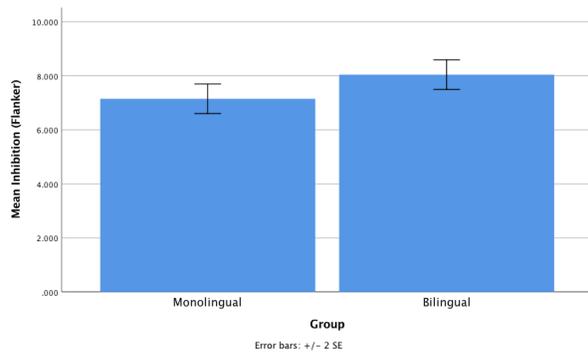


Figure 1 – Bilinguals Have Greater Inhibition (Flanker) Scores Than Monolinguals ($t(20) = -2.300, p = .032$)

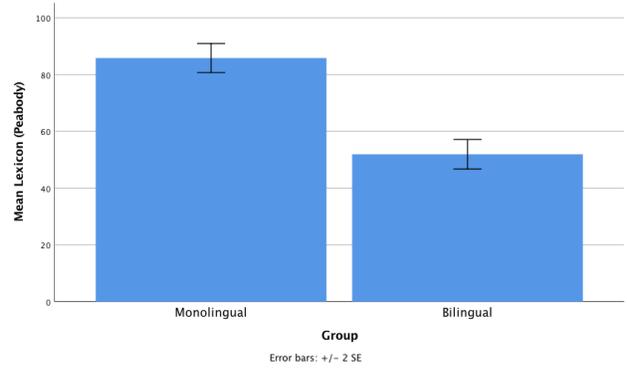


Figure 2 – Monolinguals Have Greater Lexical (TVIP) Scores Than Bilinguals ($t(20) = 9.287, p < .001$)

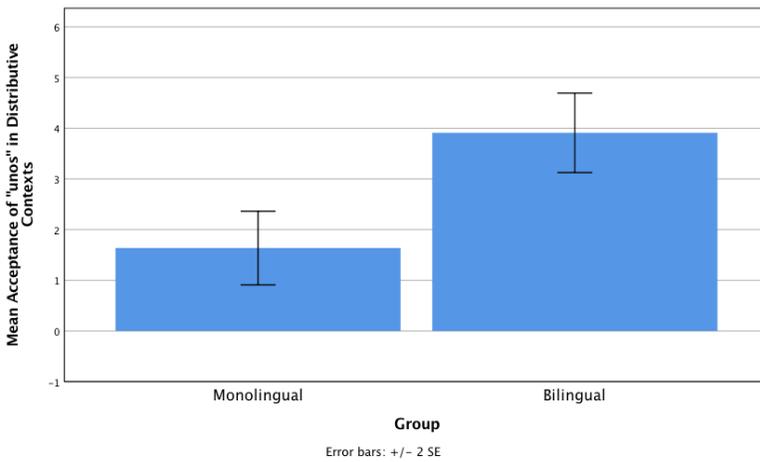


Figure 3 – Monolinguals Generate More Collective Implicatures than Bilinguals ($t(20) = -4.250, p < .001$)

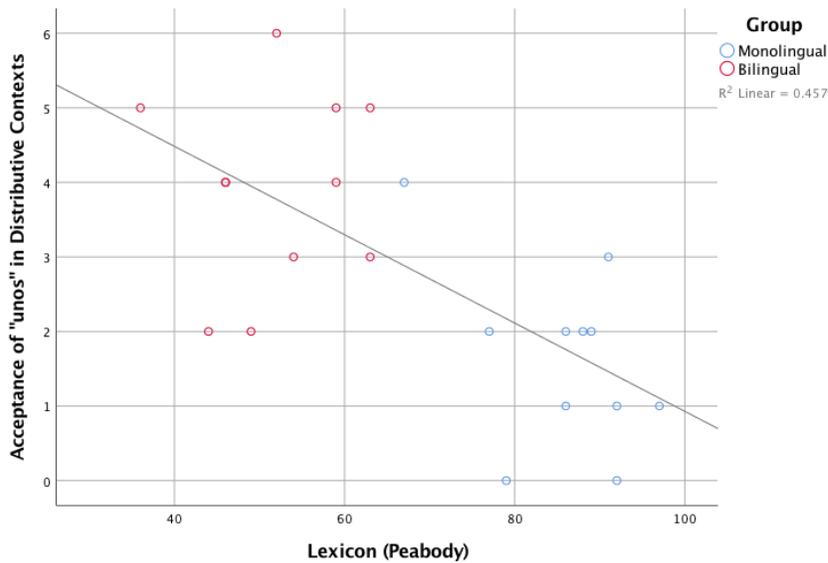


Figure 4 – Lexicon (TVIP) Predicts Collective Judgements with "unos" in Distributive Contexts ($B = -.059, SE = .014, p = .001$)

Intransitive Causatives in English and Romance

Maria Eugenia Mangialavori
CONICET

Josep Ausensi
UPF

We focus on Intransitive Causatives (ICs), an underexplored argument structure realization in verbs entering causative alternation, with key implications for argument structure and verb formation/derivation. We focus on the asymmetry between English (1) and Romance (2).

- (1) a. Smoking kills. b. Bleach disinfects. c. Alcohol dehydrates. d. Rice constipates.
e. Shaving creams irritate. f. Sunlight oxidizes and discolors. g. Normal dryers wrinkle.
- (2) a. Fumar mata. b. La lejía desinfecta. c. El alcohol deshidrata. d. El arroz estríñe. e. Las cremas de afeitar irritan. f. La luz solar oxida y destiñe. g. Las secadoras arrugan.

ICs feature an (inanimate) subject interpreted as possible cause of change of state (COS). Crucially, the undergoer is semantically and syntactically unrealized. ICs challenge major claims on argument structure: e.g., **(i)**the idea that the internal argument is an invariable constituent in the causative alternation (Haley & Keyser 2002); **(ii)**the prediction that unique arguments in COS verbs are, by default, undergoers (Levin & Rappaport Hovav 2005), as ICs show a defective cause interpretation of sole arguments in causative verbs is possible/natural.

Properties. (a) ICs pattern as states: they cannot appear in perception reports or located in space (3) or license habitual readings (4), are odd in contexts forcing eventive readings (5), and modals generate epistemic (not deontic) readings (6) (Maienborn 2005, Rothmayr 2007).

- (3) a. #John saw smoking kill. (cf. *I saw John kill Tom*)
b. #Shaving creams irritate in the bathroom.
- (4) a. Alcohol (#regularly) dehydrates. (cf. *John regularly kills animals*)
b. Smoking (#regularly) kills.
- (5) a. #What the dryer did was wrinkle. (cf. *What John did was kill animals*)
b. #What happened was that rice constipated.
- (6) a. Smoking must kill (OK It probably has property x | #It is under obligation to kill)
b. John must kill Tom (# He probably has property x | OK He is under obligation to kill)

(b) Like middles, ICs are restricted to generic tenses, as expected from states (**Smoking killed/#This vase broke easily*). Further, both constructions pattern as dispositional generics in that they do not entail a deontic reading, but rather report a property of the subject (Lekakou 2015)(i.e., a state). IC sentences are true in virtue of the properties inherent to the subject, rather than whether there were actual events of the specific type in the past (cf. *Chromic acid burns (that is why it has never been used before)* and *This vase breaks easily (that is why it is kept inside the box)*). ICs and middles thus contrast with dispositional habituals that “assert the existence of a pattern of regularly recurring events” (Krifka et al. 1995) (true insofar as there were actual *helping* events in the past, e.g., *John helps homeless people*). Yet, ICs differ from middles as the sole DP is not an internal but an external argument: hence, the property is not attributed to the undergoer, but to the cause(r). ICs reflect the definition of dispositional causation (Copley 2018) relating a disposer *y* (holder of a property), a dispositional state *e*, a manifestation *e'*, and a (nonepisodic) eventuality description *p*. This captures ICs restrictions: i.e., the cause(r) must have the relevant property to produce the COS of the verb (Fara 2001).

- (7) Dispositional causation: (a) *y* is the holder of *e*, (b) *e* is a state that directly causes *e'* ceteris paribus, (c) *e'* instantiates *p* (d) *y* is disposed toward *p*. (Copley 2018: 13)

(c) The key **difference between Romance and English** lies in IC productivity: while Romance freely allows ICs from classes standardly related to the causative alternation like psych verbs, English has to make recourse to stative-attributive *ing*-predicates, suggesting nontrivial differences yet to be explained and raising questions on crosslanguage availability of ICs (8).

- (8) a. La playa cansa. (lit. *The beach tires) b. Tarantino aburre. (lit. *Tarantino bores)
'The beach is tiresome/makes you tired.' 'Tarantino is boring/makes you bored.'

In English, *be-ing* appears to be the default mechanism to denote that a cause(r) has potential to trigger COS (cf. *La leche engorda* 'Milk is fattening (≠is getting fat)' vs. ??Milk fattens). ICs are possible when *be-ing* does not yield IC interpretations: as *be-ing* is not available for dispositional causation, English resorts to the intransitive causative variant to express IC (9).

- (9) a. Sad movies are depressing. (OK Causational reading = #IC: #Sad movies sadden)
 b. Normal dryers are wrinkling. (#Causational reading = OK IC: Normal dryers wrinkle)

(d) Even if **genericity** is a common property shared with another dyadic/monadic argument structure alternation (Unexpressed/Null Object alternations [UNOA], Levin 1993), ICs are clearly different (e.g. vs. Characteristic Property of Agent Alternation (10), Levin 1993) in: **(i)** verb type (activity/manner verbs in (10) vs. COS/result verbs in ICs); **(ii)** interpretive and selectional restrictions on the subject (animate/volitional actor in (10) vs. inanimate causer in ICs). Several facts indicate they constitute a radically different type of intransitive alternation.

- (10) a. This dog bites (#but hasn't bitten anybody yet).
 b. Stand back! This horse kicks (#but hasn't kicked anybody yet).

(e) Unlike UNOA, ICs are not Null-Object constructions: ICs do not allow Null-Object-oriented depictives (11) (OK with null/arbitrary implicit arguments: *Il dottore visita [] nudi*, 'The doctor visits [] naked', Rizzi 1996) and null object quantification (12) (e.g., bare *molti*, Italian). *Neccliticization* and inchoative/passive morphology (Romance) are also disallowed (13). Last, ICs fail to bind reflexive pronouns (vs. the anticausative (*se-cl*) form) (14).

- (11) a. *Smoking kills dead/depressed. b. John cooks healthy. c. John buys cheap.
 (12) a. *Smoking kills a lot. b. John eats a lot. c. John bought some.
 (13) a. Fumar (*en/*es) mata. b. El Joan en cuina/compra (cada dia). (Catalan)
 (14) a. Bad news sadden (*myself). b. Take a crepe. Cover one half with the jam.
 Fold [] over onto itself. (Massam & Roberge 1989)

Proposal. We propose ICs are monadic (atransitive) realizations where the external-argument-introducing head responsible for the causative component is merely complemented, not by a theme, but by mere rhematic information (RhemeP, Ramchand 2008, 2013) \checkmark specifying the COS potentially triggered by the subject (vP [DP CAUSE/TRIGGER [v INIT^o, RHEME \checkmark]]). As for syntax-semantics interface and direct mapping between semantic (event) composition and argument structure realizations, ICs show that nonrealization of the internal argument correlates with lack of COS (sub)event instantiation (attributed to the internal-argument licensing head, Levin&Rappaport 1995; Hale&Keyser 2002). The noneventive denotation, along with pure stative behavior, simply follow. ICs crucially show that if there is no theme, there is no COS-event-encoding component in the semantic/syntactic makeup of the VP. This allow us to avoid a derived analysis of stativity, with direct empirical evidence and more parsimony (assuming economy is a result desirable in GG). Syntactically, it reveals that Property of Agent Alternations (10) and middles are not true argument structure alternations as they are alternate expressions of the same set of arguments: the former can be analyzed as Null/Unexpressed Object alternations, i.e., they are underlyingly transitive structures, while ICs behave instead as original monadic. We also note that, across languages, ICs appear in verbs denoting COS-caused state, rather than (manner of) action (Rappaport&Levin 2010). If the verb encodes a manner of action, and there is no possible state-like interpretation the subject is able to trigger, ICs fail and yield Property of Agent Alternation interpretations (10) in monadic frames (16). Namely, verbs like *kill* which only allow a result/state interpretation, freely allow both constructions depending on subject type (15). Conversely, verbs like *murder* cannot allow ICs as the verb makes references to both a manner and a result (Ausensi 2019) and only yield Property of Agent Alternation readings in monadic frames in consequence, as they restrict the subject to a specific type, an Agent in this case (cf. (10)).

- (15) a. John kills (impulsively). (Property of Agent alt.) b. Smoking kills. (IC)
 (16) a. CIA spies murder silently. (Property of Agent alt.) b. #This poison murders. (IC)

ICs thus challenge the long-held constraint (Rappaport&Levin 2010) that COS verbs disallow unrealized theme (**John breaks*) showing atransitive constructions with expected properties.

Conclusion. ICs: ● establish interesting crosslanguage regularities; ● reveal important structural (syn/sem) consistencies; ● raise questions on lexical coding of relevant features; ● uncover a necessary contrast within intransitive alternations (UNOA vs. true atransitive (IC)).

Sound change and rhythm in Altiplateau Mexican Spanish

The issue of vowel compression in Spanish is much debated. In this paper, we present evidence from Altiplateau Mexican Spanish (AMS) that coda-driven vowel compression is a variety-specific phenomenon in Spanish.

Vowel compression refers to the phonetic shortening of vowels in specific phonological contexts. In languages in which it occurs (i.e. English [1]), compression has been shown to depend on syllable structure: i.e. the greater number of segments within a syllable constituent, the more extreme the shortening. Thus, the nuclear vowel in an English CCVCC syllable like *dreads* exhibits more extreme compression than in a CVC syllable like *dead*. Maddieson's [1] *Closed Syllable Shortening* principle in fact claims that compression is coda-driven in all languages, i.e. vowels in closed syllables are predicted to be universally shorter than those in open syllables.

Language-specific research, however, shows that this principle is not as universal as originally claimed. Katz [2] showed that, in English, durational differences between vowels in open and closed syllables were negligible, and that coda complexity has no effect on degree of compression. Moreover, data reported in Aldrich & Simonet [3] reveal that onset complexity causes greater compression than coda complexity in Spanish. In both studies, onset-induced compression was consistent across all speakers whilst coda-induced compression was more variable and speaker-specific in nature. Particular to the present study is the following: "we are aware of no published finding suggesting that, in Spanish, coda presence (or complexity) drives compensatory vowel shortening" ([3]:268).

Whilst this statement holds true for Aldrich & Simonet's [3] aggregate findings—based on a dialectal cross-section of Spanish varieties—data from AMS reveals that coda-induced compression does indeed occur in this variety. Accordingly, we observe pairs such as those in (1) below in which the closing of an open syllable (here through inflectional operations) triggers compression of the nuclear vowel.

(1) Vowel compression in AMS

[ˈgran.ðe]	[ˈgran.ðēs]	[es.ˈta.βa]	[es.ˈta.βã]
'big' (SING)	'big' (PLURAL)	'he/she/it was' (3 rd PERSON SING)	'they were' (3 rd PERSON PLURAL)

Our data comes from an experiment designed to test for coda compression effects in AMS. This originated in impressionistic reports that this variety, spoken in Mexico's Central Highland region, displayed perceptible vowel reduction of closed-syllable contexts. Realisations of mid and low stressed and unstressed vowels (/e, o, a/) were compared in word-final contexts in a dataset comprising recordings of continuous read speech from 6 female speakers of AMS. Comparison of durational and formant-frequency measurements reveal that coda-driven vowel compression shortens, centralises and reduces the intensity of unstressed vowels in this variety. More specifically, it is noted that these effects are:

1. Phonologically constrained: compression occurs in word-final, post-tonic closed syllables. Although complex onsets partially shorten the vowel, these are negligible when compared to coda effects.
2. Phonetically constrained: compression is most perceivable when the coda is /s/.
3. Prosodically constrained: compression is most extreme in domain-final positions of large prosodic units, i.e. utterance phrases, rather than smaller ones, i.e. syllables.

These results therefore suggest that coda-driven compression does occur AMS, and that this is potentially a variety-specific phenomenon. The existence of the compression effects summarised

above bear further on a theoretical point relating to phonological timing in Spanish: namely, that with regard to compression and vowel reduction, dialect-specific phonology-phonetics interactions cause certain varieties of Spanish, like AMS, to behave in a way typically associated with stress-timed languages, i.e. English. This is significant given that previous analyses of phonological timing categorised Spanish as a syllable-timed language ([4]; [5]).

In line with this, we link our results to wider debates concerning phonological rhythm, specifically the role of timing and prominence. Firstly, the variation in both quantitative and qualitative properties between stressed and unstressed vowels aligns with the theory that phonological rhythm does not solely arise from timing, but rather from the marking of prominence. Although timing may play a role in marking prominence, other acoustic properties, e.g. intensity, formant-frequency, may interact to indicate stress. As such, the terms *stress-based* and *less stress-based* seem more appropriate for a cross-dialectal characterisation of Spanish than a more fixed dichotomous categorisation of stress- and syllable-timed languages ([4]). Secondly, our results contribute further evidence in support of the claim that there is not one universal metric for rhythm. On the contrary, rhythm is a complex phenomenon that may depend, to various degrees, on phonological and prosodic constraints of the language and variety, as well as dialect-specific phonetic interactions with phonology ([6]; [7]; [8]). Thus, the acoustic realisation and uniformity of rhythm (prominence) are conditioned by unique phonologies of specific language varieties ([8]; [9]).

In view of our results from AMS, we argue that further work is needed into the apparent fluidity in how languages, and specific varieties within them, express prominence ([6]; [7]; [8]; [9]). We see particular advantages in additional research that would seek to understand variation in prominence in Spanish along a dialectal continuum, potentially crossing into the phonetic territory of what have previously been thought to be rhythmically distinct languages, e.g. Portuguese [10].

References

- [1] Maddieson, I. (1985) Phonetic Cues for Syllabification. In: V. A. Fromkin (Ed.), *Phonetic Linguistics: Essays in honor of Peter Ladefoged*. 203–221, New York, NY: Academic Press
- [2] Katz, J. (2012) Compression Effects in English. *Journal of Phonetics* 40: 390–402.
- [3] Aldrich, A.C & Simonet, M. (2019) Duration of syllable nuclei in Spanish. *Studies in Hispanic and Lusophone Linguistics* 12(2): 247–280.
- [4] Duaer, R. M. (1983) Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51–62.
- [5] Pike, K.L. (1945) *The intonation of American English*, Ann Arbor: University of Michigan
- [6] Arvaniti, A. (2009) Rhythm, Timing and the Timing of Rhythm. *Phonetica* 66: 46–63.
- [7] Arvaniti, A. (2012) The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40: 351–373.
- [8] Turk, A. & Shattuck-Hufnagel, S. (2013) What is speech rhythm? A commentary inspired by Arvaniti & Rodriguez, Krivokapi, and Goswami & Leong. *Laboratory Phonology*, 4(1): 93–118.
- [9] Smith, R. & Rathcke, T. (2020) Dialectal phonology constrains the phonetics of prominence, *Journal of Phonetics*, 78: 1–17.
- [10]. Frota, S. & Vigário, M. (2001), On the correlates of rhythmic distinctions: The European/Brazilian Portuguese case, *Probus* 13: 247–275.

On the Acquisition of Spanish Psych Predicates: When Intervention Makes 'Subject' Extraction Harder

Victoria Mateu (UCLA)

Background. Prior studies have reported an advantage in the comprehension of subject over object *wh*-questions (e.g., Friedmann et al. 2008). Under a theory of (featural) Relativized Minimality (fRM; Rizzi, 1990, 2004) the subject>object asymmetry can be explained as an effect of intervention and the assumption that children's grammars are more restrictive than adults'.

Experimental L1 acquisition studies typically test this hypothesis using actional verbs and S/O extraction (ACTs, e.g., *chase, kick, hug...*), and not psychological predicates (PSYs, e.g., *love, bother, fear...*) or other types of argument extraction. Additionally, there is no known experimental work on the L1 acquisition of *wh*-questions in Spanish and very little on psychological verbs like *gustar* ('to please') (cf. Soler, 2012; Torrens et al., 2006).

In Spanish, Class III PSYs like '*gustar*' (Belletti & Rizzi, 1988) (1a) project a different structure than ACTs (1b): the 'object' of PSYs, i.e. the experiencer DP, obligatorily preceded by the dative marker 'a', '*a la niña*' ('the girl') in (1) is generated in a *higher* syntactic position (high applicative phrase, Cuervo, 2003) than the 'subject', i.e. the nominative theme DP, with which the verb agrees in person and number, '*las maestras*' ('the teachers') in (1a).

- (1) a. Canonical word order for PSY *gustar*: 'Object' (experiencer) - Verb - 'Subject' (theme):
A la niña le gustan las maestras.
to the girl 3DAT.SG like-3.PL the-PL teacher-PL
'The girl likes the teachers' or 'The teachers please the girl'
- b. Canonical word order for ACT *gritar*: Subject (agent) - Verb - Object (goal):
La niña les grita a las maestras.
the girl 3DAT.PL yell-3.SG to the-PL teacher-PL
'The girl yells at the teachers'

Therefore, while fRM predicts we should find a S>O asymmetry with Spanish ACTs (1b), it predicts an 'O'>'S' asymmetry with PSYs (1a).

Corpus study. We first conducted an exhaustive search of all 19 Spanish-speaking corpora on CHILDES (MacWhinney, 2000) comparing children (<7yo) to adults with respect to the frequency with which each group produces the order Experiencer-clitic-*gustar*-Theme as opposed to Theme-clitic-*gustar*-Experiencer. The intervention hypothesis leads to expect children will avoid moving the Theme past the Experiencer to a higher degree than adults and so children will produce fewer Theme-over-Experiencer sentences than expected given their input. Results show that while both adults and children produce Experiencer-first utterances ($N = 256, 139$) more so than Theme-first utterances ($N = 81, 8$), children do so to a significantly higher degree than adults (Fisher's exact test, $p < .001$), even when we exclude all instances of A'-movement ($p = .006$), in line with RM.

Experimental study. We followed up with an online picture-selection task paired with d-linked *wh*-questions that featured the Class III PSY verbs '*gustar*' ('please') and '*molestar*' ('bother') and the ACT verbs '*gritar*' ('yell') and '*leer*' ('read'), which can be superficially analogous (see 2-3). So far, we have collected data from 11 monolingual Spanish-speaking adults and 12 children aged 4-6 (average age = 5.3). The task consisted of 4 training items, and 32 test items, which were balanced for Verb Type (PSY/ACT), Extraction Site (S/O), and Feature Match (M: SG-SG/ MM: SG/PL). The prediction of fRM is that we will replicate the S>O asymmetry with ACTs but will find an 'O'>'S' asymmetry with PSYs. Also, morphosyntactic feature mismatches between the moved element and the intervener should ameliorate children's performance on intervening configurations more so than on non-intervening configurations.

- (2) a. *Qué niña le gusta a la maestra _?* PSY ‘Subject’ *wh*-Q
 ‘What girl does the teacher like _?’ or ‘What girl _ pleases the teacher?’
 b. *A qué niña _ le gusta la maestra?* PSY ‘Object’ *wh*-Q
 ‘What girl _ likes the teacher?’ or ‘What girl does the teacher please _?’
- (3) a. *Qué niña _ le grita a la maestra?* ACT Subject *wh*-Q
 ‘What girl _ yells at the teacher?’
 b. *A qué niña le grita la maestra _?* ACT Object *wh*-Q
 ‘What girl does the teacher yell at _?’

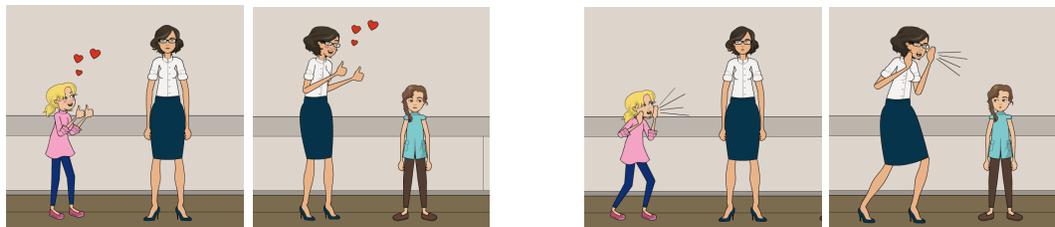


Figure 1. Scenario setups for *gustar* and *gritar* (Match condition).

Results. The results we have obtained thus far are in Tables 1 (adults) and 2 (children).

Table 1. Adult results.

	ACT	PSY
S, M	90.9%	75.0%
S, MM	100.0%	88.6%
O, M	93.2%	90.9%
O, MM	97.7%	95.5%
S	95.5%	81.8%
O	95.5%	93.2%
GR AVE	95.5%	87.5%

Table 2. Child results.

	ACT	PSY
S, M	91.7%	14.6%
S, MM	95.8%	31.3%
O, M	31.3%	95.8%
O, MM	64.6%	91.7%
S	93.8%	22.9%
O	47.9%	93.8%
GR AVE	70.8%	58.3%

We ran a mixed effects logistic regression model with the child data which included score as a binary dependent variable, Age, VerbType, ExtractionSite, FeatureMatch, and relevant interactions as fixed effects, and verb and participant as random intercepts (best model by fit comparison). Results show that overall children do better with ACT verbs ($p < .001$), but importantly, there is a strong interaction between VerbType and ExtractionSite ($p < .001$) such that subject-*wh*-Qs are harder than object-*wh*-Qs for ACT verbs, but the opposite is true for PSY verbs. We also find an effect of FeatureMatch such that feature matching impairs children’s performance on intervening structures (O for ACT, $p = .05$ and S for PSY, $p = .001$) but not on non-intervening ones.

Our results strongly suggest that children’s difficulties are not based on canonical word order or frequency effects, but rather support: (a) syntactic analyses proposed for Class III PSY verbs (B&R, 1988), and (b) fRM (Rizzi, 2004), which proposes that intervention effects are mitigated by morphosyntactic feature mismatches.

Selected References. Friedmann, N. A. Belletti & L. Rizzi. 2008. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119. 67-88. • Soler, I. G. (2012). Acquiring Spanish Psych-Verbs: Maturation or Continuity? In *Suppl. Proceedings of BUCLD 36*. Somerville, MA: Cascadilla. • Torrens, V., L. Escobar, & K. Wexler. (2006). The acquisition of experiencers in Spanish L1 and the external argument requirement hypothesis. In *The Acquisition of Syntax in Romance Languages*. John Benjamins. 183-202.

Which factors affect the production of Overt pronouns in Catalan?

Laia Mayol (Universitat Pompeu Fabra)

Background. The literature on Romance null-subject languages has often postulated a division of labor between Null and Overt pronouns: Nulls prefer to retrieve an antecedent in subject position, whereas Overts prefer an antecedent in a lower syntactic position (Carminati, 2002). However, in some previous work (Authors 2018), we found that even if Overts showed a clear non-subject bias interpretation, they were scarcely produced to refer to non-subjects and that speakers preferred to produce Names. The goal of this research is to further study the production of Overt pronouns in contexts in which they are optional (i.e. non-focal), by means of a discourse completion study. In particular, we manipulate three factors that may affect antecedents' prominence: (i) their syntactic function, (ii) the implicit causality bias of the verb and (iii) whether the two entities are introduced by names conventionally associated with the same gender.

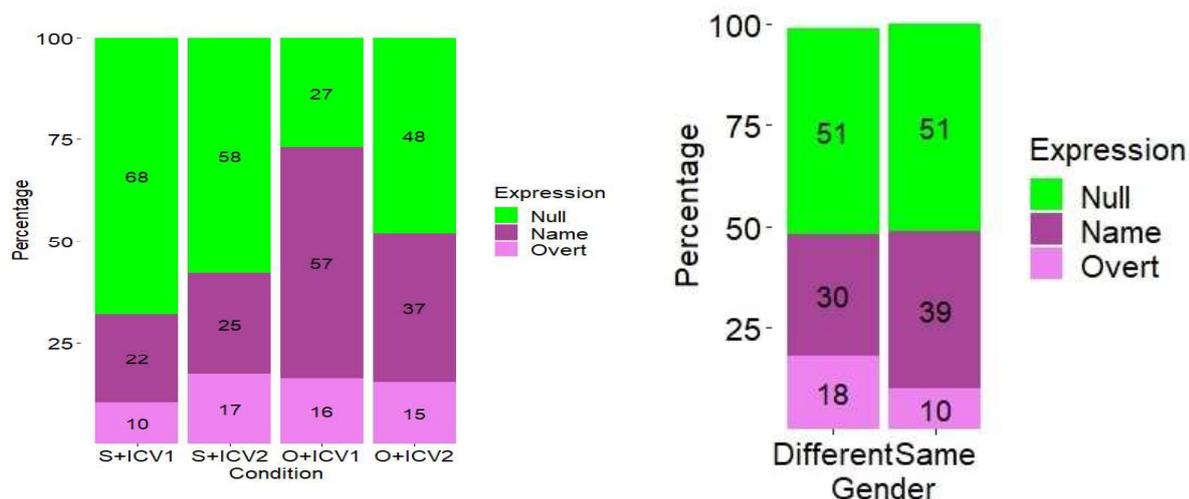
Hypotheses. (i) Based on previous work, we expect that Overts will be mostly used to refer to the referent in object position. (ii) The role of verb bias on the choice of referring expressions is controversial. In their Bayesian account, Rohde & Kehler (2014) argue that, while structural factors (e.g. syntactic position) have an effect on the production of referring expressions, semantic factors (e.g. verb bias) do not. In contrast, the Expectancy Hypothesis of Arnold (2010) claims that semantic factors do affect production of expressions. (iii) Since Overts carry gender features, we expect Overts to be produced more in context in which they have the potential to disambiguate: i.e. when the two potential antecedents have different genders.

Experiment design. A discourse-completion study was carried out in which participants were asked to imagine they were talking to a friend and had to answer a question their friend was posing. The context sentences used implicit causality verbs, which impute the cause of the event they denote either to the subject (ICV1; *intimidate*) or the object (ICV2; *congratulate*). Since ICVs elicit their biases in the context of an Explanation rhetorical relation, the question posed by the friend was always a causal one ('Why?') in the critical items. In addition, either the subject or the object of the context sentence was underlined and participants were told to write a continuation about the underlined referent by placing this referent in the subject position of their completion. Finally, the two entities in the context sentence were introduced with proper names of either different gender or same gender. (1) shows examples of the Subject + ICV1 + Same Gender (Different Gender) conditions. The experiment contained 16 critical items (8 ICV1s and 8 ICV2s) and 16 fillers. Four lists were generated, following a 2x2 design.

- (1) a. Tu: El Pol intimida el Carles (la Maria). 'You: Pol intimidates Carles (Maria).'
b. El teu amic: Per què? 'Your friend: Why?'

Data coding. 72 Catalan native speakers participated in the experiment, yielding 1152 completions. Since the design was fairly complex, the data was checked to see if the subject of the completion referred to the underlined entity in the context sentence. 57 completions were excluded from the analysis because they either contained a nonfinite clause, were ambiguous or were about a third referent. Of the remaining 1105 completions that were about one of the two referents in the context sentence, 90% correctly referred to the underlined entity and 10% did not. Mistakes concentrated on the combinations in which the verb bias and the syntactic position of the underlined entity did not match: ICV1 + O (44% of the mistakes) and ICV2 + S (38% of the mistakes). In the 991 correct completions, we coded the type of referring expression used in subject position. The observed referring expressions are the following: Null pronouns (50.9%, N=504), Names (34.8%, N=345), Overt pronouns (14.1%, N=140), demonstratives (0.1%, N=1), right-dislocations (0.1%, N=1). Given the extremely low number of demonstratives and right-dislocations, they are excluded from the analysis.

Results. The figure on the left shows the distribution of the three examined referring expressions in each syntactic position and verb bias combination. Nulls show a clear subject preference, although modulated by verb bias: the subject preference is stronger with the expected verb bias (ICV1) and the object dispreference is stronger when the verb bias is unexpected (ICV1). This



is reversed for Names, which show a clear object preference, which is stronger with the unexpected verb bias (ICV1). In contrast, the effect of syntactic position is milder with Overt, which disprefer highly salient antecedents (subject antecedents with the expected ICV1 verb bias).

The figure on the right shows the effects of same/different gender of the potential antecedents. While this manipulation does not affect the distribution of Null pronouns, it does affect Names and Overt: Names are favored in contexts where the two names in the context sentence are conventionally associated to the same gender, while Overt pronouns are favored in contexts where they are associated to different genders.

In order to test for the statistical significance of the data, we performed an analysis on two subsets of the data to make the choice of referring expression binary: (i) choice of Overt vs. Null and (ii) choice Overt vs. Name. Mixed-effect logistic regressions were performed, using R and *lme4* (Bates et al., 2015). We obtained the optimal model by comparing models with a likelihood-ratio test. For the first subset, the optimal model is the full model containing syntactic function, gender sameness and verb bias and their interaction as fixed effects. It also contained items and participants as random effects, as well as random slopes for all predictors in the case of participants and syntactic function and gender sameness in the case of items. Compared to Nulls, Overt are used significantly more in cases of different gender, $b = 1.32$, $p = .02$, and to refer to the object, $b = -2.08$, $p = .005$. There is also a syntactic function*verb bias interaction, $b = 2.43$, $p = .006$: the probability of Overt is higher in the S+ICV2 and O+ICV1 combinations. For the second subset, the optimal model contained syntactic function and gender sameness (and did not contain verb bias). It also contained items and participants as random effects, as well as random slopes for all predictors. Compared to Names, Overt are used significantly more in cases of different gender, $b = 1.29$, $p = .003$, and for subject antecedents, $b = 1.34$, $p = .003$.

Discussion. Hypothesis (iii) is borne out: as opposed to both Nulls and Names, Overt are produced more when they have the potential to unambiguously select a referent, which may explain why they were produced scarcely in previous experiments. As for hypothesis (i), while we do replicate the finding that Overt are more object-biased than Nulls, we also find that they are less object-biased than Names. Finally, we find some evidence that verb bias has a mild effect in the production of Overt: compared to Nulls, the use of Overt is favoured when there is a mismatch between the verb bias and the antecedent's position (S+ICV2 and O+ICV2). In contrast, verb bias is not significant when comparing Nulls and Names. In conclusion, while we do find evidence for the role of semantic factors in the production of Overt, its role is secondary when compared to structural factors, which calls for an intermediate proposal between the Bayesian account and the Expectancy Hypothesis.

Selected references. Arnold, J. E. (2010). *How speakers refer: The role of accessibility*. *Language and Linguistics Compass*, 4(4):187203. • Carminati, M. N. (2002). *The processing of Italian subject pronouns*. PhD thesis, UMass. • Rohde, H. and Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, 29(8):912–927.

An acoustic account of Rhotic assibilation in Chihuahua Spanish

Natalia Mazzaro

University of Texas at El Paso

Rhotic assibilation is a common feature in different Spanish speaking countries such as Argentina (Colantoni, 2006), Ecuador (Bradley, 2004), México (Amastae *et al.*, 1998; Willis & Bradley, 2012), among others. These studies reported that rhotic assibilation alternates with the flap (e.g. in final position: mejo[r] vs. mejo[r̥] ‘better’) and/or with the trill (pe[r]o vs. pe[r̥]o ‘dog’). However, a more recent sociophonetic analysis (Mazzaro and González de Anda, 2020) of production data reveals that the variation in Chihuahua Spanish involves several variants and that a finer acoustic analysis will provide a better account for rhotic assibilation in this dialect of Spanish. The goal of this study is to explore the acoustic properties of rhotic variants in Chihuahua Spanish and look for a way to classify them based on these properties.

In this study, we explore two aspects of rhotic assibilation in Chihuahua Spanish, Mexico /El Paso Spanish, United States: 1) an acoustic analysis of rhotic variants including intensity, duration, center of gravity, and spectral moments, and 2) the acoustic parameters that best correlate with perception of assibilation

The participants of this study include 49 native Spanish speakers recruited in the El Paso, Texas—Ciudad Juárez, Mexico border area. Thirty-one participants were women and eighteen were men with ages that ranged between 18 and 69 years. They were recorded while reading a list of 59 isolated words (36 stimuli + 23 distracters) with the rhotic in utterance final position preceded by all vowels except /u/. The statistical analysis was conducted in R and it included a combination of linear discriminant analysis (LDA), ANOVA and Tukey’s tests.

Our results based on 1,755 tokens of /r/ in absolute final position revealed seven variants: the tap and the trill, the fully assibilated rhotic, the partially assibilated rhotic, the devoiced trill, the approximant rhotic, and the glottal fricative. The acoustic analysis showed that the best predictor for assibilation is a combination of standard deviation (measured in the third quarter of the rhotic) and the logarithm of duration (Figure 1). Log Duration also allowed us to distinguish 80% of taps from trills (Figure 2). Unfortunately, we did not find a reliable acoustic parameter to distinguish weakened variants (approximants and glottal fricative) from others. Results support the analysis of weakened variants (approximants and glottal fricative) as members of an articulatory continuum from standard variants (Colantoni 2006). They also show that assibilated rhotics form a continuum moving away from the standard rhotics in a different direction.

References

- Amastae, J., Escalera, A., Arceo, A., Gutiérrez, Z., Silva, H., & Talamantes, B. (1998). Cambio desde abajo, cambio desde arriba. *IV Encuentro Internacional de Lingüística en el Noroeste*, ed. Acosta Félix, A., Estrada Fernández, Z., Figueroa Esteva, M., López Cruz, G. 269-280. Hermosillo, Mx: UniSon.
- Boersma, P. & Weenink, D. (2020). Praat: doing phonetics by computer [Computer program]. Version 6.1.11, retrieved 6 April 2020 from <http://www.praat.org/>.

Bradley, T. G. (2004). Gestural Timing and Rhotic Variation in Spanish Codas. *Laboratory Approaches to Spanish Phonology*, ed. Timothy L. Face, 195-220. Berlin: Mouton de Gruyter.

Colantoni, L. (2006). Increasing Periodicity to Reduce Similarity: An Acoustic Account of Deassibilation in Rhotics. *Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology*, ed. Manuel Diaz-Campos, 22-34. Somerville, MA: Cascadilla Proceedings Project.

Willis, T. G., & Bradley, E. W. (2012). Rhotic variation and contrast in Veracruz Mexican Spanish. *Estudios de Fonética Experimental*, XXI, 43-74.

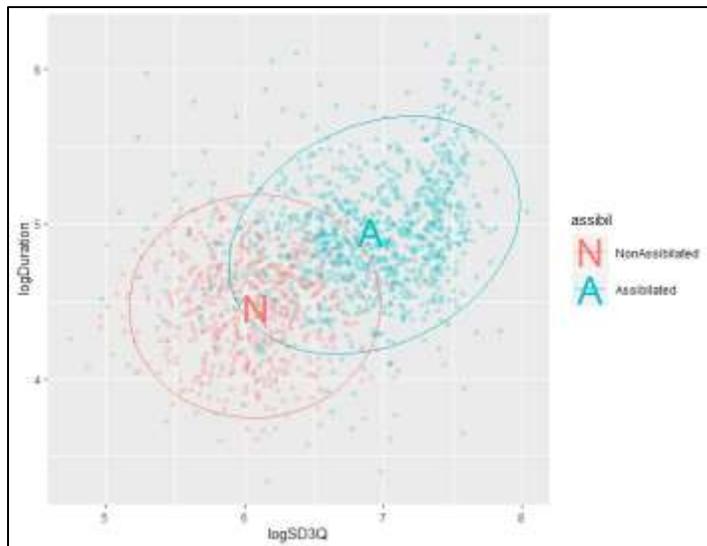


Figure 1. Standard deviation in the third quarter and log duration can classify well 86.2% of our assibilated rhotics from the non-assibilated rhotics –LDA in R.

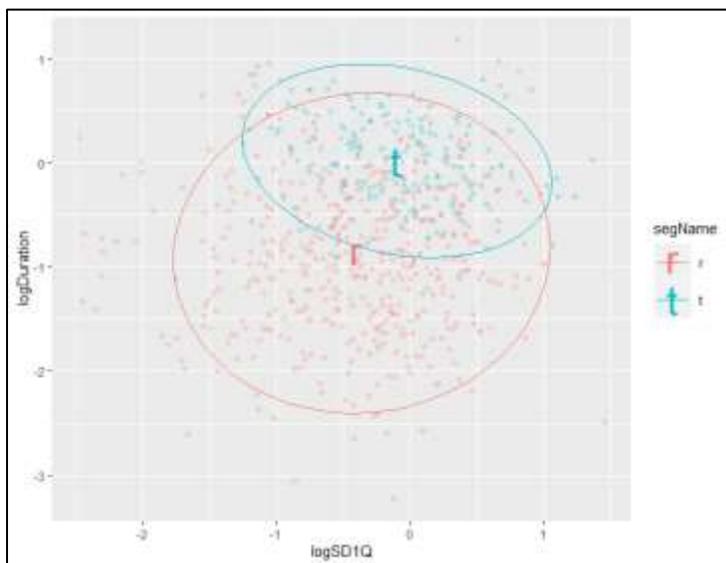


Figure 2. LogDuration can distinguish 79.8% of taps from trills. –LDA in R.

Singular, plural, modal

Puzzle. The Spanish indefinite *algún* (henceforth: ALG-SG), the German indefinite *irgendein* (henceforth: IRGEND-SG), and the English indefinite *some* (henceforth: SOME-SG) are all what is called *modal* indefinites: They all give rise to a speaker ignorance (or indifference, but we will put this aside) effect in seemingly episodic contexts. However, they vary with respect to the strength of this effect: While they are all compatible with a context where there is actually specific negative nonvariation (henceforth: 1-LOSER scenario; cf. *but not x* below), only SOME-SG is also compatible with a context where there is actually specific *positive* nonvariation (henceforth: 1-WINNER scenario). The same seems to be true of their plural variants (henceforth: __-PL): While their modal status is not often discussed, they seem to have the same kind of variation. Finally, we sometimes find differences between the SG and the PL variants of the same item: ALG-SG is not compatible with 1-WINNER, but ALG-PL is.

- | | |
|---|--|
| <p>(1) Jo vive con algún estudiante,
Jo lives with ALG-SG student-SG</p> <p>a. # en concreto, con Ada.
namely with Ada</p> <p>b. ✓ pero no con Ada.
but not with Ada</p> <p>(2) Jo wohnt mit irgendeiner Studentin,
Jo lives with IRGEND-SG student-SG</p> <p>a. # und zwar mit Ada.
namely with Ada</p> <p>b. ✓ aber nicht Ada.
but not Ada</p> <p>(3) Jo lives with some student,</p> <p>a. ✓ namely Ada.</p> <p>b. ✓ but not Ada.</p> | <p>(1') Jo vive con algunos estudiantes,
Jo lives with ALG-PL student-PL</p> <p>a. ✓ en concreto, con Ada y Bea.
namely with Ada and Bea</p> <p>b. ✓ pero no con Ada y Bea.
but not with Ada and Bea</p> <p>(2') Jo wohnt mit irgendwelchen Studenten,
Jo lives with IRGEND-PL student-PL</p> <p>a. # und zwar mit Ada und Bea.
namely with Ada and Bea</p> <p>b. ✓ aber nicht Ada und Bea.
but not Ada and Bea</p> <p>(3') Jo lives with some students,</p> <p>a. ✓ namely Ada and Bea.</p> <p>b. ✓ but not Ada and Bea.</p> |
|---|--|

Existing literature and this talk. All of these patterns are known, and there have also been many accounts to explain them. For example, Alonso-Ovalle and Menéndez-Benito (2010, 2011) discuss and account for ALG-SG/PL in full. However, the patterns have never been considered in full, which has resulted in inconsistent diagnosis and analysis. For example, it has never been noticed that SOME-SG (modal character known since Strawson 1974, endorsed in Alonso-Ovalle and Menéndez-Benito 2015; compatibility with 1-WINNER noted in Becker (1999), Marty, Picat, and Mascarenhas, p.c., ongoing experimental work) shows that weak modal indefinite patterns may also include compatibility with specific *positive* certainty, which may in turn suggest that __-PL, which for IRGEND-PL is clearly modal, might be generally modal too. As a result, there is no account that would capture them in full. In this talk I argue, based on *some-SG*, that all these items, both SG and plural, are modal. And that they raise 3 questions: How do we derive the between-item (non)variation in the SG? How do we derive the between-item (non)variation in the PL? And how do we derive the within-item (non)variation between the SG and the PL? The goal of this talk is to answer all these questions.

Proposal: The between-item variation in the SG. Building on Alonso-Ovalle and Menéndez-Benito (2010), Chierchia (2013), etc., and refs. therein, I propose the following:

★ A singular NP ranges over atoms: E.g., $\llbracket \text{student} \rrbracket = \{a, b\}$. The indefinite quantifies existentially over the domain of atoms resulting from this: $\exists x \in \{a, b\}[\dots]$.

★ Replacing the domain in the truth conditions with its subsets yields subdomain alternatives, DA : E.g., $DA = \{\exists x \in \{a\}[C(j, x)], \exists x \in \{b\}[C(j, x)]\}$. Abbreviating: $DA = \{a, b\}$. Replacing the scalar element, \exists , with its scalmate, \forall , similarly yields scalar alternatives, SA.

★ All these alternatives are factored into meaning via a silent exhaustivity operator O. O asserts the prejacent and negates the non-entailed alternatives. The DA of all our indefinites must be factored in in a pre-exhaustified form, ExhDA (obtained by applying O to individual DA; I assume pre-exhaustification of a DA is done relative to other DA of the same size).

★ $O_{\text{ExhDA+SA}}$ without an intervening operator leads to a crash, but with an intervening modal leads to a Free Choice (FC) effect. Our seemingly episodic utterances are actually prefixed with a null epistemic necessity modal \Box_S (cf. Gricean *Bel_S* ‘the speaker believes ...’), so O_{ExhDA} proceeds across this modal and yields an epistemic FC effect aka ignorance. ★ The FC effect thus obtained is total. How do we then get specific negative and positive nonvariation? I propose that this comes from DA -pruning (and only for domains larger than 2, as for smaller domains DA -pruning would destroy the domain). In particular, note that using just the singleton DA yields 1-LOSER (replicating other results from the literature) and using just the non-singletons yields 1-WINNER (a new result). The latter is illustrated below, showing only the crucial ExhDA -implicatures. If ALGUN/IRGEND-SG only allow pruning of non-singleton DA whereas SOME-SG allows pruning of either SgDA or NonSgDA, this captures the variation.

(4) $O_{\text{ExhNonSgDA+SA}} \Box_S (a \vee b \vee c)$ **just NonSgDA \Rightarrow specific positive certainty = 1-WINNER ✓**

$$= \Box_S (a \vee b \vee c) \wedge \neg \underbrace{O \Box_S (a \vee b)}_{\substack{\Box_S (a \vee b) \wedge \neg \Box_S (a \vee c) \wedge \neg \Box_S (b \vee c) \\ \Box_S (a \vee b) \rightarrow \Box_S (a \vee c) \vee \Box_S (b \vee c)}} \wedge \neg \underbrace{O \Box_S (a \vee c)}_{\substack{\Box_S (a \vee c) \wedge \neg \Box_S (a \vee b) \wedge \neg \Box_S (b \vee c) \\ \Box_S (a \vee c) \rightarrow \Box_S (a \vee b) \vee \Box_S (b \vee c)}} \wedge \neg \underbrace{O \Box_S (b \vee c)}_{\substack{\Box_S (b \vee c) \wedge \neg \Box_S (a \vee b) \wedge \neg \Box_S (a \vee c) \\ \Box_S (b \vee c) \rightarrow \Box_S (a \vee b) \vee \Box_S (a \vee c)}}$$

verified, e.g., by $\Box_S a \wedge \neg \Box_S / \Box_S \neg b \wedge \neg \Box_S / \Box_S \neg c$

Proposal: The between-item variation in the PL. A plural NP ranges over atoms and pluralities. E.g., $[\text{students}] = \{a, b, ab\}$. Everything else follows as for the SG (though to see this the domain needs to be even larger).

Proposal: The within-item variation between the SG and the PL. Note that $O_{\text{ExhNonSgDA}}$ actually verifies two 1-WINNER scenarios: Specific positive certainty about one element of the domain with *ignorance* about the rest (e.g., $\Box_S a \wedge \neg \Box_S b \wedge \neg \Box_S c$) or specific positive certainty about one element of the domain with *negative certainty* about the rest (e.g., $\Box_S a \wedge \Box_S \neg b \wedge \Box_S \neg c$). This should be the case for both SG and PL. However, I propose that the SG part of __-SG imposes a presupposition that there is a unique witness of the existential claim: $\exists!x \in D[P(x)]$. This essentially means that, for a SG epistemic indefinite, compatibility with a 1-WINNER scenario can only be of the latter kind. Thus, in SG but not PL modal indefinites, compatibility with 1-WINNER destroys the FC nature of the item. This explains why an item, such as ALG-, which disallows 1-WINNER in the SG might nevertheless allow it in the PL (and possibly also why when this effect is attested in the SG it tends to be accompanied by another mysterious modal effect—speaker *indifference*).

Conclusion and outlook. I have argued that weak modal indefinites vary with respect to the strength of the modal effect, and there is between-item variation in the singular, between-item variation in the plural, and within-item variation between the singular and the plural. I propose a fully unified account where all these patterns come from obligatory exhaustification relative to pre-exhaustified subdomain alternatives coupled with variation with respect to subdomain alternative pruning, and a uniqueness presupposition in singular indefinites.

References. Alonso-Ovalle, L. & Menéndez-Benito, P. (2010). Modal indefinites. NLS. Alonso-Ovalle, L. and Menéndez-Benito, P. (2011). Domain restrictions, modal implicatures and plurality: Spanish *algunos*. JoS. Chierchia, G. (2013). Logic in grammar.

Prosodic constraints on wh-extraction from infinitival clauses

Carlos Muñoz Pérez (Pontifical Catholic University of Chile)
Matías Verdecchia (IIF/SADAF/CONICET & University of Buenos Aires)
Fernando Carranza (CONICET & University of Buenos Aires)

General issue. This paper discusses a type of wh-extraction restriction in Spanish that has remained unnoticed in the literature. The pattern involves wh-movement from within an infinitival clause functioning as a preverbal subject. While the restriction is reminiscent of more traditional constraints on movement, e.g., Huang’s (1982) *Condition on Extraction Domain* (CED), the construction displays a series of amelioration effects that set it apart from apparently similar phenomena. In particular, we argue that the restriction has a strong correlation with prosodic structure: extraction from the infinitival subject clause S_{inf} is impossible if the sentence receives an (S_{inf} VO) phrasing. We provide a conjecture explaining this connection.

Data. The sentences in (1) and (2) seem to show that wh-extraction from a subject infinitival TP is disallowed in Spanish if the TP is preverbal; in all cases, the data belong to Rioplatense Spanish.

- (1) * ¿Qué libros creés que [TP leer ~~qué~~ libros] causa problemas?
what books believe that read that books causes trouble
- (2) ¿Qué libros creés que causa problemas [TP leer ~~qué~~ libros]?
what books think that causes problems read what books
‘Reading what books do you think is troubling?’

However, such a configuration becomes transparent for extraction under a series of apparently heterogeneous conditions. To begin with, the acceptability of sentences like (1) improves significantly if additional material, e.g., an adjunct PP, appears together with the infinitive.

- (3) ¿Qué libros creés que [TP leer ~~qué~~ libros en el secundario] causa problemas?
what books think that read what books in the secondary causes trouble
‘What books do you think it is troubling to read in high school?’

Second, if the infinitive in (1) is interpreted as a contrastive focus and receives the corresponding intonation, the sentence also becomes rather acceptable.

- (4) A: Creo que comprar esos libros causará problemas.
‘I believe that buying those books will cause trouble.’
B: Bueno, ¿pero qué libros creés que [TP LEER ~~qué~~ libros] causará problemas?
well but what books think that read what books will.cause trouble
‘Well, but what books do you think that READING will cause trouble?’

A third and final amelioration factor is the presence of a parenthetical adjunct between the subject infinitival phrase and its verb.

- (5) ¿Qué libros creés que [TP leer ~~qué~~ libros], de acuerdo con tu experiencia, causa problemas?
what books think that read what books of according with your experience causes trouble
trouble
‘What books do you think that reading, according to your own experience, causes trouble?’

Discarding structural explanations. The pair in (1) and (2) resembles a traditional subject island effect already attested in Spanish: as observed by Starke (2001: 57), Spanish preverbal subjects are opaque for extraction, while postverbal subjects are rather transparent.

- (6) ??* ¿De qué autor crees que [DP varios libros] han recibido premios internacionales?
of which author think that several books have received awards international
- (7) ? ¿De qué autor crees que han recibido premios internacionales [DP varios libros]?
of which author think that have received awards international several books
‘By which author do you think several books have received international awards?’

According to Haegeman et al. (2014), the problem with (6) is that the subject occupies a derived position, and thus violates the *Freezing Principle* (which says that a moved constituent is opaque for extraction). However, this type of analysis does not apply to (1), as it does not capture the ameliorating effects attested in (3) to (5), e.g., there is no obvious reason why adding a PP within the subject as in (3) would repair a movement-based violation. The same holds for other structural

constraints on subject extraction, e.g., Chomsky’s (1973) *Subject Condition*, or the many incarnations of the CED (Huang 1982): e.g., there seems to be no relevant structural distinction between (1) and (3), so a purely configurational restriction fails at predicting the acceptability of the latter.

Moreover, subject island violations like (6) are insensitive to the repairing strategies in (3) to (5), e.g., contrastive focus on the subject does not improve acceptability (8); this shows that the active constraints in (1) and (6) are of distinct nature.

- (8) *¿De qué autor crees que [DP varios LIBROS] recibieron premios internacionales?
of which author think that several books received awards international
‘By which author do you think several BOOKS received international awards?’

Amelioration effects are prosodic. We argue that the restriction in (1) and the ameliorating effects in (3), (4) and (5) are related to the prosodic phrasing of the subject infinitival clause. We propose the following descriptive generalization.

- (9) Wh-extraction from a subject infinitival clause S_{inf} in preverbal position is impossible if S_{inf} is mapped into a prosodic word ω that is immediately dominated by the phonological phrase φ containing the VP, i.e., no wh-extraction from S_{inf} if ($S_{inf}VO$).

Thus, the problem with (1) is that the infinitive *leer* ‘to read’ is mapped into a prosodic word ω that forms a phonological phrase with the rest of the clause; more on this below.

- (10) ... (leer $_{\omega}$ causa $_{\omega}$ problemas $_{\omega}$) $_{\varphi}$ cf. (1)

Syntactic branchingness is one of the key factors intervening in phonological phrase formation in Spanish (D’Imperio et al. 2005). Since the infinitival clause in (3) contains two prosodic words ω , it is mapped into a phonological phrase φ independent from that containing the finite verb.

- (11) ... (leer $_{\omega}$ en-el-secundario $_{\omega}$) $_{\varphi}$ (causa $_{\omega}$ problemas $_{\omega}$) $_{\varphi}$ cf. (3)

Contrastive focalization and insertion of parenthetical adjuncts also allow to separate the prosodic domain of the subject infinitival clause from that of the predicate. In (4), we take that contrastive focus introduces a boundary tone. Thus, the focused infinitive is mapped into a preceding phonological phrase (12). In (5), a parenthetical constituent intervenes between the infinitival subject and the VP; crucially, this type of adjunct introduces its own prosodic domain (Truckenbrodt 2015), and therefore separates the infinitive from the embedded verb.

- (12) ... LEER) $_{\varphi}$ (causa $_{\omega}$ problemas $_{\omega}$) $_{\varphi}$ cf. (4)

- (13) ... leer) $_{\varphi}$ (de acuerdo con tu experiencia) $_{\varphi}$ (causa $_{\omega}$ problemas $_{\omega}$) $_{\varphi}$ cf. (5)

The conjecture. It remains to explain why (10) is problematic for extraction; notice that there is nothing wrong *per se* with such a phrasing, as it is independently attested in declarative sentences. We speculate that (10) creates a mismatch at the syntax–prosody interface. At the syntactic level, the extracted wh-phrase requires to be linked to a trace in the complement position of the infinitive. For this, the infinitive must head a verbal projection. If syntactic phrases map into phonological phrases (Selkirk 2011), the prosodic parsings in (11), (12) and (13) successfully express the fact that the infinitive heads a phrase of its own, from which a constituent could have been extracted, i.e., the phonological phrases headed by the infinitive can be taken to host wh-traces. On the other hand, the phrasing in (10) treats the infinitival clause as an element with no internal structure, from where movement cannot have taken place. In a nutshell, if the infinitival clause is assigned the prosody of a word, this may disrupt syntactic dependencies that require it to be a phrase.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart & Winston.

D’Imperio, Mariapaola, Gorka Elordieta, Sónia Frota, Pilar Prieto & Marina Vigário. 2005. Intonational phrasing in Romance: The role of syntactic and prosodic structure. In Sónia Frota, Marina Vigário & Maria João Freitas (eds.), *Prosodies. With special reference to Iberian languages*, 59–97. Berlin: Mouton de Gruyter.

Haegeman, Liliane, Ángel L. Jiménez-Fernández & Andrew Radford. 2014. Deconstructing the subject condition in terms of cumulative constraint violation. *The Linguistic Review* 31(1). 73–150. doi:10.1515/tlr-2013-0022.

Huang, C. T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: MIT dissertation.

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In John Goldsmith, Jason Riggle & Alan Yu (eds.), *The handbook of phonological theory*, vol. 2, 435–483. Oxford: Wiley-Blackwell Malden.

Starke, Michal. 2001. *Move dissolves into merge: a theory of locality*. Geneva: University of Geneva dissertation.

Truckenbrodt, Hubert. 2015. Intonation phrases and speech acts. In Marlies Kluck, Dennis Ott & Mark de Vries (eds.), *Parenthesis and ellipsis*, 301–349. Berlin: Mouton de Gruyter.

The phonetic reduction of Spanish stressed vowels in spontaneous speech

Scott James Perry, Matthew C. Kelley, Benjamin V. Tucker
University of Alberta

The majority of research investigating the acoustic properties of Spanish vowels has focused on the difference between vowels in stressed and unstressed syllables. While Spanish does not have metrically-induced vowel reduction, there is evidence that unstressed vowels are shorter and more centralized than their stressed counterparts (e.g., Santiago and Mairano, 2018), although this effect may differ across the vowel inventory (Nadeu, 2014). Thus, while we have ample evidence concerning how stress influences Spanish vowel production, there is a dearth of research investigating what predicts variation in stressed vowels. The present study addresses this gap by investigating how lexical and probabilistic factors that influence speech production in other languages affect the acoustic properties of Spanish stressed vowels.

A growing body of research indicates that factors such as word frequency and predictability in context can partially predict sub-phonemic variation in speech production. Within the temporal domain, which has received the most attention, it has been found that more frequent and more predictable parts of speech are shorter, all else being equal (e.g., Cohen Priva and Gleason, 2020; Aylett and Turk, 2004; Bell et al., 2009; Tremblay and Tucker, 2011). Regarding vowel quality, it has been found that more frequent (Wright, 2004) and more predictable (Aylett and Turk, 2006) words have more centralized vowels. Many studies have also investigated the effect of grammatical class, and have found that function words are shorter and more likely to contain reduced vowels (e.g., Bell et al., 2009; Jurafsky et al., 2001). Thus, based on previous research, we predicted that stressed vowels in Spanish would be shorter and more centralized in more frequent and predictable words, as well as being shorter and more centralized in function words.

To investigate how word frequency, word predictability, and grammatical class affect stressed vowel production in Spanish, we analyzed recordings from the Nijmegen Corpus of Casual Spanish (Torreira and Ernestus, 2010). This corpus comprises recordings of spontaneous conversations between groups of friends from Madrid, Spain, with a total of 60 recordings of approximately 90 minutes each. Professional orthographic transcriptions accompany all conversations. We aligned the recordings at the word and phone levels using the Montreal Forced Aligner (McAuliffe et al., 2017). The duration and first two formant values were then extracted from all stressed monophthongal vowels using a script in Praat, with the formant ceilings being set manually for each recording at a value where the formant tracker aligned with visible formants. As both the forced alignment and formant tracking were sources of error for formant values, we flagged and removed outliers in the multivariate formant space using a robust method of calculating Mahalanobis distance: the *minimum covariance determinant*. This was calculated using the `rrcov` package (Todorov, 2020) for each speaker and vowel individually. After removing outliers in terms of vowel quality, we removed temporal outliers that were more than 1.5 times outside the interquartile range for each vowel. The remaining 206,615 tokens constituted $\approx 79\%$ of the original data. To account for differences in formant values stemming from vocal tract length, we normalized formant values using the `neareyI` method as implemented in `phonTools` (Barreda, 2015).

To analyze vowel production, we fit two sets of generalized additive mixed models using the `mgcv` package (Wood, 2020). The first set of models had logged vowel duration as the dependent variable, and the second set had the euclidean distance from the normalized vowel

space centre. The independent variables of interest in both models were grammatical class, word frequency, and the word’s conditional probability of occurrence based on the previous word, with frequency and conditional probability being calculated from the `OpenSubtitles` Spanish corpus (Tiedemann, 2012). Grammatical class was binary, either content or function word, and was based on Hallebeek (1986). We fit separate smooths for the effects of word frequency and predictability for the two levels of grammatical class. The word’s length in syllables and the local speech rate were entered into all models as control variables, with local speech rate being calculated automatically for each utterance using a modified version of the Praat script developed by de Jong and Wempe (2009). Random effects were included that accounted for pseudo-replication and phonetic environment.

Results indicate that frequency significantly affects stressed vowel production in Spanish, with more frequent words containing shorter and more centralized vowels. Word predictability influenced the temporal aspects of function words and the vowel quality of content words. There was no main effect of grammatical class on vowel duration ($p=0.844$), but there was a main effect of grammatical class on centralization ($p=0.006$). Plots for the smooth terms for frequency and predictability by grammatical class from the final models appear in Figure 1.

These results constitute evidence that the production of stressed vowels in Spanish varies as a function of word frequency. The effect of predictability was less consistent, and had a smaller effect size than did frequency. Grammatical class did not affect vowel duration, which is in contrast with findings from other languages. While this discrepancy may point to a cross-linguistic difference, another potential explanation is that previous studies failed to control for phrasal prominence. Many function words in Spanish are destressed in running speech, and as such were excluded from the present analysis, which concerned stressed vowels. These results will be discussed in the context of hypotheses of phonetic reduction.

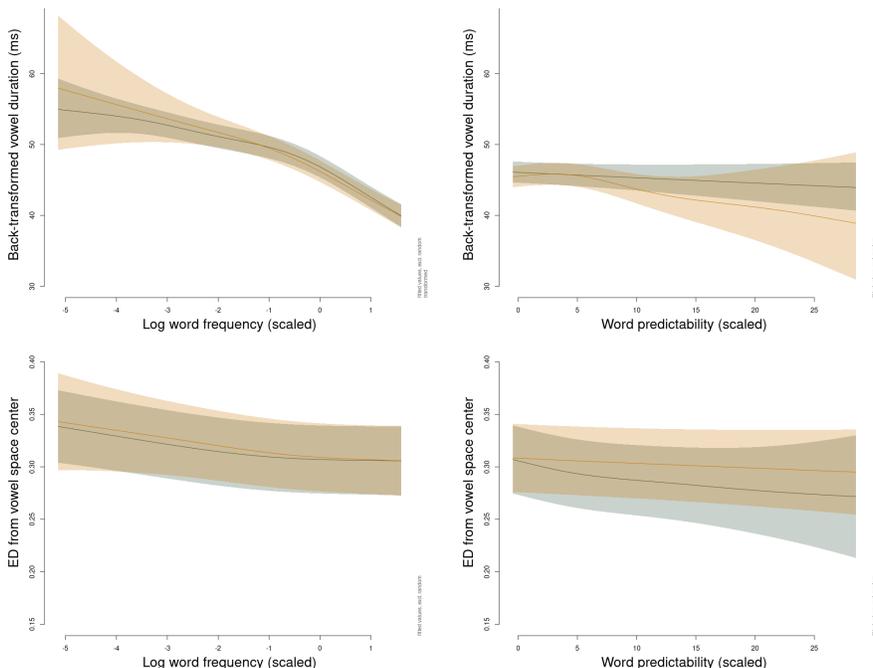
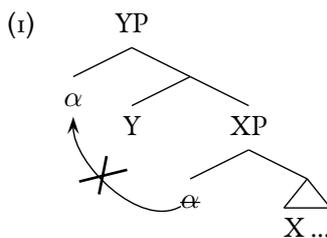


Figure 1: Predicted values from final GAMM models. For all plots, smooths for content words are plotted in green and smooths for function words are plotted in gold.

A unified analysis of failed operator movement in Spanish in terms of anti-locality

Erika Petersen (Stanford University)

1. INTRODUCTION. There are in Spanish a series of configurations which disallow \bar{A} -movement across them: (i) P + CP clauses; (ii) D + CP clauses; (iii) clauses with two complementizers; and (iv) finite adjunct clauses with parasitic gaps. Separate, often construction-specific explanations have been provided heretofore for this pattern. In this talk, I provide a unified analysis for the islandhood of these configurations in terms of the interaction between an anti-locality constraint on movement, which bans movement that is too short (Abels 2003, Bošković 2015, Grohmann 2003, Erlewine 2020, a.o.), and successive cyclicity, which forces a moving XP to stop at the specifier position of each intermediate phase head on its way to its final landing site. For concreteness, I adopt Erlewine’s (2020) Spec-to-Spec anti-locality constraint, which states that “[m]ovement of a phrase from the specifier of XP must cross a maximal projection other than XP” (1).



2. P + CP CLAUSES. CP complements of adjectives and some verbs are introduced by a P. *Wh*-extraction out of these complements is impossible (2). Spec-to-Spec anti-locality accounts for the impossibility of *wh*-movement in these configurations, on the assumption that P and C are phasal heads (Bošković 2015, Martínez Vera 2020).

- (2) *¿Qué_i se olvidó Carmen [PP **de** [CP **que** la chica aprobó t_i]]?
 what REFL.3SG forgot Carmen of that the girl passed t
 ‘What did Carmen forget that the girl passed?’

The analysis correctly predicts that when either the C or the P is absent, extraction becomes possible. First, in line with the Minimal Structure Principle (Bošković 1996, a.o.), infinitival clauses can be considered to be TPs rather than CPs. They are therefore not phases. *Wh*-extraction out of a P + TP_{-FIN} complement should then be possible – a prediction which is borne out (3). Second, for some speakers, Vs and As which take P + CP complements may also select directly for a CP complement. *Wh*-extraction in this latter case should be possible, and, in fact, it is (4) (Bošković 2015).

- (3) ¿Qué_i se olvidó Carmen [PP **de** [TP_{-FIN} haber aprobado t_i]]?
 what REFL.3SG forgot Carmen of have.INF passed t
 ‘What did Carmen forget passing?’
- (4) ¿Qué_i estás segura ([PP ***de**] [CP **que** la chica aprobó t_i]())?
 what are sure of that the girl passed t
 ‘What are you sure that the girl passed?’

3. D + CP CLAUSES. Some *que*-CP complements to Vs may optionally be introduced by the D *el* ‘the’. While *wh*-extraction is possible when D is absent, it is impossible when D is present (5) (Leonetti 1999: 826). Assuming that D and C are phases, the impossibility of *wh*-movement across *el* + *que* complements follows from Spec-to-Spec anti-locality.

- (5) ¿Qué_i lamentó Ana ([DP ***el**] [CP **que** Carmen reprobara t_i]())?
 what regretted Ana the that Carmen failed t
 ‘What did Ana regret that Carmen failed?’

4. CLAUSES WITH TWO COMPLEMENTIZERS. In Spanish, *wh*-extraction is possible out of embedded declarative CPs headed by *que*. Certain verbs accept clauses with two complementizers, when a ‘left-dislocated’ XP is between the two *que*-Cs: [CP *que* [CP XP [CP *que* ...]]] (Fontana 1994, Villa-García 2012). In this case, *wh*-extraction from the lower clause is impossible (6). On the standard assumption that C is a phase head (Chomsky

status of infinitival complements', *NLLT*. • Bošković 2015 'From the Complex NP constraint to everything', *TLR*. • Chomsky 1986 *Barriers*. • Chomsky 2000 'MI'. • Chomsky 2001 'DbP'. • Chomsky 2008 'OP'. • Erlewine 2020 'Anti-locality and subject extraction', *Glossa*. • Fontana 1994 'A residual A-bar position in Spanish', *Proc. 12th WCCFL*. • García Mayo & Kempchinsky 1994 'Finiteness in Romance vs. English *pg* constructions', *IIRL*. • García Mayo 1993 'A new look at *pgs*', *LPRL*. • Grohmann 2003 *Prolific Domains*. • Leonetti 1999 'La subordinación sustantiva', *GDLE*. • Martínez Vera 2020 'On Phase-over-Phase configurations and phase collapsing', *TLR*. • Nissembaum 2000 *Investigations of Covert Phrase Movement*, MIT diss. • Villa-García 2012 'Recomplementation and locality of movement in Spanish', *Probus*.

Acquiring pragmatic prosody: Evidence from 3- to 4-year-old Catalan speaking children.

Mariia Pronina¹, Iris Hübscher², Ingrid Vilà-Giménez^{1,3} and Pilar Prieto^{4,1}

¹ Dept. of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Catalonia, ² URPP Language and Space, University of Zurich, ³ Dept. of Subject-Specific Education, Universitat de Girona, Girona, Catalonia, ⁴ Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia

INTRODUCTION: Even though it is well-known that prosodic features are central in the conveyance of pragmatic meaning across languages, most developmental research has separately assessed the development of prosodic and pragmatic abilities, focusing either on children's pragmatic communicative abilities or on the understanding and use of prosodic cues, mainly intonation. From the point of view of clinical evaluation, likewise, assessment tools have separately assessed pragmatic and prosodic abilities. Moreover, the main focus of the past research was on prosodic development of early infant abilities (see Chen et al., 2020 for a review), while less is known about the acquisition of prosodic patterns for the expression of pragmatic meanings (what we will call from now on "pragmatic prosody") later in development. This study paper aims to fill this gap and contribute to our knowledge of prosodic development patterns by investigating pragmatic prosody patterns in a comprehensive way in young Catalan speaking preschool children.

METHOD: A total of 42 3- to 4-year-old typically developing Catalan-speaking children participated in the study. Children's pragmatic prosody skills were assessed by means of the Audiovisual Pragmatic Test (APT; Pronina, Hübscher, Vilà-Giménez, & Prieto, 2019). This test assesses prosody in relation to social contexts by using a comprehensive pragmatic coverage, as well as a picture-supported set of the Discourse Completion Task in which the participant is asked to imagine an everyday social context and then to respond to it as naturally as possible. All items intended to elicit a pragmatically appropriate phrase/set of phrases which correspond to a certain speech act. We distinguish between 4 types of speech acts, specifically, *assertions*; *requests*; *basic expressives* such as greeting, calling or thanking; and *complex expressives* that evolve around complex social situations like expressing empathy, compassion, condole or congratulations. Requests and assertions can be either *unbiased* or *biased*. Unbiased (i.e., unmarked) requests and assertions have no additional pragmatic meanings (e.g., an example of *unbiased request* is a command; an example of *unbiased assertion* is an unmarked declarative statement). Biased (i.e., marked) requests and assertions convey additional pragmatic biases such as different types of epistemic meanings (e.g., a *biased request* expressing incredulity or a *biased assertion* expressing obviousness or uncertainty), marked informational structure (e.g., a *biased assertion* conveying contrastive focus), or negation. The children were tested individually; the prosodic component of the answer was evaluated perceptually in terms of the appropriateness of the prosody and was given a score from 0 to 2.

RESULTS: We examine the overall distribution of appropriate pragmatic and prosodic responses and report them in terms of percentage distributions across pragmatic types. More specifically, following the APT scoring system, we distinguish between three types of responses: (1) pragmatically appropriate answers produced as indirect speech (i.e., not enacted through prosody), (2) pragmatically infelicitous answers enacted through prosody but not prosodically appropriate (e.g., non-expressive or produced with non-adequate prosody), and (3) pragmatically and prosodically felicitous answers. Results showed that while the majority of children manage to successfully produce basic speech acts such as *unbiased assertions*, *basic expressives*, and *unbiased requests*, *biased speech acts* (biased assertions and requests) and *complex expressive* result more difficult (see Figure 1). Overall, 37% of prosodically felicitous answers were given for unbiased assertions, 34% for basic expressives, 31% unbiased requests, 16% for biased assertions, 16% for complex expressives, and 14% for biased requests, demonstrating that biased pragmatic meanings are more challenging for young preschoolers. Interestingly the graph also reveals a close relation between pragmatic and prosodic scores, that is, the more pragmatically appropriate answers are obtained for a pragmatic area, the more prosodically felicitous answers there are.

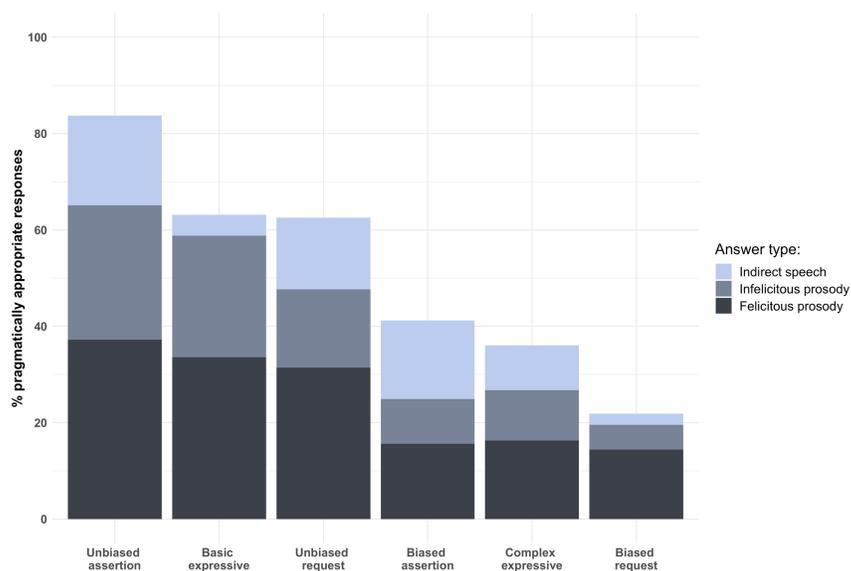


Figure 1. Percentage of pragmatically and prosodically felicitous responses by children, separated by speech acts.

DISCUSSION: Our findings allow us to sketch out the prosodic profiles of preschoolers, contributing to the integration of prosody and pragmatics and their joint consideration by investigating pragmatic prosody in Catalan-speaking preschool children. Results showed that Catalan-speaking preschoolers deal well with unbiased speech acts, expressing basic pragmatic meanings, such as basic expressive (e.g., greetings, calling), unbiased requests (e.g., commands) and unbiased assertions (e.g., unbiased declarative statements with no biased meanings). These results corroborate the findings of work on intonational development, confirming that by 3 years of age Catalan children use pragmatically appropriate prosody for basic speech acts, specifically, for basic expressives, as well as unbiased requests and assertions (e.g., see Frota et al., 2016 for Portuguese; and Prieto et al., 2012 for Catalan and Spanish). However, young preschoolers tend to have more trouble producing prosodic cues related to the expression of pragmatic biases such as information structure (corrective/contrastive focus), belief states (e.g., incredulity, uncertainty, obviousness), and negation. Preschoolers also show difficulties in expressing felicitous responses in complex social scenarios (e.g., praising comments, congratulations, concerns). These results indicate the developmental path of the ability to prosodically express pragmatic biases takes time, in line with previous studies that separately explored the acquisition of specific prosodic skills such as for example the expression of prosodic focus (e.g., see Armstrong & Hübscher, 2018; Chen, 2018). Overall, the comprehensive assessment of the acquisition of pragmatic prosody by young preschoolers demonstrates the relevance of bridging the gap between prosody and pragmatics when accounting for prosodic developmental profiles.

REFERENCES

- Armstrong, M., & Hübscher, I. (2018). Children's development of internal state prosody. In P. Prieto & N. Esteve-Gibert (Eds.), *The Development of Prosody in First Language Acquisition* (pp. 271–293). John Benjamins.
- Chen, A. (2018). Get the focus right across languages: acquisition of prosodic focus-marking in production. In P. Prieto & N. Esteve-Gibert (Eds.), *The Development of Prosody in First Language Acquisition* (pp. 295–316). John Benjamins.
- Chen, A., Esteve-Gibert, N., Prieto, P., & Redford, M. A. (2020). Development of phrasal prosody from infancy to late childhood. In C. Gussenhoven & A. Chen (Eds.), *The Oxford Handbook of Language Prosody*. Oxford University Press.
- Frota, S., Matos, N., Cruz, M., & Vigário, M. (2016). Early prosodic development: Emerging intonation and phrasing in European Portuguese. In M. Armstrong, N. C. Henriksen, & M. Vanrell (Eds.), *Intonational grammar in Ibero-Romance: Approaches acrosslinguistic subfields* (pp. 277–294). John Benjamins.
- Prieto, P., Estrella, A., Thorson, J., & Vanrell, M. D. M. (2012). Is prosodic development correlated with grammatical and lexical development? Evidence from emerging intonation in Catalan and Spanish. *Journal of Child Language*, 39(1), 1–37. <https://doi.org/10.1017/S030500091100002X>
- Pronina, M., Hübscher, I., Vilà-Giménez, I., & Prieto, P. (2019). A new tool to assess pragmatic prosody in children: Evidence from 3- to 4-year-olds. *Proceedings of the 19th International Congress of Phonetic Sciences. 5-9 August 2019, Melbourne, Australia*.

Strengthening and weakening at domain edges: evidence from Spanish

Michael Ramsammy & Matthew King
University of Edinburgh

Overview. In this paper, we report the results of experimental work that examines articulatory strengthening and weakening at the edges of prosodic domains in Spanish. /l/-articulations in a range of contexts were studied using Ultrasound Tongue Imaging (UTI). Results from a dialectal cross-section of Spanish speakers reveal both general and speaker-specific patterns. We focus firstly on the relationship between dorsum retraction, front-body advancement and utterance-edge lengthening. Secondly, we address the question of to what extent utterance-medial reduction in intervocalic /l/ is predictable on the basis of the location of prosodic boundaries.

Background. Previous research on a number of languages has shown that segmental lengthening and shortening—and concomitant articulatory strengthening and weakening—are dependent on prosodic boundaries (Byrd et al. 2005; Keating et al. 2003; Tabain 2003). However, research on these phenomena in Spanish is limited: our study therefore set out to examine to what extent articulatory domain-edge effects are observable cross-dialectally in Spanish. Given its relatively unrestricted distribution in the language, and relatively limited cross-dialectal variability (i.e. in comparison to other consonants like /s/ or /n/), we focus on realisations of /l/ in this paper.

Experimentation. Simultaneous audio and UTI data were collected from 16 speakers of Spanish (EchoBlaster 128 machine, AAA software package). The speaker sample included speakers of European, Chilean, Colombian, Mexican and Peruvian varieties. Participants read meaningful sentences containing /l/ in the environments listed in Table 1 below.

Table 1: Test contexts

(a) [v] _ω l...	<i>Utterance-initial</i>	(b) ...l _ω]v]	<i>Utterance-final</i>
(c) ...[_ω ...l...]	<i>Utterance-medial, word-medial</i>	(d) ... _ω][_ω l...	<i>Utterance-medial, word-initial</i>
(e) ...l _ω][_ω ...	<i>Utterance-medial, word-final</i>		

Contexts (a) and (b) test the realisation of /l/ at the utterance periphery. This follows from the hypothesis that domain-edge lengthening and articulatory strengthening effects, if they occur, would be most consistently observable at the edges of the maximal utterance domain (Byrd et al. 2005; Keating et al. 2003). (c–e) test /l/ in utterance-medial contexts in which the locations of word and syllable boundaries were varied. This permitted word-medial, syllable-initial realisations of /l/ to be compared to word-initial and word-final realisations, respectively.

Analysis. Acoustic durations of /l/-tokens were calculated in Praat by script. Ultrasound splines were semi-automatically fitted to visible tongue contours in each ultrasound frame over the full acoustic duration of each /l/-realisation in AAA. Polar coordinate data (Mielke 2015) for a maximum of 42 fanpoints were then extracted from these ultrasound frames and subjected to a Principal Components Analysis. This permitted tongue-configuration and /l/-duration data to be compared across contexts. Additionally, coordinates from single ultrasound frames corresponding most closely to the temporal midpoint of each /l/ were compared using visualisations of fitted splines generated using *loess* smoothing (Turton 2017).

Results & Discussion. We observe that /l/-realisations across contexts span a large durational range. In agreement with existing research on other languages, /l/-articulations are longest in utterance-peripheral contexts. Lengthening also occurs in the word-initial utterance-medial

environment; however, this is less extreme than at the utterance periphery. The shortest realisations of /l/ occur utterance-medially in word-medial and word-final contexts (i.e. (c) and (e) in Table 1).

In the UTI data, two main patterns emerge in utterance-peripheral contexts. One subset of speakers articulate utterance-final /l/ with a significant degree of tongue-dorsum retraction (e.g. (a) in Figure 1 below). A second set of speakers, such as (b) below, do not display this pattern, but instead show a tendency to produce utterance-initial /l/ with a degree of tongue-tip advancement and tongue-body lowering. Although these strategies are clearly different, we argue that both occur as the result of prosodically driven articulatory enhancement. Individual speakers, however, display variation (that may be linked to dialectal factors) in terms of whether utterance-initial or utterance-final articulations are most saliently distinct from other, non-peripheral /l/-articulations.

Similarly, a high degree of speaker-specific variation is observed in utterance-medial /l/. Some speakers (e.g. (d) in Figure 1) show minimal variation in tongue position across test contexts. This variation is more extreme for others: e.g. (c) in Figure 1, for whom the articulatory profile of word-final /l/ is significantly distinct from other utterance-medial /l/-realisations. We discuss these findings in the context of theoretical claims about the prosodification of consonants in phrasal environments in Spanish, such as the operation of coda-capture processes driven by resyllabification across word boundaries.

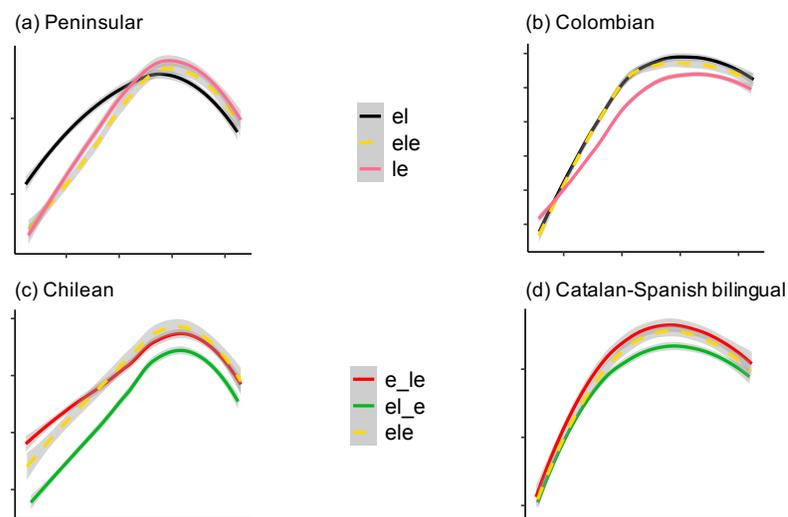


Figure 1: /l/-articulations across test contexts.

References

- Byrd, D., S. Lee, D. Riggs & J. Adams. 2005. Interacting effects of syllable and phrase position on consonant articulation. *Journal of the Acoustical Society of America* 118: 3860–3873.
- Keating, P., T. Cho, C. Fougeron & C. S. Hsu. 2003. Domain-initial strengthening in four languages. In J. Local, R. Ogden & R. Temple (eds.), *Papers in Laboratory Phonology 6*, pp. 143–161. Cambridge: Cambridge University Press.
- Mielke, J. 2015. An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons. *Journal of the Acoustical Society of America* 137: 2858–2869.
- Tabain, M. 2003. Effects of prosodic boundary on /aC/ sequences: articulatory results. *Journal of the Acoustical Society of America* 113: 2834–2849.
- Turton, D. 2017. Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English. *Laboratory Phonology* 8: 1–31.

Do Spanish codas completely resyllabify? A look at Mexican Spanish

Gemma Repiso-Puigdelliura

University of California, Los Angeles

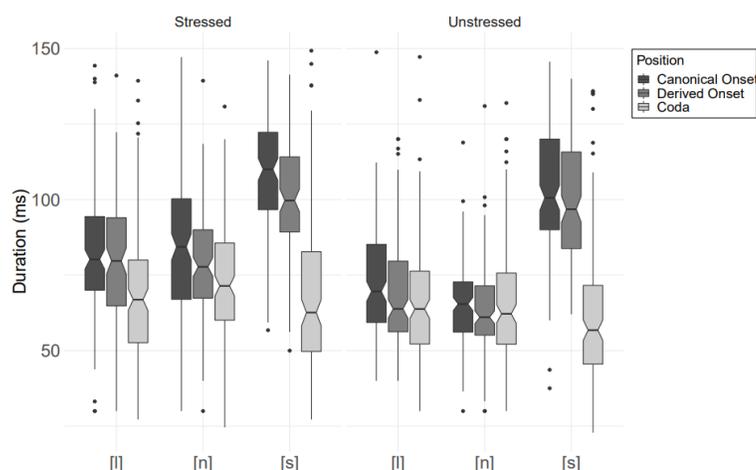
There exists a consensus in the Spanish phonology literature that coda consonants preceding vowels (e.g., /Vc#V/) completely resyllabify into the subsequent syllable and become onsets (i.e., *los osos* ‘the bears’ [los#o.sos] - [lo.so.sos]) (Colina, 1997; Harris, 1983, a.o.). From this view, it follows that consonants in derived onset position (i.e., /Vc#V/) will phonetically pattern with consonants in canonical onset position (i.e., /V#cV/) and diverge from consonants in the position of coda (i.e., /Vc#C/). However, recent phonetic analyses on the durational properties of /s/ have found that the consonant /s/ in derived onset position is longer than /s/ in canonical onset position (Hualde & Prieto, 2014; Strycharczuk & Kohlberger, 2016) and shorter than /s/ in the position of coda (Strycharczuk & Kohlberger, 2016). These findings suggest that resyllabification in Peninsular Spanish may not be complete. This study seeks to contribute to the current debate on resyllabification and expand the body of work on the phonetics of derived onsets with an analysis of the durational properties of /s/, /n/ and /l/ in Mexican Spanish.

Sixteen Mexican Spanish speakers (11 F, 5 M, Mean age = 20.5 years; SD = 1.93 years) participated in an online reading-aloud task eliciting canonical onsets (e.g., *mi silla* ‘my chair’), derived onsets (e.g., *mi_s higos* ‘my plums’), and canonical codas (e.g., *mi_s timones* ‘my rudders’) for /s/, /n/ and /l/ followed by an initial stressed syllable (e.g., *mi_s hilos* ‘my threads’) and an initial unstressed syllable (e.g., *mi_s inicios* ‘my beginnings’). Participants recorded themselves using an external microphone on Audacity (R) (AudacityTeam, 2020) (sampling rate 44.1 kHz, sample size of 16 bits). The speech data was forced aligned and manually corrected. The duration of the segments /s/, /n/, and /l/ was extracted. The resulting data (N = 2545) was fitted in three linear mixed effects models (i.e., one per consonant). Position (i.e., derived onset, canonical onset, and canonical coda) and stress (i.e., stressed and unstressed) were entered as fixed effects. Item and participant were included as random effects.

The model for /l/ shows that /l/ is longer in derived and canonical onset positions than in coda position ($p = 0.013$, $p < 0.001$ respectively). It also shows that /l/ is longer in the stressed condition than in the unstressed position ($p = 0.002$). The model for /n/ shows that /n/ is longer in the canonical onset position than in the coda position ($p = 0.018$), and that /n/ is longer in the stressed condition than in the unstressed position ($p < 0.001$). The model for /n/ also demonstrates that stress lengthens /n/ duration more when /n/ appears in canonical and derived onset positions than when it appears in the position of coda ($p = 0.03$, $p < 0.001$ respectively). The model for /s/ shows that /s/ is longer in canonical and derived onset positions than in the coda position ($p < 0.001$, $p < 0.001$ respectively), and that /s/ in the stressed condition is longer in the stressed condition than in the unstressed position ($p < 0.001$).

Our findings show different phonetic behaviours among the target consonants (See Figure 1). For the consonants /s/ and /l/, we found that derived onsets pattern canonical onsets and that they both diverge from codas, supporting a phonological account of complete resyllabification in Mexican Spanish. For the consonant /n/, we found that derived onsets pattern both with canonical onsets and codas. However, the interaction between stress and position shows that /n/ shares some stress properties with the following syllable, indicating that this consonant is at least partially resyllabified into the following syllable.

Figure 1. Consonant duration (ms) for consonant, stress and position.



References

- AudacityTeam. (2020). *Audacity(R): Free Audio Editor and Recorder* (2.4.2).
- Colina, S. (1997). Identity constraints and Spanish resyllabification. *Lingua*, 103(1), 1–23. [https://doi.org/10.1016/S0024-3841\(97\)00011-9](https://doi.org/10.1016/S0024-3841(97)00011-9)
- Harris, J. (1983). Syllable structure and stress in Spanish. a non-linear analysis. In *Linguistic Inquiry Monographs* (Vol. 8). MIT Press.
- Hualde, J. I. (1991). On Spanish syllabification. In Héctor Campos and Fernando Martínez-Gil (Ed.), *Current Studies in Spanish Linguistics* (Issue 2, pp. 475–493). Georgetown University Press.
- Hualde, J. I., & Prieto, P. (2014). Lenition of intervocalic alveolar fricatives in Catalan and Spanish. *Phonetica*, 71(2), 109–127. <https://doi.org/10.1159/000368197>
- Strycharczuk, P., & Kohlberger, M. (2016). Resyllabification Reconsidered: On the Durational Properties of Word-Final /s/ in Spanish. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1), 1–24. <https://doi.org/10.5334/labphon.5>

The aspectual properties of <no + event deverbal nominal> in Spanish

Laura Ros García
Complutense University of Madrid

This paper deals with the interaction of negation and event deverbal nominals in Spanish, and, in particular, with the aspectual properties of the constructions illustrated in (1). I show that these constructions share properties both with events and states. Thus, it is not possible to argue that we get a state, contra Horn (1989) or De Swart & Molendijk (1999), among others. I argue that the mixed aspectual properties of the construction <no + event deverbal nominal> follow from the fact that *no* occupies a Neg(ative)P that is merged with Init(iator)P and refutes the causative relation between InitP and Proc(ess)P.

- (1) La no verificación de los datos por parte de los científicos
the no verification of the data by part of the scientists
'The non-verification of the data by the scientists.'

As said before, <no + event deverbal nominal> shares properties both with events and states. Among the eventive properties we find the compatibility with *tener lugar* ('to take place') (2a), with the anaphoric construction *esto sucedió* ('this happened') (2b) and with perception verbs (2c).

- (2) a. La no verificación de los datos tuvo lugar ayer.
the no verification of the data took place yesterday
'The non-verification of the data took place yesterday.'
- b. La no verificación de los datos alteró a los científicos. Esto sucedió cuando...
the no verification of the data agitated DOM the scientists. This happened when...
'The non-verification of the data agitated the scientists. This happened when...'
- c. Los periodistas presenciaron la no verificación de los datos.
the journalists witnessed the no verification of the data
'The journalists witnessed the non-verification of the data.'

<No + event deverbal nominal> also patterns with states, given that it rejects modifiers that refer to the velocity at which the eventuality takes place (*rápido* 'fast', *lento* 'slow'...) (3a) and it cannot be the subject of the predicates *acabar* ('to end') or *parar* ('to stop') (3b).

- (3) a. *La {rápida / lenta} no construcción del puente por parte de los obreros.
the {slow / fast} no construction of.the bridge by part of the labourers
'*The {slow / fast} non-construction of the bridge by the labourers.'
- b. *La no construcción del puente ha {acabado / parado}.
the no construction of.the bridge has {ended / stopped}
'*The non-construction of the bridge has {ended / stopped}.'

I assume Ramchand's (2008) VP structure and a syntactic analysis for nominalizations, that is, one in which the nominalizator merges once the functional phrases of the verbal base have been projected (Alexiadou 2001; Fábregas 2016), as in (4). Given that a verbal base such as *verificar* ('to verify') denotes a dynamic eventuality, ProcP must be projected, as it is the projection that codifies the dynamic component of the eventuality. I also assume that the external argument — here the initiator in terms of Ramchand (2008)— is placed in the InitP projection. Once the verbal projections have been merged, NP merges with InitP, so the verbal domain is left behind.

- (4) [NP -ción [InitP por parte del científico [ProcP ... [$\sqrt{\text{verificar}}$]]]]

My proposal is that, in the structure <no + event deverbal nominal>, *no* occupies a NegP which is merged with InitP, so it refutes the relation between the causative subevent (InitP) and the

process subevent (ProcP) (5), in line with Fábregas & González Rodríguez's (2020) proposal for the verbal domain. As a result, these subevents are now related through an inhibition relation, and not a causative one. In other words, the interpretation obtained is the so-called *negative event reading* (Higginbotham 1983; Przepiórkowski 1999) or *inhibited eventuality reading* (Fábregas & González Rodríguez 2020). The structure <no + event deverbal nominal> denotes that there is an eventuality that takes place, which consists in the initiator inhibiting itself from triggering the corresponding positive eventuality, which was necessarily expected to happen in the first place, but it does not finally occur. Thus, (1) denotes an eventuality that consists in the scientists inhibiting themselves from verifying the data. The refutation of a causative relation gives rise to an inhibited relation.

(5) [NP -ción [NegP no [InitP por parte del científico [ProcP ... [√verificar]]]]]

The analysis proposed can account for the mixed aspectual properties of <no + event deverbal nominal>. Regarding the eventive properties, *tener lugar* ('to take place'), *esto sucedió* ('this happened') and perception verbs request that its subject, its antecedent and their internal argument, respectively, refer to an event that codifies ProcP within its syntactic structure. <No + event deverbal nominal> is compatible with these predicates, and hence it presents eventive properties, given that the requisites imposed by them are satisfied by the construction, as ProcP is codified within the structure of the nominal, its subevent being identified by these predicates. As for the stative properties, modifiers such as *rápido* ('fast') or *lento* ('slow') and the predicates *acabar* ('to end') and *parar* ('to stop') select dynamic predicates. Even though the structure of <no + event deverbal nominal> codifies ProcP, the merge of NegP with InitP makes that the subevents codified within InitP and ProcP are related through an inhibition relation. This means that the subevent associated with ProcP does not take place, so <no + event deverbal nominal> lacks of dynamicity.

In addition, the analysis sketched predicts that negation requires an initiator to be present in the argument structure of the event nominal, and this prediction is borne out. Event deverbal nominals that alternate between the causative and the anticausative readings will only accept the negation in the causative version, but not in the anticausative one, as shown in (6).

(6) a. *La no bajada de los precios por sí solos fue lo que hizo que Luis
the no lowering of the prices by themselves was that what made that Luis
se enfadase.

PRON. got angry

'*The non-lowering of the prices by themselves was what made Luis angry.'

b. La no bajada de los precios por parte de la inmobiliaria fue lo que hizo que
the no lowering of the prices by part of the estate agency was that what made that
Luis se enfadase.

Luis PRON. got angry

'The non-lowering of the prices by the estate agency was what made Luis angry.'

SELECTED REFERENCES. ALEXIADOU, A. (2001): *Functional structure in nominals: Nominalizations and ergativity*. Amsterdam/Philadelphia: John Benjamins. DE SWART, H. & A. MOLENDIJK (1999): "Negation and the temporal structure of narrative discourse", *Journal of Semantics* 16, 1-42. FÁBREGAS, A. (2016): *Las nominalizaciones*. Madrid: Visor. FÁBREGAS, A. & R. GONZÁLEZ RODRÍGUEZ (2020): "On inhibited eventualities", *Natural Language & Linguistic Theory* 38 (3), 729-773. HIGGINBOTHAM, J. (1983): "The logic of perceptual reports: an extensional alternative to situation semantics", *Journal of Philosophy* 80, 100-127. HORN, L. (1989): *A natural history of negation*. Stanford, California: CSLI Publications. PRZEPIÓRKOWSKI, A. (1999): "On negative eventualities, negative concord and negative yes/no questions", in T.

Matthews & D. Strolovitc (eds.), *Proceedings of Semantics and Linguistic Theory 9*. Ithaca, New York: CLC Publications. RAMCHAND, G. C. (2008): *Verb meaning and the lexicon: A first phase syntax*. Cambridge: Cambridge University Press.

Agree, Expletives, and Person Restrictions in Bolognese

Edward J. Rubin – University of Utah

In Bol(ognese), the Gallo-Italic grammar of Bologna, Italy, the Free Inversion structure in (1-2) shows a special clitic, glossed as AI and distinct from its subject clitics (SCL), lacks agreement of the tensed verb with the postverbal subject (VS), and forbids VSs of first or second (1, 2) person (π), which precludes any simple ‘partial’ (π only) agreement account of (1), or (2b) should be well-formed. (1-2) thus differ from better studied Free Inversion structures found in Italian (3), Spanish and others, even Bol (4), where there is full agreement of the tensed verb and VS, no π -restriction, and, in Bol (and others), a SCL that agrees with the VS. (3-4) are, in these ways, identical to clauses with preverbal subjects (5).

- | | |
|--|---|
| <p>(1) a. Ai=à dscàurs la dóna.
AI=has.3S spoken the woman
‘The woman spoke.’</p> <p>b. Ai=à dscàurs äI dôn.
AI=has.3S spoken the women
‘The women spoke.’</p> | <p>(2) a. *Ai=à dscàurs mé/té/nó/vó.
AI=has.3S spoken.3S I/you.S/we/you.P
‘I/you/we/you spoke.’</p> <p>b. *Ai=è dscàurs té/vó.
AI=have.2S spoken you.S/you.P
‘You spoke.’</p> |
| <p>(3) a. Hanno parlato le donne. (Italian)
have.3P spoken the women
‘The women spoke.’</p> <p>b. Ho parlato io.
has.1S spoken I
‘I spoke.’</p> | <p>(4) a. T=è dscàurs té.
SCL.2S=have.2S spoken you.S
‘You spoke.’</p> <p>b. A=dscurän nó.
SCL.1P=speak.1P we
‘We are speaking.’</p> |
| <p>(5) a. Le donne hanno parlato.
the women have.3P spoken
‘The women spoke.’</p> <p>b. ÄI dôn äli=an dscàurs.
the women SCL.3PF=have.3P spoken
‘The women spoke.’</p> | <p>(6) a. L= é bél ch’ al= piôv.
SCL.3SM= is beautiful that SCL.3SM= rains
‘It’s great that it’s raining.’</p> <p>b. La P, al= pèr ch’ la= séppa môrta
the P SCL.3SM= seems that SCL.3SF= is.SBJ died
‘Petrarca, it seems that she died (in France).’</p> |

Analyses of (3-4) standardly have an expletive *pro* (*expl*) where a corresponding preverbal subject occurs (5) (Rizzi 1982, 1986, Burzio 1986, Cardinaletti 1997, 2004, Belletti 2005, Roberts 2010, etc). In (5), Agree (and movement) underlies the subject’s Case-Licensing, control of agreement, and EPP (Chomsky 2000, 2004, 2008). In (3-4), however, *expl* is deficient and only satisfies EPP, while long-distance Agree(T,VS) explains the Licensing of and agreement with *expl*’s ‘associate’. We adopt this analysis for (3-5), and propose that there is an *expl* also in (1-2). When agreement-lacking data like (1) is discussed at all (e.g. in Fiorentino, Brandi & Cordin 1989, Belletti 2005, Roberts 2010), similar analyses are supplied, but the *expl* is usually asserted to have ‘default’ 3SM features and to control agreement. This raises difficult questions: If *expl*_{3SM}, presumably not deficient, fully values T and thus inactivates it, how is VS Licensed (Lasnik 1995, Chomsky 2004:114)? If 3SM is default, then why does (1) not show it (via an SCL, as in Fiorentino)? Bol does have a SCL.3SM with (the *expl* with) weather verbs and clausal complements (6), just not in (1). Therefore, which *expl* is default, and what is the other? How does each interact with Agree? It is not simply that there are ‘deficient’ *expl*s in data like (3-4) and all other *expl*s are ‘default’: a deeper analysis is needed.

We propose that (1-2), (3-4), and (6) each involve *expl*s with distinct features and thus distinct interactions with Agree and resulting morpho-syntactic properties. As discussed above, the *expl* in (3-4) is deficient, only able to satisfy EPP but not to participate in Agree. The others have richer feature structures, are independent nominals needing Licensing, and interact with Agree: *Expl*.3SM occurs in (6), where it is the only nominal available for Agree with $T_{u\phi}$; *Expl*.3S occurs in (1-2) (signaled by AI and merged with it below T), i.e. when there is another non-associate nominal in the domain. These two both satisfy EPP and, by Agree, determine the agreement, which presumably simultaneously Case-Licenses them. To address the Licensing of VS and π -restrictions in (1-2), we argue that such data involves Multiple/Cyclic Agree (MA/CA) (Béjar

& Rezac 2003, 2009, Nevins 2007, 2011), an elaboration of Agree in which a probe relates to more than one goal (cf. Tortora 1999). We also adopt the ideas that individual φ -features can probe separately and that π is responsible for Case-Licensing. MA/CA is explicitly tied to Case-Licensing (Béjar & Rezac 2009:47, Nevins 2011:955), as we will develop here, though in Romance, its most familiar effect, the Person Case Constraint (PCC), seen in Bol (7), involves clitics, which are generally held to be independent of it.

- (7) a. Pèvel al= {m= , s= , t= , v= , i= } la= dà.
 P. SCL.3SM= DCL.1S= DCL.1P= DCL.2S= DCL.2P= DCL.3= ACL.3SF= gives.3S
 ‘Pèvel gives it to me/us/you/you/him/her/them.’
- b. *Pèvel al= i= {m= , s= , t= , v= } dà.
 P. SCL.3SM= DCL.3= ACL.1S= ACL.1P= ACL.2S= ACL.2P= gives.3S
 ‘Pèvel gives him/her/them me/us/you/you.’

In Bol, as in many Romance (and other) grammars, a direct object/accusative clitic (ACL) may not be 1 or 2 when there is also an indirect object/dative clitic (DCL). Simplifying the accounts, these two clitics Agree with a single probing head, and ill-formedness arises either through depletion of the probe’s features by the ACL before Licensing the DCL (CA) or through a locality effect imposed by the c-commanding DCL (MA).

We extend the same ideas to (1-2), where a c-commanding *expl.3S* is incompatible with a lower $\underline{VS}_{1/2}$. For the exposition below, we adopt MA, but CA is equally compatible with with our proposal. In (1), the \underline{VS} is 3, and thus π on T successfully Agrees with both it and a $\underline{VS}_{3S/P}$, and thereby Licenses them (8). # on T probes separately, and finds the local *expl.3S*. (This involves standard Agree, and not MA/CA, which may be operative in the significantly different Icelandic data discussed in Anagnostopoulou 2003.) In (2), where \underline{VS} is 1/2, probing by π on T encounters the locality effect (or the full deletion effect with CA) mentioned above (9). One of the two nominals is not Licensed, and the derivation crashes. Unlike in (1-2), with two nominals needing Licensing, in (3-4) and (6) there is only one nominal in the domain of T. The *expl* in (3-4) is deficient, and Agree(T, \underline{VS}) values/deletes $u\varphi$ on T and Licenses the associate \underline{VS} (10). The *expl.3SM* in (6), as the only nominal in T’s domain, Agrees with it, and thus is Licensed and values/deletes $u\varphi$ on T (11).

- (8) [*expl.3S* ai=à-T_{*uπ, u#*} [*expl.3S* ai=à [dscàurs la dôna.3SF/âl dôn.3PF]] Structure for (1)
- (9) [*expl.3S* ai=à-T_{*uπ, u#*} [*expl.3S* ai=à [dscàurs mé.1S/té.2S/nó.1P/vó.2P]] Structure for (2a)
- (10) [*expl* t=è-T_{*uφ*} [è [dscàurs té_{2S}]] Structure for (4), & all Case Transmission
- (11) [*expl.3SM* al=piôv/pér-T_{*uφ*} [piôv/pér ch’] Structure for (6), weather/clausal comp. Vs

Cross-linguistic Variation. Other grammars may lack (some of) these *expls* (English) or the features of *expls* available in a grammar may differ. E.g., Italian has the deficient *expl* for (3) and *expl.3SM* for data like (6), but not *expl.3S* nor the MA/CA in T to support it; Fiorentino has defective *expl*, and *expl.3SM* for structures like both (1) and (6); Rodoretto di Prali seems to have *expl.3SF* with at least weather verbs. Under the current analysis, the observable variations are always due to Agree (and/or its elaboration, MA/CA) and the way in which it interacts with the particular properties of T and the available items in each grammar.

Anagnostopoulou 2003. *The syntax of ditransitives: Evidence from clitics*. Mouton. Béjar & Rezac 2003. Person licensing and the derivation of PCC Effects. In Pérez-Leroux & Roberge (eds) *Romance Linguistics: Theory and Acquisition (LSRL 32)*. Benjamins. 49-62. Béjar & Rezac 2009. Cyclic agree. *Linguistic Inquiry* 40.1:35-73. Belletti 2005. Extended doubling and the VP periphery. *Probus* 17:1-35. Burzio 1986. *Italian syntax: A government-binding approach*. Reidel. Cardinaletti 1997. Subjects and clause structure. In Haegeman (ed) *The New Comparative Syntax*. 33-63. Cardinaletti 2004. Toward a cartography of subject positions. In Rizzi 2004, *The Structure of CP and IP: The Cartography of Syntactic Structures, Volume 2*. Oxford. Chomsky 2000. Minimalist Inquiries. In Martin, et. al. (eds) *Step by Step*. MIT. Chomsky 2008. On Phases. In Freiden et. al. (eds) *Foundational Issues in Linguistic Theory*. MIT. Lasnik 1995. Case and Expletives Revisited. *Linguistic Inquiry* 26.4:615-633. Nevins 2007. The Representation of Third Person an Its Consequences for Person-Case Effects. *Natural Language and Linguistic Theory* 25.5:273-313. Nevins 2011. Multiple agree with clitics: person complementarity vs. omnivorous number. *Natural Language and Linguistic Theory* 29.4:939-971. Rizzi 1982. *Issues in Italian syntax*. Foris. Rizzi 1986. On the status of subject clitics in Romance. In Jaeggli & Silva-Coravalan (eds) *Studies in Romance linguistics*. 391-420. Roberts 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. MIT Press. Tortora 1999. Agreement, Case, and i-subjects. In Tamanji, Hirotani, & Hall (eds) *NELS 29*. 397-408.

The Influence of Language Typology in Multilinguals' Language Learning Strategies

Sara Sáez Fajardo

University of Illinois at Urbana-Champaign

This ongoing project investigates how native Spanish speakers from different linguistic backgrounds use language learning strategies (LLS) to learn French as a foreign language. Particularly, this study looks at how the typology of the learner's L1 and L2 influences their use of LLS.

In this study, I compare two populations of high school students (11 to 18 years old): learners of French in a bilingual context (Spanish-Catalan) and learners of French in a monolingual context (Spanish). Both groups have studied English as an L2 and learn French as an additional language in high school.

Many previous studies compare how monolinguals and bilinguals use LLS to learn an L2/L3 (e.g., Sanz 2000; Cenoz, 2003). However, most studies have focused on English learning and have not compared bilingual and multilingual populations that were matched in age and instruction on foreign languages. Moreover, there have not been studies that observe the variable of language typology (more specifically, language proximity) in the different use of LLS. Nevertheless, studies of the role of background languages in the acquisition of an L3 suggest that language typology is one of the main influential variables in the L3 learning process (e.g., Falk & Bardel, 2010; Ringbom, 2011).

This study compares a group of L1 speakers of Spanish to a group of L1/L2 speakers of Spanish and Catalan. Given that Spanish and Catalan are Romance languages, together with French, but Catalan is more similar to French than Spanish is, it would be expected that its speakers would learn French differently than the monolingual Spanish speakers. A visible difference in their learning process could be in their use of LLS.

In order to investigate a potential difference in the use of LLS by multilinguals with different linguistic backgrounds, I used Dmitrenko's (2017) adaptation for multilinguals of Oxford's (1990) Strategy Inventory for Language Learning (M-SILL). While Oxford's survey has been the most popular instrument to evaluate the use of language learning strategies in L2 learning, and it has also been employed in L3 research (e.g., Haukas, 2012; Sanz, 2000), it was constructed for the L2 learner. Dmitrenko's adaptation is more appropriate for multilingualism research and is the chosen instrument in this project.

The M-SILL survey includes 70 statements that describe activities and beliefs related to language learning. Participants rate on a Likert Scale 'how true of them' those affirmations are. The survey is being administered to the previously mentioned groups of participants in their respective high schools (one in the province of Valencia and the second in the province of Huesca, in Spain). Approximately 100 students of French from each high school (classified by class) will take the test, after completing a language background questionnaire and a short translation task, French to Spanish, that activates the students' awareness of their learning strategies.

This study aims to answer one specific question on the use of LLS by different groups of learners (i.e., which group uses strategies more often). This question will be answered by comparing the two groups (L1 Spanish speakers and L1/L2 Spanish – Catalan speakers). Considering that many studies reported more frequent use of strategies by multilinguals (e.g., Kemp, 2007), will there be a difference when comparing groups of multilinguals with different linguistic backgrounds?

Moreover, a comparison within the bilingual Spanish-Catalan group will also be run to observe potential differences across sequential and simultaneous bilinguals in their use of LLS when learning French. Previous research (Merkelbach, 2011) has shown that simultaneous bilinguals use more LLS and more frequently than sequential bilinguals, who are also expected to use more strategies and more frequently than the learners in the monolingual region.

References:

- Cenoz, J. (2003). "The additive effect of bilingualism on third language acquisition: A review." *International Journal of Bilingualism*, 7, 1, 71–87.
- Dmitrenko, V. (2017). "Language learning strategies of multilingual adults learning additional languages". *International Journal of Multilingualism*, 14:1, 6-22.
- Falk, Y. and Bardel, C. (2010). The study of the role of the background languages in third language acquisition. The state of the art. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 48, 185-219.
- Haukas, A. (2012). "A Comparison Between L2 and L3 learners' Strategy Use in School Settings". *The Canadian Modern Language Review*, 71, 4, 383-405.
- Kemp, C. (2007). Strategic processing in grammar learning: Do multilinguals use more strategies? *International Journal of Multilingualism*, 4(4), 241–261.
- Merkelbach, C. (2011). "Wie unterscheiden sich die Lernstrategien beim Erlernen von L2 und L3? Ergebnisse einer empirischen Studie bei taiwanischen Deutsch-als-L3-Lernenden." *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 16 (2), 126–146.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. New York, NY: Newbury House; Harper & Row.
- Ringbom, H. (2011). Perceived Redundancy or Crosslinguistic Influence? What L3 Learners' Material Can Tell Us About the Causes of Errors. In Angelis, G. and Dewaele, J., editors, *New Trends in Crosslinguistic Influence and Multilingualism Research*, 19-24.
- Sanz, C. (2000). "Bilingual education enhances third language acquisition: Evidence from Catalonia". *Applied Psycholinguistics*, 21, 23-44.

Long passives of causatives and perception verbs in Italian: implications for phase theory

Michelle Sheehan (*Anglia Ruskin Cambridge*) & Jan Casalicchio (*Palermo*)

In Romance, long passivization of perception/causative verbs is highly restricted. One potential explanation for this is that the complements of these verbs are often phasal voicePs which do not allow movement of embedded arguments to the main clause (Sheehan & Cyrino 2018). Italian, however, is a notable exception to this general pattern, since ECM complements of *vedere* ‘see’/*sentire* ‘hear’ and clause union complements of causative verbs both productively allow long passives. We propose that this intra-Romance difference is due to microvariation in the syntax of causatives/perception verbs: reduced complements of these verbs be of four different sizes, depending on the language and on the matrix verb: TP, voiceP, vP or VP, with only voiceP complements blocking long passivization. Italian is key to understanding this microvariation because it allows all four sizes of complement. In the other Romance languages discussed, on the other hand, the voiceP complement occurs in most cases, with the only exception being the *faire-par* construction, which (when it is available) selects a bare VP (see Folli & Harley 2007).

Italian differs from French, Spanish and European Portuguese in much more freely allowing passivization of perception and causative verbs taking reduced clausal complements involving ECM (1a) or clause union (1b) (see Burzio 1986, Guasti 1993, Cinque 2003) – note that *fare* requires clause union, while *vedere/sentire* allow either ECM or clause union.

- 1) a. Gino è stato **visto** bere quel vino. (ECM)
 Gino is been seen drink.INF that wine ‘Gino was seen drinking that wine.’
 b. La mela è stata **fatta** mangiare al bambino (da Maria). (Clause union)
 the apple is been.F made.F eat.INF to.the child by Maria
 ‘The apple has been made to be eaten by the child (by Maria).’

Considering first the **ECM complements** of perception verbs which surface with a preverbal accusative subject, promotion of the external argument of a transitive verb is banned/degraded in all the other languages (illustrated here for Spanish only) (see Kayne 1975 on French, Hornstein, Martins & Nunes 2006, 2010 on European Portuguese and Burzio 1986, Guasti 1993 on Italian):

- 2) ??María fue vista robar el coche.
 Maria was seen.F steal.INF the car

Following Sheehan & Cyrino (2018), this is because **ECM complements are phasal** voicePs and long A-movement cannot cross two phase heads because of the Phase Impenetrability Condition 2 (Chomsky 2001). As schematized in (3), at the point at which the matrix T probes for a DP to agree with, all of the arguments of the embedded ECM complement have been spelled out and so are not visible to T, leading to a derivational crash:

- 3) [TP T_[uPhi*] [voiceP **fue** [vP Vista [voiceP **voice** [~~vP María v~~ [~~vP robar~~ [~~DP el coche~~]]]]]]]]]

We propose that ECM complements of passivisable perception verbs in Italian (*vedere* and *sentire*) differ from other ECM complements in being larger, including a T-related projection with an EPP feature. The presence of this EPP feature feeds long A-movement by making the embedded ‘subject’ visible to matrix T, in the same way that the presence of *to* facilitates long passivisation in English (*John was made/seen *(to) leave.*):

- 4) [TP T_[uPhi*] [voiceP **stato** [vP visto [TP **Gino** T [voiceP **voice** [~~vP Gino v~~ [~~vP bere~~ [~~DP quel vino~~]]]]]]]]]

We provide evidence for this proposal by contrasting *vedere* ECM complements with the ECM complements of other Italian perception verbs that do not allow passivisation and are thus expected to be voicePs. For example, the distribution of high adverbs and modal verbs shows that the ECM complement of *vedere* is larger than that of *guardare* (which does not permit long passivization).

- 5) L' ho {vista/*guardata} dover rinunciare all' università.
her=have.1SG seen.F/watched.F must renounce to.the university
- 6) L' abbiamo {vista/*guardata} sicuramente crescere fino a diventare la [fiera] più grande
Her=have.1PL seen/watched surely grow until to become the fair most big

In relation to **clause union complements of the faire-infinitive kind**, we propose that Italian permits long passivization because these complements are smaller than voiceP and so non-phasal. While all the languages under discussion ban periphrastic passives under FACERE (for semantic reasons), Italian also disallows the expression of voice as *si* (Lepschy 1976, Burzio 1986):

- 7) La hanno fatta pentire / *pentirsi.
her= have.3PL made repent repent-SE
'They made her repent.' (Lepschy 1976: 159)

Of the other languages, only European Portuguese bans *se* in this context, but this is part of a broader pattern of *se*-suppression in Portuguese non-finite clauses, which we take to be a distinct phenomenon (Martins and Nunes 2017a, b). Again, we can compare the behaviour of Italian *fare* both with the FACERE cognate in other Romance languages but also with other verbs in Italian. For example, while *vedere* can also participate in clause union, it does not behave like *fare* when it does, but rather like FACERE in other Romance languages; e.g., *si* cannot be suppressed and long passivization is ruled out in the presence of a dative causee.

- 8) La mela è stata vista mangiare {*a Gianni / da Gianni}
the apple is been seen eat.INF to G. by G.

On the other hand, long passives of the faire-par (*da Gianni* in (8)) are possible under *vedere*. Long passives of this kind are available across all the Romance languages under discussion where *faire par* is available. As a result, only internal arguments (objects or unaccusative subjects) can be promoted (9) – we take this to result from the fact that causative/perception verbs can embed a VP with no external argument in the *faire-par* construction (Folli & Harley 2007, Guasti 2017).

- 9) a. Todos estes palácios foram mandados construir a arquitectos italianos.
all these palaces were had build.INF to architects Italian.PL
Lit. 'All these palaces were had built by Italian architects.'
- b *Depois da prova de salto, o atleta foi mandado correr.
after of.the test of jump, the athlete was had run.INF (Gonçalves 1999, 422)

Across the languages, the possibility of long passives is closely connected to the availability of *faire-par*. Thus, in European Portuguese, only *mandar* surfaces in the *faire-par* (Raposo 1981, 239; fn 13) meaning that only it permits long passivization (the dative in (9a) is a benefactive):

- 10) Os meus pais mandaram/*fizeram/*deixaram construir uma casa.
the my parents had made let build a house
'My parents had a house built.'

In short, long passives are also possible where the matrix verbs selects a VP (in the *faire-par*) or a vP (in the Italian *faire-inf*) because A-movement can cross one clausal boundary:

- 11) [TP T_[uPhi*] [voiceP **foi** [vP mandado [vP construir [DP o palácio]]]]]]]]

Long passives are blocked only where the complement of the matrix verb is a phase (CP/voiceP).

Selected references: Chomsky, N. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*. MIT Press. Hornstein, N, J Nunes, and A-M Martins. 2010. Perception and causative structures in English and European Portuguese. *Syntax*. Sheehan, M and S Cyrino. 2018. Why do some ECM verbs resist passivisation? A phase- based explanation. *Proceedings of NELS 48*.

Long passives of causatives and perception verbs in Italian: implications for phase theory

Michelle Sheehan (*Anglia Ruskin Cambridge*) & Jan Casalicchio (*Palermo*)

In Romance, long passivization of perception/causative verbs is highly restricted. One potential explanation for this is that the complements of these verbs are often phasal voicePs which do not allow movement of embedded arguments to the main clause (Sheehan & Cyrino 2018). Italian, however, is a notable exception to this general pattern, since ECM complements of *vedere* ‘see’/*sentire* ‘hear’ and clause union complements of causative verbs both productively allow long passives. We propose that this intra-Romance difference is due to microvariation in the syntax of causatives/perception verbs: reduced complements of these verbs be of four different sizes, depending on the language and on the matrix verb: TP, voiceP, vP or VP, with only voiceP complements blocking long passivization. Italian is key to understanding this microvariation because it allows all four sizes of complement. In the other Romance languages discussed, on the other hand, the voiceP complement occurs in most cases, with the only exception being the *faire-par* construction, which (when it is available) selects a bare VP (see Folli & Harley 2007).

Italian differs from French, Spanish and European Portuguese in much more freely allowing passivization of perception and causative verbs taking reduced clausal complements involving ECM (1a) or clause union (1b) (see Burzio 1986, Guasti 1993, Cinque 2003) – note that *fare* requires clause union, while *vedere/sentire* allow either ECM or clause union.

- 1) a. Gino è stato **visto** bere quel vino. (ECM)
 Gino is been seen drink.INF that wine ‘Gino was seen drinking that wine.’
 b. La mela è stata **fatta** mangiare al bambino (da Maria). (Clause union)
 the apple is been.F made.F eat.INF to.the child by Maria
 ‘The apple has been made to be eaten by the child (by Maria).’

Considering first the **ECM complements** of perception verbs which surface with a preverbal accusative subject, promotion of the external argument of a transitive verb is banned/degraded in all the other languages (illustrated here for Spanish only) (see Kayne 1975 on French, Hornstein, Martins & Nunes 2006, 2010 on European Portuguese and Burzio 1986, Guasti 1993 on Italian):

- 2) ??María fue vista robar el coche.
 Maria was seen.F steal.INF the car

Following Sheehan & Cyrino (2018), this is because **ECM complements are phasal** voicePs and long A-movement cannot cross two phase heads because of the Phase Impenetrability Condition 2 (Chomsky 2001). As schematized in (3), at the point at which the matrix T probes for a DP to agree with, all of the arguments of the embedded ECM complement have been spelled out and so are not visible to T, leading to a derivational crash:

- 3) [TP T_[uPhi*] [voiceP **fue** [vP Vista [voiceP **voice** [~~vP María v~~ [~~vP robar~~ [~~DP el coche~~]]]]]]]]]

We propose that ECM complements of passivisable perception verbs in Italian (*vedere* and *sentire*) differ from other ECM complements in being larger, including a T-related projection with an EPP feature. The presence of this EPP feature feeds long A-movement by making the embedded ‘subject’ visible to matrix T, in the same way that the presence of *to* facilitates long passivisation in English (*John was made/seen *(to) leave.*):

- 4) [TP T_[uPhi*] [voiceP **stato** [vP visto [TP **Gino** T [voiceP **voice** [~~vP Gino v~~ [~~vP bere~~ [~~DP quel vino~~]]]]]]]]]

We provide evidence for this proposal by contrasting *vedere* ECM complements with the ECM complements of other Italian perception verbs that do not allow passivisation and are thus expected to be voicePs. For example, the distribution of high adverbs and modal verbs shows that the ECM complement of *vedere* is larger than that of *guardare* (which does not permit long passivization).

- 5) L' ho {vista/*guardata} dover rinunciare all' università.
 her=have.1SG seen.F/watched.F must renounce to.the university
- 6) L' abbiamo {vista/*guardata} sicuramente crescere fino a diventare la [fiera] più grande
 Her=have.1PL seen/watched surely grow until to become the fair most big

In relation to **clause union complements of the faire-infinitive kind**, we propose that Italian permits long passivization because these complements are smaller than voiceP and so non-phasal. While all the languages under discussion ban periphrastic passives under FACERE (for semantic reasons), Italian also disallows the expression of voice as *si* (Lepschy 1976, Burzio 1986):

- 7) La hanno fatta pentire / *pentirsi.
 her= have.3PL made repent repent-SE
 'They made her repent.' (Lepschy 1976: 159)

Of the other languages, only European Portuguese bans *se* in this context, but this is part of a broader pattern of *se*-suppression in Portuguese non-finite clauses, which we take to be a distinct phenomenon (Martins and Nunes 2017a, b). Again, we can compare the behaviour of Italian *fare* both with the FACERE cognate in other Romance languages but also with other verbs in Italian. For example, while *vedere* can also participate in clause union, it does not behave like *fare* when it does, but rather like FACERE in other Romance languages; e.g., *si* cannot be suppressed and long passivization is ruled out in the presence of a dative causee.

- 8) La mela è stata vista mangiare {*a Gianni / da Gianni}
 the apple is been seen eat.INF to G. by G.

On the other hand, long passives of the faire-par (*da Gianni* in (8)) are possible under *vedere*. Long passives of this kind are available across all the Romance languages under discussion where *faire par* is available. As a result, only internal arguments (objects or unaccusative subjects) can be promoted (9) – we take this to result from the fact that causative/perception verbs can embed a VP with no external argument in the *faire-par* construction (Folli & Harley 2007, Guasti 2017).

- 9) a. Todos estes palácios foram mandados construir a arquitectos italianos.
 all these palaces were had build.INF to architects Italian.PL
 Lit. 'All these palaces were had built by Italian architects.'
- b *Depois da prova de salto, o atleta foi mandado correr.
 after of.the test of jump, the athlete was had run.INF (Gonçalves 1999, 422)

Across the languages, the possibility of long passives is closely connected to the availability of *faire-par*. Thus, in European Portuguese, only *mandar* surfaces in the *faire-par* (Raposo 1981, 239; fn 13) meaning that only it permits long passivization (the dative in (9a) is a benefactive):

- 10) Os meus pais mandaram/*fizeram/*deixaram construir uma casa.
 the my parents had made let build a house
 'My parents had a house built.'

In short, long passives are also possible where the matrix verbs selects a VP (in the *faire-par*) or a vP (in the Italian *faire-inf*) because A-movement can cross one clausal boundary:

- 11) [TP T_[uPhi*] [voiceP **foi** [vP mandado [VP construir [DP o palácio]]]]]]]]

Long passives are blocked only where the complement of the matrix verb is a phase (CP/voiceP).

Selected references: Chomsky, N. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*. MIT Press. Hornstein, N, J Nunes, and A-M Martins. 2010. Perception and causative structures in English and European Portuguese. *Syntax*. Sheehan, M and S Cyrino. 2018. Why do some ECM verbs resist passivisation? A phase- based explanation. *Proceedings of NELS 48*.

Prosody and visual perception in the processing of gapping sentences in Brazilian Portuguese

This study investigates how prosodic and visual cues can influence the interpretation of ambiguous Brazilian Portuguese (BP) conjoined sentences like (1).

(1) O Pedro levou a Julia na festa e o Bruno no churrasco da empresa.

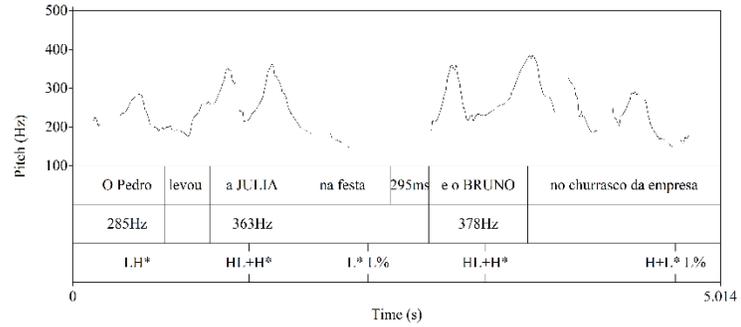
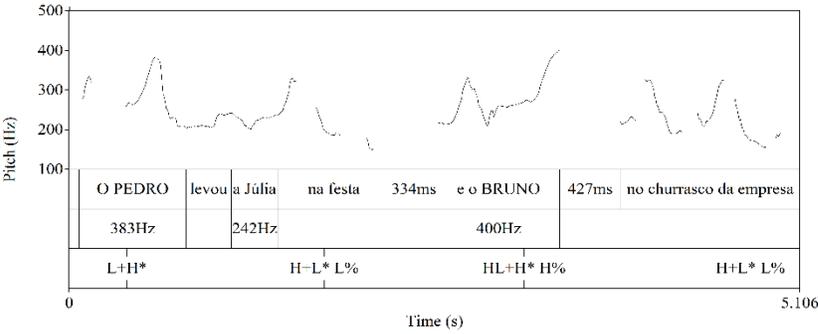
(Peter took Julia to the party and Bruno to the company's barbecue.)

a. *subject reading*: Peter took Julia to the party and Bruno ~~took Julia~~ to the company's barbecue.

b. *object reading*: Peter took Julia to the party and ~~Peter took~~ Bruno to the company's barbecue.

The subject reading shows a gapping structure (Sag, 1977); there was deletion of the VP (*visited Julia*) in the second conjunct. The object reading leads to a conjunction reduction (Johnson, 2018); two objects are conjoined within the same VP (*Peter [took Julia and Bruno]*). Carlson (2002) found a preference for the object reading, accounted by the Minimal Attachment Principle (Frazier 1987): speakers tend to choose the simplest syntactic structure (i.e., number of nodes) compatible with the input. On the other hand, the manipulation of prosody can affect the interpretation of these sentences. In an auditory questionnaire, Carlson (2002) manipulated the prosodic parallelism between the arguments in different conjuncts with pitch accents. A promoting subject prosody was shown to raise the percentage of gapping responses from 38% to 44%. In our study, we crossed prosodic structure with biasing pictures. In the subject prosody, the ambiguous DP (*Bruno*) aligns with the subject DP (*Pedro*); while in the object prosody, it aligns with the object DP (*Julia*). The aligned DPs have similar pitch accents and increased duration and intensity as well. Subject prosody has IPh boundaries after the first conjunct and the ambiguous DP, whereas object prosody has an IPh only after the first conjunct – see pictures 1-2 for pitch tracks. Additionally, three pictures (3-5) were designed: one compatible with the first conjunct (*Peter took Julia*), one biasing subject reading (*Bruno took Julia*), and another biasing object reading (*Peter took Bruno*). Half of the time pictures 3 and 4 were shown together and the other half pictures 3 and 5. Their position on the screen was counterbalanced. By crossing prosody and pictures, there are four conditions: *Subject Prosody Subject Picture* (SPSP); *Subject Prosody Object Picture* (SPOP); *Object Prosody Object Picture* (OPOP); *Object Prosody Subject Picture* (OPSP). The experiment (N=28) was a spoken language comprehension task with the Visual World Paradigm (Tanenhaus & Trueswell, 2006) which tested 20 sentences like (1). BP native speakers listened to the sentences while two pictures were being shown side by side on the screen. After listening to the sentences and seeing the pictures, they answered a comprehension question: *What happened on the company's barbecue? a) Peter took Bruno there or b) Bruno took Julia there*. We measured the participants' eye movements (total fixation duration/TFD and fixation count/FC) to both pictures on the screen, while they were listening to the second conjunct (*and Bruno to the company's barbecue*). Picture 3 was the control picture in all conditions. Picture 4 was the target in SPSP and OPSP, while picture 5 was the target in OPOP and SPOP. The results showed that participants looked more to the target pictures in all conditions (Graphs 1-2). There was a main effect of image type in TFD ($\beta=-413.46$, $SE=70.34$, $t=-5.878$, $p < .01$) and FC ($\beta=-0.41$, $SE=31.57$, $t=-12.871$, $p < .01$). A Tukey post-hoc test also shown a significant difference (target x control) in the four conditions. When prosodic and visual cues pointed to the same direction, subject responses were 73% in SPSP. When the cues conflicted, the biasing picture seems to have influenced more; subject responses dropped

to 60% in OPSP (Graph 3). These results suggest that listeners were able to use prosodic and visual cues in processing to interpret the ambiguous gapping sentences.

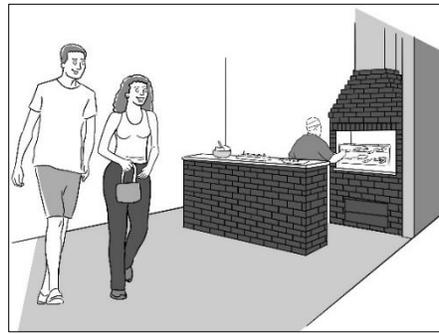


Picture 1: Example of Subject Prosody

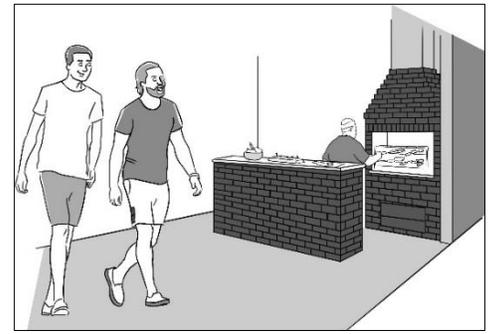
Picture 2: Example of Object Prosody



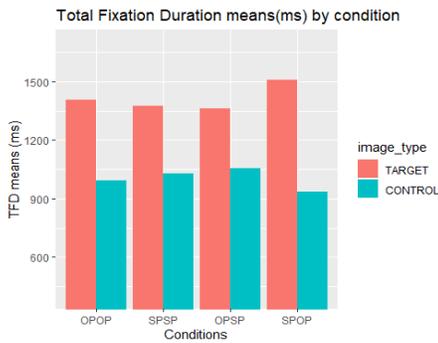
Picture 3: First Conjunct



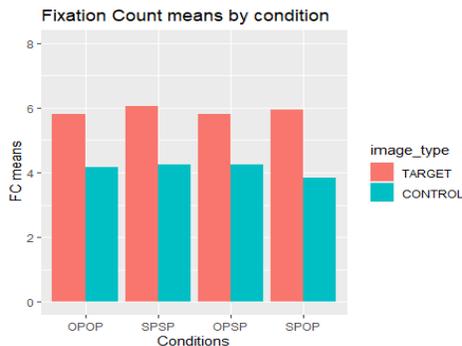
Picture 4: Subject Bias



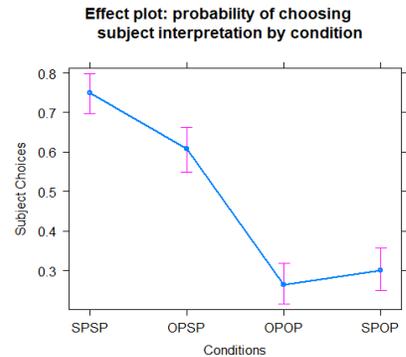
Picture 5: Object Bias



Graph 1: TFD means



Graph 2: FC means



Graph 3: subject choices

References

Carlson, K. (2002) *Parallelism and prosody in the processing of ellipsis sentences*. New York: Routledge.

Frazier, L. (1987) Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention & performance XII* (pp. 601–681). Hove, UK: Lawrence Erlbaum.

Johnson, K. (2018) Gapping and Stripping. In: Craenenbroeck, J. & Temmerman, T (Eds.) *The Oxford Handbook of Ellipsis*. Oxford University Press, p. 1-56. DOI: 10.1093/oxfordhb/9780198712398.013.24

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Sag, I. (1977) *Deletion and logical form*. Bloomington, IN: Indiana University Linguistic Club Publications.

Tanenhaus, M. K. & Trueswell, J. C. (2006) Eye Movements and Spoken Language Comprehension. In: Traxler, M. J. & Gernsbacher, M. A. (Ed.) *Handbook of Psycholinguistics*. 2nd Edition. Elsevier Academic Press. p. 863-900.

**The grammar and use of the epistemic future:
a comparison between two Italo-Romance speech communities**

OVERVIEW AND GOALS

In the evolution from Latin to Romance, the future indicative of Latin (i.e. AMABO ‘I will love’, AUDIAM ‘I will hear’), of synthetic nature, was replaced by innovative formations in Romance, including a periphrastic structure formed with the infinitive of the verb and the inflected forms of the verb ‘have’ (HABEO), which underwent grammaticalization and resulted in the synthetic future found in most Romance languages, e.g. (*elle*) *parlera* (French) ‘she shall speak’, *volverá* (Spanish) ‘s/he will come back’, *escreveremos* (Portuguese) ‘we shall write’, *andremo* (standard Italian) ‘we will go’. According to the traditional view, in the Italo-Romance landscape this type of future is only indigenous to Tuscan (and standard Italian) and northern Italian varieties, while in southern varieties no original synthetic future is attested (D’Ovidio 1878:183; Rohlfs 1968:333; Tekavčić 1980:332-3). Yet, according to more recent and accurate investigations, the synthetic future, which is attested in southern Italian varieties in a patch-like distribution and through defective paradigms, has to be seen as a relic trace of forms that were more largely spread in the past. The synthetic future that we observe today in southern Italian dialects underwent a semantic shift from expressing futurity to conveying epistemic modality only.

This paper discusses a case study based on two distinct varieties of the same upper-southern Italo-Romance variety, i.e. the dialect of Verbicaro, spoken indigenously in north-western Calabria (Verbicarese 1) as well as in the community of Verbicarese speakers of São Paulo (Verbicarese 2), who migrated to Brazil in two waves during the 20th century. The evidence discussed in this paper was collected in different phases during field-work trips to the two relevant localities. Most of the data relative to Verbicarese 2 was gathered in June 2010 through one-on-one interviews with speakers of Verbicarese in three different neighborhoods of São Paulo.

EVIDENCE

The two varieties of Verbicarese display several points of microvariation with respect to the synthetic epistemic future. In both varieties two forms of this verb are attested, a simple one and a compound one ((1) vs (2)). Also, all the forms exhibit more than one morphological type depending on the different make-up of the endings (see ‘Type A’ and ‘Type B’ in Table 1).

In both varieties the original future value is lost and the synthetic future only conveys an epistemic interpretation, i.e. a value ascribable to the category of non-factual, *irrealis* modality: it expresses the speaker’s subjective evaluation of the probable/possible truthfulness of the propositional content of the utterance.

- (1) a. *Question*: What time do you think the babies eat usually?
b. *Answer*: Mandzərrəjənə/Mandzərrajənə i tria. (Verbicarese 1 and 2)
eat.FUT.3PL the three
‘Usually they eat at three./As far as I know, they eat at three.’
- (2) a. *Question*: What time do you think they came yesterday?
b. *Answer*1: Avərrəjənə/Avərrəjənə vənutə a matina. (Verbicarese 1)
have.FUT.3PL come.PPT the morning

How fast did Cicero speak?

The speech rate of Classical Latin versus its Romance descendants

Daniel Stelzer, University of Illinois at Urbana-Champaign

1 Overview

In 2019, Coupé et al demonstrated experimentally that, while rate of speech (syllables spoken per second) and information density (bits conveyed per syllable) vary significantly between languages, their product (bits conveyed per second) does not. This “information rate” seems to be a property of the communicative niche of human language in general, a true linguistic universal.

Coupé approximated information density as the *syllable-conditional entropy*: the average amount of information conveyed by a single syllable, given knowledge of what syllable came before it. And while his method of measuring speech rate is impossible without living native speakers, the syllable-conditional entropy can be calculated from a purely written corpus (after Oh (2015)), by measuring the frequencies of all syllables and contexts.

In their previous studies, Coupé and Oh mostly used corpora with tens or hundreds of millions of tokens. However, most corpora for dead languages are significantly smaller than this; the entire surviving body of Classical Latin literature, for example, has barely over two million, on an order of magnitude with just two weeks of the *New York Times*.

To compensate for this, we developed new methods to extrapolate from a limited corpus, showing what the entropy would be if the corpus were infinitely large, and how much uncertainty results from the limited sample size. Using these new methods on the Packard Humanities Institute corpus, we were able to estimate the information density for Classical Latin. From this we can predict the natural rate of speech of the Romans during the classical era, and compare it to that of modern Romance languages.

2 Methods

The first step in calculating information density involved converting the Latin text to a phonemic representation, generally following Allen (1978). The phonemic form is almost entirely predictable from the written form, with only a few exceptions: standard Latin orthography doesn’t indicate vowel length or distinguish vowels from approximants.

These exceptions were handled automatically using a system developed by Winge (2015), which has been shown to disambiguate Latin orthography with over 98% accuracy. The rest of the preprocessing involved giving every phoneme an unambiguous representation and calculating syllable breaks, using a modified version of an algorithm by Johnson (2019). The only non-phonemic detail included was neutralization: if the distinction between two phonemes was completely lost in a particular environment, that distinction was removed in preprocessing.

Next, we had to compensate for the size of the corpus. As Oh noted, as the corpus size increases, the estimated entropy grows sharply at first, then levels off and converges. By artificially reducing larger corpora (discarding a certain number of tokens at random), we determined experimentally that this growth follows a negative power function $H = a_1 - a_2(x - a_3)^{-a_4}$; fitting this function to the data with least squares then shows what value the entropy would converge to if our corpus were infinitely large.

Although this method allows us to extrapolate from a small corpus, there’s still a risk that the corpus may not be representative. To test this, we used “author jackknifing”: we gathered a list of the 14 authors who contributed at least 100,000 tokens to the corpus, then performed the calculation repeatedly, removing a different one of these authors each time. The standard deviation of the results gives an approximation of how much the entropy value could be swayed by any particular author’s style.

3 Results and Discussion

With these methods, we estimated a conditional entropy of 6.32 bits per syllable, with a standard error of 0.033. Combining this with Coupé’s proposed universal information rate, we were able to determine the rate at which Classical Latin would have been spoken by native speakers thousands of years ago: 6.19 syllables per second, with a standard deviation of 0.81. While the variance is relatively large, it’s almost entirely due to differences in speech rate within a language, with the uncertainty in our extrapolation being negligible.

Notably, our results indicate that Classical Latin was spoken at a significantly slower rate than modern Romance languages—Coupé’s data shows a mean value of 7.73 syl/sec for Spanish, for example, and 7.16 syl/sec for Italian. This makes sense from a diachronic perspective, as historical sound changes generally reduced the size of the syllable inventory, decreasing the amount of information provided by each syllable.

This suggests multiple avenues for further research. These methods could be applied directly to other languages for which only written corpora exist; the primary difficulty lies in automatically creating phonemic representations from ambiguous writing systems, as it’s unclear how well Winge’s methods can be generalized. With Latin in particular, it should also be possible now to calculate the effects of the various sound changes that led to modern Romance languages, and determine how much the speech rate was affected by vowel shifts, epenthesis, coda deletion, and so on. Putting these changes into their historical context, this would allow us to predict speech rate across time and see when and why it changed over the centuries.

References

- Allen, W. S. (1978). *Vox Latina: A guide to the pronunciation of classical Latin* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511620348>
- Coupé, C., Oh, Y. M., Dediu, D., & Pellegrino, F. (2019). Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances*, 5(9). <https://doi.org/10.1126/sciadv.aaw2594>
- Johnson, K. P., Burns, P., Stewart, J., & Cook, T. (2014–2020). CLTK: The classical language toolkit. <https://github.com/cltk/cltk>
- Oh, Y. M. (2015). *Linguistic complexity and information: Quantitative approaches* (Doctoral dissertation). University of Lyon. http://www.ddl.cnrs.fr/fulltext/Yoonmi/Oh_2015_1.pdf
- Packard Humanities Institute. (1991). PHI Latin texts. <https://latin.packhum.org/>
- Winge, J. (2015). *Automatic annotation of Latin vowel length*. Uppsala University. <https://cl.lingfil.uu.se/exarb/arch/winge2015.pdf>

Grammatical encoding of agency in middle contexts in Ibero-Romance

Middle constructions in Ibero-Romance show imperfective tenses, agreement between the verb and the preverbal theme DP, and the impossibility of introducing an agent by means of a *by*-phrase, although these sentences denote the participation of a generic external argument in the event. Additionally, while the reflexive clitic *se* is an obligatory feature in these contexts in Ibero-Romance (1a.b), it appears to be optional in Asturian (1c) (ALLA 2001).

- (1) a. Este pan *(se) desmiga fácilmente (*por Juan). *Spanish*
 this bread REFL crumbles easily by Juan
 b. Aquest pa *(s)'esmolla fàcilment (*per Joan). *Catalan*
 this bread REFL-crumbles easily by Joan
 c. Esti pan esmigáya(se) fácil (*por Xuan). *Asturian*
 this bread crumbles.REFL easily by Xuan
 ‘This bread crumbles easily.’

A careful examination of the data shows the apparent optionality of the reflexive in Asturian is rather systematic. For instance, these sentences allow the insertion of a purpose clause controlled by the implicit external argument in the structure (Bhatt & Pancheva 2006); interestingly, such clause forces the presence of the reflexive clitic in Asturian.

- (2) a. Esti pan esmigáya*(se) fácil *pa empanar cachopos*. *Asturian*
 this bread crumbles.REFL easy for bread steaks
 b. Este pan *(se) desmiga fácilmente *para empanar cachopos*. *Spanish*
 this bread REFL crumbles easily for bread steaks
 ‘This bread is easy to crumble to bread steaks.’

Moreover, middle sentences containing change of state/location predicates allow a non-agentive reading, which may be enhanced by means of a PP like *por sí mismo/mesmu* (‘by itself’). Crucially, this PP is only licensed in the absence of the reflexive clitic in Asturian.

- (3) a. Esti pan esmigáya(*se) fácil *por sí mesmu*. *Asturian*
 this bread crumbles.REFL easy by itself
 b. Este pan *(se) desmiga fácilmente *por sí mismo*. *Spanish*
 this bread REFL crumbles easily by itself
 ‘This bread crumbles easily by itself.’

On the contrary, predicates denoting activities or accomplishments that notionally imply the participation of an agent in the event, necessarily require the reflexive.

- (4) a. Les noveles de misterio lléen*(se) con facilidá. *Asturian*
 the novels of mystery read.REFL with ease
 ‘Mystery novels read easily.’

I propose that the presence of the reflexive in middle constructions in these languages is associated with the projection of a passive Voice head, spelled out by *se*, which encodes the participation of a generic external argument in the event (Kratzer 1996; Schäfer 2008). Thus, I explain that two possible configurations can yield middle sentences containing change-of-state predicates: (i) a generic *se*-passive (5a) containing a passive Voice head and a bieventive structure comprising an activity subevent and a stative one ($v_{DO}+v_{BE}$) (Cuervo 2003); and (ii) a generic

NOT SO INCLUSIVE: The Debate on Inclusive-Writing on French Wikipedia

In recent years, inclusive-writing has been at the center of public debates in France. Often referred to as gender-neutral language, the phenomenon of linguistic inclusivity served as a vehicle for women's emancipation throughout the twentieth century (e.g. the feminization of occupational terminology), and lately, as a medium to reflect progressive social changes in the language (e.g. the inclusion of non-binary pronouns) (Kosnick, 2019). At the same time, members of the Académie française strongly oppose the implementation of inclusive-writing, demonizing it as a mortal danger to the French language, in particular evoking its structural disruptiveness to well-defined norms. Proponents of inclusive-writing argue that language is built on social codes and therefore ought to reflect progressive social change (Viennot, 2020). In 2019, French Wikipedia editors created a poll to ascertain whether inclusive-writing practices are to be featured on the website. The representation of inclusive forms on a platform with 700 million views per year could bear a remarkable effect on the perception of egalitarian language use among French speakers. In this paper, I investigate language ideologies in the 2019 Wikipedia poll on inclusive-writing by applying corpus linguistic approaches and statistical textual analysis in order to examine purist attitudes, or the lack thereof, toward gender-neutral language among influential Wikipedia contributors.

Data is scraped from the poll's webpage and analyzed in R. The corpus consists of 2011 votes spread over 6 questions (Table 1) and cast by more than 310 editors having contributed at least 50 articles to French Wikipedia. The poll was open between December 1, 2019 and January 5, 2020. Votes are distributed on a Likert-scale from *strongly support* to *strongly against*. Figure 1 demonstrates that French-speaking editors overwhelmingly rejected all six proposed inclusive-writing techniques. *Question 1* (epicene terms) and *Question 3* (rare feminine forms) received the most positive reactions, which may be explained by their relatively traditional linguistic forms. On the other hand, morphological neologisms in *Question 5* (masculine-feminine portmanteaus) and *Question 6* (non-binary terms) are considered disruptive to current French norms imposed by the Académie française. Interestingly, while *Question 2* and, to a certain extent, *Question 4* both conform to established rules: editors oppose them because they would make an article unnecessarily lengthy and heavy ("*alourdissement inutile*"). On Figure 2, preliminary linear regression models also shed light on a correlation between time and votes: over one month, editors cast increasingly more positive votes on each issue, with the strongest increase observed for *Question 2* (double inflection), *Question 4* (proximity or majority agreement) and *Question 5* (masculine-feminine portmanteaus). Further ordinal logistic regression models are applied to the data in order to examine how individual editors score across all six questions.

The second part of the analysis proposes to combine quantitative computational approaches and qualitative textual analysis to investigate topics related to each inclusive-writing form. From a methodological perspective, the efficacy of topic models using Latent Dirichlet Allocation (Blei et al., 2003) is compared with the more recent powerful tool of universal sentence encoder (Cer et al., 2018). Initial results suggest two major recurring topics both in favor of and opposed to the proposed changes: linguistic and social rationales. The linguistic topics include comments on language norms ("*WP n'a pas à inventer de nouvelles règles de grammaire*"), while the social arguments incorporate the need to reflect progressive changes in French society ("*la militance serait AUSSI de garder les anciennes formules et wiki ne doit pas s'associer à la discrimination*").

Results of the analysis thus corroborate previous scholarship that reveals strong purist attitudes among French speakers signaling that progressive social behavior does not necessarily translate into support of or advocacy for inclusive language use on Wikipedia.

Selected references:

- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3.
- Cer, D., Yang, Y., Kong, S., Hua, N., Limtiaco, N., St. John, R., Constant, N., Guajardo-Céspedes, M., Yuan, S., Tar, C., Sung, Y.-H., Strope, B., & Kurzweil, R. (2018). *Universal Sentence Encoder (Version 2)* [Computer software]. Google Research; arXiv:1803.11175v2.
- Kosnick, K. (2019). The everyday poetics of gender-inclusive French: Strategies for navigating the linguistic landscape. *Modern & Contemporary France*, 27(2), 147–161. <https://doi.org/10.1080/09639489.2019.1588869>
- Viennot, É. (2020). Langage égalitaire: Vers une rationalisation des procédés et des approches. *Cahiers d'histoire. Revue d'histoire Critique*, 146, 149–160. <https://doi.org/10.4000/chrhc.14838>

Table 1. Six questions targeting gender-inclusive language use on Wikipedia

Question	Objective	Examples
Question 1	Epicene terms	<i>le personnel, les novices</i>
Question 2	Double inflection	<i>les contributeurs et les contributrices</i>
Question 3	Rare feminine forms	<i>la cheffe / la chève</i>
Question 4	Proximity or majority agreement	<i>l'étudiante et l'étudiant inscrits</i>
Question 5	Masculine-feminine portmanteaus	<i>salarié.e.s, salarié-es, salarié(e)s</i>
Question 6	Non-binary terms	<i>iel, celleux, contributeurice</i>

Figure 1. Boxplot of the votes shows that French-speaking Wikipedia editors overwhelmingly reject the use of gender-inclusive language. *Question 5* (masculine-feminine portmanteaus) and *Question 6* (non-binary terms) are voted the most strongly against.

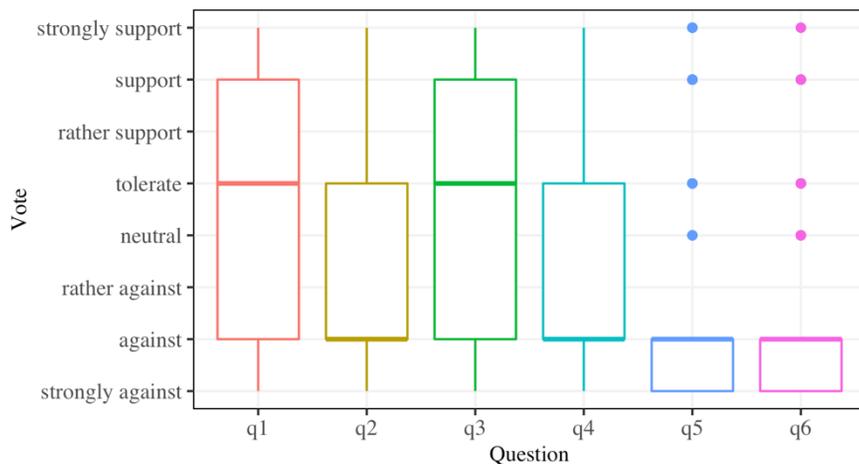


Figure 2. Linear regression estimates reveal that during the one month that the poll was open, every question received slightly more favorable votes over time. This is particularly true for *Question 2* (double inflection), *Question 4* (Proximity or majority agreement) and *Question 5* (masculine-feminine portmanteaus).

