Catalan nativization patterns in the light of Weighted Scalar Constraints*

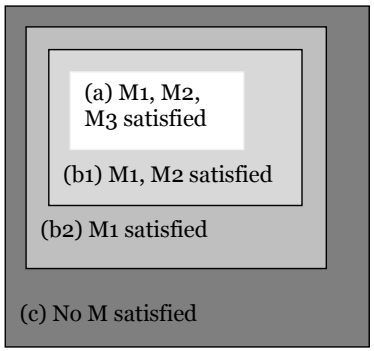
Clàudia Pons-Moll[1] & Francesc Torres-Tamarit[2]
Universitat de Barcelona[1], CNRS/Université Paris 8[2]

Going Romance 2018
Utrecht Institute of Linguistics – OTS
11-14 December 2018

1. Introduction and goals

- Loanwords usually pattern differently than native words with respect to markedness.
- Loanwords usually comply with only a subset of the markedness constraints satisfied by native words, or, in many cases, with none of these markedness constraints.
- This situation brings about a nested core-periphery structure of the lexicon, with different strata (Itô & Mester 1999, 2008 / 2009): (a) the *core stratum*, in which loanwords behave as native words and satisfy all markedness constraints (*nativized loanwords*); (b) *the intermediate stratum(s)*, in which loanwords satisfy only a subset of the markedness constraints active in the core strata (*partially nativized loanwords*); (c) the *peripheral stratum*, in which loanwords do not satisfy any of the markedness constraints active in the previous strata (*non-nativized loanwords*).

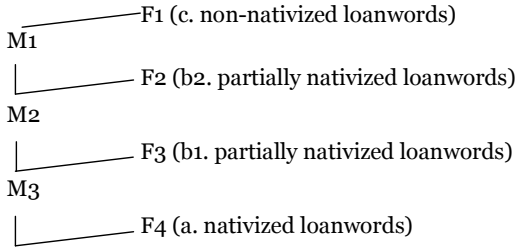
(1) Core-periphery structure of the lexicon (Itô & Mester 1999, 2008 / 2009)



- The differences according to each of these strata are explained by the variable position of a block of faithfulness constraints F1, F2, F3..., to which lexical items in each stratum are indexed, with respect to a language-particular fixed hierarchy of markedness constraints (M1 >> M2 >> M3).

* We are grateful to Vlad Martin-Diaconescu, who helped us with the the treatment of the data in Excel and Python (§ 3), and who did the calculations of the weights depending on the scaling factor variable in Excel (§ 4). This reasearch is supported by the Catalan Government (2014SGR918) and by the project FFI2016-76245-C3-3-P.

(2) Differences across strata



- Such a structure gives rise to implicational patterns in the adaptation of loanwords.

  "Structures [...] are built out of a network of implicational relations involving lexical items and phonological constraints of the following kind: items that are subject to constraint *A* are also always subject to constraint *B*, but not all items subject to *B* are also subject to *A*." (Itô & Mester 2008 / 2009: p. 554).

- In this talk we deal with two cases involving implicational patterns of this sort in the adaptation of loanwords in Catalan (see Pons-Moll 2015).
- The purpose of the talk is to present the results of two surveys supporting quantitatively these kinds of patterns and to attempt a formalization of them under the *Weighted Scalar Constraints* version of *Harmonic Grammar*, following the recent proposals by Hsu & Jesney (2017, 2018).

2. Data

2.1. Word-final posttonic /n/ deletion (ND) and vowel reduction (VR)

Word-final posttonic /n/ deletion and vowel reduction are general processes in the native lexicon of Catalan.

(3) ND (Mascaró 1976, Bonet & Lloret 1998)

pla[n]s ~ pla[n]íssim ~ pla[∅]	'flat PL.' ~ flat SUPERL.' ~ 'flat SG.'
cosi[n]s ~ cosi[n]et ~ cosí[∅]	'cousin PL.' ~ 'cousin DIM." ~ 'cousin SG.'

(4) VR (Mascaró 1976, Bonet & Lloret 1998)

c[á]sa ~ c[ə]seta	'house SG.' ~ 'house DIM.'
t[ɛ́]rra ~ t[ə]rrestre	'earth SG.' ~ 'terrestrial'
f[é]ra ~ f[ə]rós	'beast SG.' ~ 'fierce'
p[ɔ́]rta ~ p[u]rtal	'door SG.' ~ 'hallway'
p[ó]ma~ p[u]mera	'apple SG.' ~ 'apple tree'

2.2. Underapplication of ND and VR

- These two processes, though, tend to underapply in loanwords.

(5) Underapplication of ND in loanwords (Pons-Moll *et. al* 2018)

diva[n]	taliba[n]	Pakista[n]
futo[n]	catipe[n]	Afganista[n]
canca[n]	mato[n]	Suda[n]
xama[n]	canto[n]	Vuitto[n]
catamara[n]	pasto[n]	Nissa[n]

(6) Underapplication of VR in loanwords (Mascaró 2002, Cabré 2009, Pons-Moll 2012, Pons-Moll *et. al* 2018)

cutr[e]	Goog[e]l	m[o]jit[o]
gor[e]	pilat[e]s	pest[o]
fly[e]r	típ[e]x	jud[o]
gadg[e]t	clín[e]x	sad[o]
hípst[e]r	ram[e]n	cron[o]
màst[e]r	youtub[e]r	tac[o]
cút[e]r	t[e]mpura	parkins[o]n
blíst[e]r	s[e]rotonina	gastr[o]bar
Twitt[e]r	c[o]ntàin[e]r	c[o]llage

Interestingly enough, loans susceptible to undergo both processes show a consistent behavior in which underapplication of both processes is the most common solution (*t*[o]*b*[o]*ga*[n]), followed closely by just underapplication of ND (*t*[u]*b*[u]*ga*[n]), followed by far by application of both processes (*t*[u]*b*[u]*ga*[ø]), and in which underapplication of VR and application of ND (**t*[o]*b*[o]*ga*[ø]) is **unattested**.

(7) Implicational relations between ND and VR, and tendencies

Most common	Underapplication of ND and VR	*t*[o]*b*[o]*ga*[n]	PatA1
Less common	Underapplication of ND and application of VR	*t*[u]*b*[u]*ga*[n]	PatA2
Least common	Normal application of ND and VR	*t*[u]*b*[u]*ga*[ø]	PatA3
Unattested (impossible nativization)	Underapplication of VR but not of ND	**t*[o]*b*[o]*ga*[ø]	PatA4

2.2. Mid vowel laxing (VL) and VR

In Catalan, there is a notable tendency to prefer [–ATR] mid vowels in stressed position ([ɛ́], [ɔ́]), over the [+ATR] counterparts ([é], [ó]), which is manifested through a wider distribution of the former across the Catalan lexicon (Mascaró 2002) and in loanword adaptation.

(8) Preference for [–ATR] mid vowels in loanword adaptation (Mascaró 2002, Pons-Moll et al. 2018)

top t[ɛ́]n	postd[ɔ́]c
tr[ɛ́]ndy	p[ɔ́]st-it
tr[ɛ́]kking	l[ɔ́]ft
s[ɛ́]lfie	Power P[ɔ́]int
l[ɛ́]ggings	
gill[ɛ́]tte	
r[ɛ́]iki	

This tendency, which we interpret as a process of sonority-driven vowel laxing (VL) in stressed position of an underlying /e/ or /o/ also interacts with VR in loanwords (see Pons-Moll 2015).

In these cases, the most common solution is underapplication of both processes ([é]*ur*[o], *p*[ó]*st*[e]*r*),[2] followed by far by the application of both processes ([ɛ́]*ur*[u], *p*[ɔ́]*st*[ə]*r*); on the contrary, mixed patterns with underapplication of VL and application of VR ([é]*ur*[u], *p*[ó]*st*[ə]*r*), or with application of VL and underapplication of VR ([ɛ́]*ur*[o], *p*[ɔ́]*st*[e]*r*) are generally avoided, although they can be found sporadically in some specific words (Cabré 2009).

(9) Implicational relations between VL and VR, and tendencies

Most common	Underapplication of VL and VR	[é]*ur*[o], *p*[ó]*st*[e]*r*	PatB1
Less common	Application of VL and VR	[ɛ́]*ur*[u], *p*[ɔ́]*st*[ə]*r*	PatB2
Very infrequent	Application of VR and underapplication of VL	?[é]*ur*[u], ?*p*[ó]*st*[ə]*r*	PatB3
Even more infrequent	Application of VL and underapplication of VR	?[ɛ́]*ur*[o], ?*p*[ɔ́]*st*[e]*r*	PatB4

3. Experimental survey

3.1. Picture-naming production task

- 16 loanwords with word-final posttonic /n/ + unstressed mid vowels (***t*o*b*o*ga*n**)
- 6 loanwords containing a stressed mid vowel + unstressed mid vowels (***eur*o, *p*ò*st*er**)
- 31 Barcelona Catalan speakers aged 18-23 during the period 2017-2018
- Students of the BA degree Comunicació i Indústries Culturals

3.2. Judgment test inquiring the naturality of the four possible patterns

- Presented in an audio file via a Google form available on Internet

[2] See Bonet *et al.* (2007) and Cabré (2009) for an alternative interpretation of this pattern based on vowel harmony.

- The same 16+6 loanwords (22 x 4 patterns = 88 items)
- Patterns valued along a Likert scale of 1-5 (very unnatural, quite unnatural, natural enough, quite natural, very natural).

Both tests were fulfilled with loanwords with just one of the relevant structures (e.g. *diva**n***, *màst**er***, etc.) and with a 50% of distractors, and were presented in a randomized way.

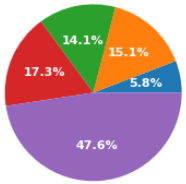
(10) Results of the picture-naming production task

Patterns A	% of answers	Patterns B	% of answers
a. PatA1 t[o]b[o]ga[n]	65,2%	a. PatB1 p[ó]st[e]r	98,9%
b. PatA2 t[u]b[u]ga[n]	25%	b. PatB2 p[ɔ́]st[ə]r	1,1%
c. PatA3 t[u]b[u]ga[∅]	9,8%	c. PatB3 p[ó]st[ə]r	0%
d. PatA4 t[o]b[o]ga[∅]	0%	d. PatB4 p[ɔ́]st[e]r	0%

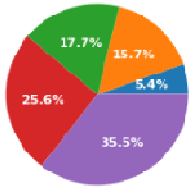
(11) Results of the judgment tests

a. *Judgment test. Patterns A*

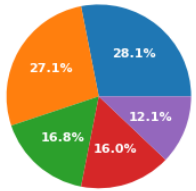
PatA1 t[o]b[o]ga[n]



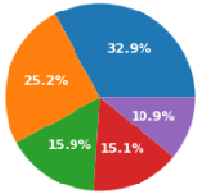
PatA2 t[u]b[u]ga[n]

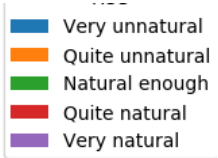


PatA3 t[u]b[u]ga[∅]



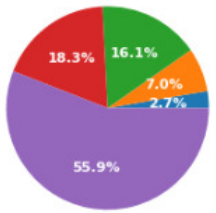
PatA4 t[o]b[o]ga[∅]



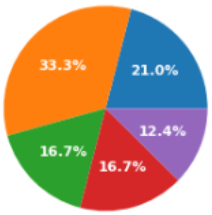


b. *Judgment test. Patterns B*

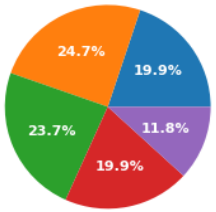
PatB1 p[ó]st[e]r



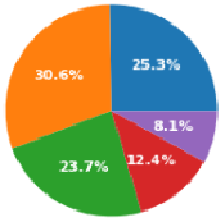
PatB2 p[ɔ́]st[ə]r

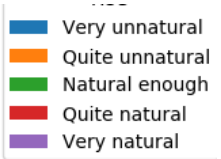


PatB3 p[ó]st[ə]r



PatB4 p[ɔ́]st[e]r





About these results, which generally fit the gradations exposed in §2, we should comment the following:

(a) Mixed patterns B3 and B4 received a high score for the neutral category "natural enough" (23,7% in both cases), which reveals the hesitation of speakers in front of this type of realizations.

(b) We attribute the low scores for PatA3 *t*[u]*b*[u]*ga*[ø] and PatB2 [ɛ́]*ur*[u] (*i.e.* nativized patterns), both in the production and in the judgment tests, to the age of the inquired speakers.

(c) Note, finally, that no significant differences were detected in patterns A with respect to the quality of the unstressed vowels (*i.e.* low /a/, as in *or*[a]*ngutan*, *vs.* mid /e/, /o/, as in [o]*rangutan*).

4. Analysis with Weighted Scalar constraints

- Implicational patterns of the sort exemplified in the previous sections are predicted to exist in a model with weighted constraints as in Harmonic Grammar (Smolensky & Legendre 2006).

- In the analysis presented here, which follows Hsu & Jesney (2017, 2018), faithfulness violations are scaled according to the definition in (12).

(12) *Scaled Faithfulness Weigthed Constraints* (Hsu & Jesney 2018: 255)

> "Given a basic constraint weight *w*, and a scaling factor *s* corresponding to distance from the core, for any input that is not realized faithfully in the output, assign a weighted violation score of *w* x *s*."

4.1. Patterns A (ND & VR)

- For the cases belonging to Pattern A, we assume a triple lexical strata in the Catalan grammar (13; 14): (*a*) the core one (for those speakers [and loans] with application of VR and ND: *t*[u]*b*[u]*ga*[ø]); (*b*) the intermediate one (for those speakers [and loans] with just application of VR: *t*[u]*b*[u]*ga*[n]); (c) the peripheral one (for those speakers [and loans] with underapplication of both VR and ND: *t*[o]*b*[o]*ga*[n]).

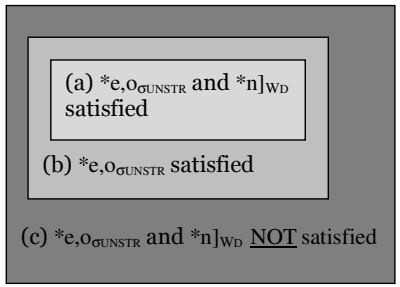
- The two M constraints involved are *e,o$_{\sigma UNSTR}$ and *n]$_{WD}$, which receive respectively a stable weight of 5.5 and 2.5 across all three possible strata.

- Scaled faithfulness ensures that the faithfulness weight values increase from the core stratum (in which *s* = 1), towards the intermediate stratum (which starts with *s* = 1.8), until reaching the **peripheral stratum** (which starts with *s* = 2.7 and which covers the **largest interval**).

- **Faithfulness values** acquire, thus, a **higher relevance** the closer to the peripheral strata.

- Given the constraint weights, no scaling factor can yield the impossible nativization PatA4 **t*[o]*b*[o]*ga*[ø] (see the strata cross overpoints in 17).

(13) Core-periphery grammar



(14) HG with *scalar weighted constraints* tableau for Patterns A

i. /tobogan/	*e,o$_{\sigma UNSTR}$ w = 5.5	*n]$_{WD}$ w = 2.5	Ident-V$_{UNSTR}$ w = 2	Max-IO w = 1.5	H	*Scaling factor for F*	*Strata*
a. [toβoɣán]	–1	–1			–8	1	*Core stratum*
b. [tuβuɣán]		–1	–1		–4.5		
☞c. [tuβuɣáØ]			–1	–1	–3.5		
d. [toβoɣáØ]	–1			–1	–7		
ii. /tobogan/	*e,o$_{\sigma UNSTR}$ w = 5.5	*n]$_{WD}$ w = 2.5	Ident-V$_{UNSTR}$ w = 2	Max-IO w = 1.5	H	*Scaling factor for F*	
a. [toβoɣán]	–1	–1			–8	1.8	*Intermediate stratum*
☞ b. [tuβuɣán]		–1	–1		–6.1		
c. [tuβuɣáØ]			–1	–1	–6.3		
d. [toβoɣáØ]	–1			–1	–8.2		
iii. /tobogan/	*e,o$_{\sigma UNSTR}$ w = 5.5	*n]$_{WD}$ w = 2.5	Ident-V$_{UNSTR}$ w = 2	Max-IO w = 1.5	H	*Scaling factor for F*	
☞ a. [toβoɣán]	–1	–1			–8	2.7	*Peripheral stratum*
b. [tuβuɣán]		–1	–1		–7.9		
c. [tuβuɣáØ]			–1	–1	–9.45		
d. [toβoɣáØ]	–1		–1		–9.55		

4.2. Patterns B (VL & VR)

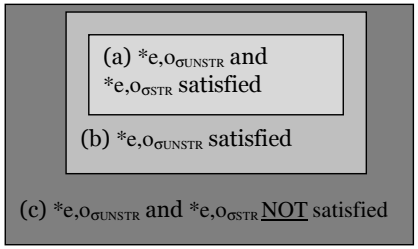
- For pattern B, we assume also a triple lexical strata (15, 16): *a*) the core one (for speakers [and loans] with application of VR and VL: [ɛ́]*ur*[u], *p*[ɔ́]*st*[ə]*r*); an intermediate one (for speakers [and loans] with application of VR but underapplication of VL: *p*[ó]*st*[ə]*r*), and *c*) the peripheral one (for speakers [and loans] with underapplication of both VR and VL: [é]*ur*[o], *p*[ó]*st*[e]*r*).

- The two markedness constraints involved are *e,o$_{\sigma UNSTR}$ and *e,o$_{\sigma STR}$, which receive both a stable weight of 5.5 across all possible strata.

- In this case, the transition *scaling factors* from one strata to the other are 1, 2.3 and 2.8.

- Given the constraint weights, no scaling factor can yield the nativization PastB4 (*p[ɔ́]st[e]r), and a very small scaling factor for the intermediate stratum with PatB3 p[ó]st[ə]r) is predicted.

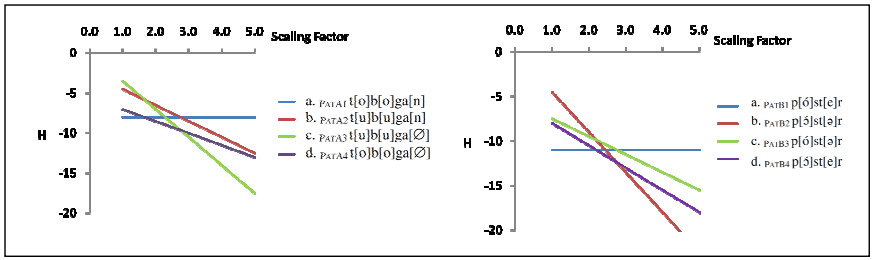
(15) Core-periphery grammar



(16) HG with *scalar weighted constraints* tableau for Patterns B

i. /poster/	*e,o$_{\sigma UNSTR}$ w = 5.5	*e,o$_{\sigma STR}$ w = 5.5	IDENT-V$_{STR}$ w = 2.5	IDENT-V$_{UNSTR}$ w = 2	H	*Scaling factor for F*	*Strata*
a. [póster]	–1	–1			–11	1	*Core stratum*
☞ b. [póstər]			–1	–1	–4.5		
c. [póster]	–1		–1		–8		
d. [póstər]		–1		–1	–7.5		
ii. /poster/	*e,o$_{\sigma UNSTR}$ w = 5.5	*e,o$_{\sigma STR}$ w = 5.5	IDENT-V$_{STR}$ w = 2.5	IDENT-V$_{UNSTR}$ w = 2		*Scaling factor for F*	
a. [póster]	–1	–1			–11	2.3	*Intermediate stratum*
b. [póstər]			–1	–1	–10.35		
c. [póster]	–1		–1		–11.25		
☞ d. [póstər]		–1		–1	–10.1		
iii. /poster/	*e,o$_{\sigma UNSTR}$ w = 5.5	*e,o$_{\sigma STR}$ w = 5.5	IDENT-V$_{STR}$ w = 2.5	IDENT-V$_{UNSTR}$ w = 2	H	*Scaling factor for F*	
☞ a. [póster]	–1	–1			–11	2.8	*Peripheral stratum*
b. [póstər]			–1	–1	–12.6		
c. [póster]	–1		–1		–12.5		
d. [póstər]		–1		–1	–11.1		

(17) Strata cross overpoints for Patterns A and Patterns B



5. Conclusions

- In this paper we have explored phonological nativization patterns in Catalan loanwords, and we have shown, on the basis of a production and a judgment test, that the three processes under scrutiny (word-final *–n* deletion [ND], vowel reduction of unstressed mid-vowels [VR], and vowel laxing of stressed mid-vowels [VL]) interact in an asymmetrical way.

- Loans susceptible to undergo ND and VR show a consistent behavior, in which underapplication of both processes is the most common solution (*t*[o]*b*[o]*ga*[n]), followed closely by just underapplication of ND (*t*[u]*b*[u]*ga*[n]), followed by far by application of both processes (*t*[u]*b*[u]*ga*[∅]), and in which underapplication of VR and application of ND (**t*[o]*b*[o]*ga*[∅]) is unattested.

- Loans susceptible to undergo VR and VL also show a consistent behavior, in which the most common solution is underapplication of both processes ([é]*ur*[o], *p*[ó]*st*[e]*r*), followed by far by the application of both processes ([ɛ́]*ur*[u], *p*[ɔ́]*st*[ə]*r*), and which mixed patterns with underapplication of VL and application of VR ([é]*ur*[u], *p*[ó]*st*[ə]*r*), or with application of VL and underapplication of VR ([ɛ́]*ur*[o], *p*[ɔ́]*st*[e]*r*) are generally avoided, although the judgment test indicates that the third pattern is slightly more tolerated than the fourth one.

- We argue that these asymmetrical interactions can be straightforwardly formalized resorting to Harmonic Grammar with Scalar Weighted Constraints (Hsu & Jesney 2017, 2018), in which faithfulness constraints acquire an increasing relevance from the core to the peripheral strata and in which the scaling factor intervals for each stratum abstractly reproduce the most frequent and the least frequent patterns.

6. References and bibliography

Bonet, Eulàlia; Maria-Rosa Lloret (1998). *Fonologia catalana*. Barcelona: Ariel.

Bonet, Eulàlia; Lloret, Maria-Rosa; Mascaró, Joan (2007). Domain and directionality in Catalan ATR harmony. Poster presented at the Workshop Harmony in the languages of the Mediterranean de la 4th Old World Conference on Phonology (OCP-4), Rhodes.

Cabré, Teresa (2002). Altres sistemes de formació de mots. In: Joan Solà, Maria-Rosa Lloret, Joan Mascaró & Manuel Pérez Saldanya (eds), *Gramàtica del català contemporani*, 89-123. Barcelona: Empúries.

Cabré, Teresa (2006 / 2009). El sistema vocàlic del català central i l'adaptació de manlleus. En: Faluba, Kalman; Szijj, Ildikó (ed.). *Actes del Catorzè Col·loqui Internacional de Llengua i Literatura catalanes: Universitat Eötvös Loránd de Budapest, 4-9 de setembre de 2006*, vol. 3. p. 111-120. Barcelona: Publicacions de l'Abadia de Montserrat.

Cabré, Teresa (2009). Vowel reduction and vowel harmony in Eastern Catalan loanword phonology. In Marina Vigário, Sónia Frota & Maria Joao Freitas (eds). *Phonetics and Phonology. Interactions and interrelations*, 267-286. Amsterdam & Philadelphia: John Benjamins.

Holden, Kyril (1976). Assimilation rates of borrowings and phonological productivity. *Language*, vol. 52, p. 131-147.

Hsu, Brian & Karen Jesney (2017). Loanword adaptation in Québec French: evidence for

weighted scalar constraints. In A. Kaplan et al. (eds), *Proceedings fo the 34th Meeting of the West Coast Conference on Formal Linguistics* (WCCFL 34), 249-258. Somerville, MA: Cascadilla.

Hsu, Brian & Karen Jesney (2018). Weighted scalar constraints capture the typology of loanword adaptation. In: G. Gallagher, M. Gouskova, S. Heng Yin (eds), *Proceedings of the 2017 Annual Meeting on Phonology*, 1-11. Linguistic Society of America.

Itô, Junko; Mester, Armin (1995). Japanese phonology. In: John Goldsmith (ed.). *The handbook of phonological theory*, p. 817-838. Cambridge, MA/Oxford: Blackwell.

Itô, Junko; Mester, Armin (1999). The phonological lexicon. In: Natsuko Tsujimura (ed.). *The handbook of Japanese linguistics*, p. 62-100. [Also in John J. McCarthy (ed.) (2008 / 2009). *Optimality Theory: a Reader*. Wiley Blackwell.]

Itô, Junko & Armin Mester (2008). Lexical classes in phonology. In: S. Miyagama & M. Saito (eds), *The Oxford Handbook of Japanese Linguistics*, 84-106. Oxford: Oxford University Press.

Kiparsky, Paul (1971). Historical linguistics. En: Dingwall, William O. (ed.). A survey of linguistic science. *Linguistics Program University of Maryland*, College Park, Md. p. 576-649.

Kiparsky, Paul (1973). Productivity in Phonology. En: Kenstowicz, Michael; Kisseberth, Charles (ed.). *Issues in Phonological Theory*, p. 69-76. The Hague: Mouton.

Mascaró, Joan (1976). Catalan phonology and the phonological cycle. Ph.D. dissertation, MIT. Published by Indiana University Linguistics Club, Bloomington, 1978.

Mascaró, Joan (2002). El sistema vocàlic. Reducció vocàlica. In J. Solà, M.-R. Lloret, J. Mascaró & M. Pérez Saldanya (eds), *Gramàtica del català contemporani*, 89-123. Barcelona: Empúries.

Pons-Moll, Clàudia ([2009, 2010] 2011). Underapplication, lexical exceptions, loanword phonology and foreign language acquisition. Which is their lowest common denominator. Paper presented at the Phonetics and Phonology in Iberia (PaPI 2009), U de Las Palmas de Gran Canaria; the Michigan Linguistic Society Conference, U of Michigan-Flint; Manchester Phonology Meeting 2011, U of Manchester.

Pons-Moll, Clàudia (2012). Loanword Phonology, Lexical Exceptions, Morphologically Driven Underapplication, and the Nature of Positionally Biased Constraints. *Catalan Journal of Linguistics*, vol. 11, p. 127-166.

Pons-Moll, Clàudia (2015). Comentaris a Joan Mascaró Regularitat i excepcions en fonologia: les reduccions vocàliques, de Joan Mascaró. p. 71-101. Barcelona: Publicacions i Edicions de la Universitat de Barcelona.

Pons-Moll, Clàudia; Torres-Tamarit, Francesc; Martin-Diaconescu, Vlad (2018). Nativitzacions (im)possibles en la fonologia del català. [Manuscript.]

Smith, Jennifer L. (2018). Impossible-nativization effects and productivity in loanword phonology. Invited talk presented at the OCP 15. University College London.

Smith, Jennifer L.; Pinta, J. (2018). Aggressive core-periphery phonology in Guarani: Implications for stratal faithfulness. The University of North Carolina at Chapel Hill & The Ohio State University. [Manuscript]

Smolensky, Paul & Géraldine Legendre. 2006. T*he harmonic mind: from neural computation to Optimality-Theoretic Grammar*. Cambridge, MA: MIT Press.