Modeling morphological epenthesis

Andrija Petrović
Department of Linguistics, Stony Brook University
andrija.petrovic@stonybrook.edu

University of Arizona DSP Linguistics Colloquium, Feb 18 2021

Classical epenthesis

- Spanish

/stop/ > [estop]

Classical vs. morphological epenthesis

- Northwestern Catalan (Artés 2013)

/l#ma$_{r}$/ > [əlmár] ‘the sea’ (f.)

Classical vs. morphological epenthesis

- Northwestern Catalan (Artés 2013)

/l#ma$_{ɾ}$/	>	[əlmár]	‘the sea’ (f.)
/l#a$_{ɾ}$b$_{ɾ}$e/	>	[larβ$_{ɾ}$e]	‘the tree’ (m.)

Classical vs. morphological epenthesis

- Northwestern Catalan (Artés 2013)

/l#ma$_{ɾ}$/	>	[əlmár]	‘the sea’ (f.)
/l#a$_{ɾ}$b$_{ɾ}$e/	>	[larβ$_{ɾ}$e]	‘the tree’ (m.)
/l#pa$_{ɾ}$e/	>	[lopá$_{ɾ}$e]	‘the father’ (m.)

Classical vs. morphological epenthesis

- Valencian hypocoristics (Artés 2014)

Name	**Central Catalan**	**Valencian**
Agustina	Tina	Tina
Francisco	Cisco	Cisco
Isabel	Bel	Bela
Miquel	Quel	Quelo
Manel	Nel	Nelo
Rosario	Sari	Sari

Classical vs. morphological epenthesis

- Brazilian Portuguese (Bachrach & Wagner 2007)

/korea/ > [koreja] ‘Korea’

Classical vs. morphological epenthesis

- Brazilian Portuguese (Bachrach & Wagner 2007)

/korea/	>	[koreja]	‘Korea’
sofá + inho	>	sofazinho	‘little sofa’
café + al	>	cafezal	‘coffee grove’

Classical vs. morphological epenthesis

- Spanish diminutives (Crowhurst 1992, Norrmann-Vigil 2012)

gato	>	gatito			
bacalao	>	bacalaíto			
cable	>	cablecito	madre	>	madrecita
sofá	>	sofacito	comadre	>	comadrita
camión	>	camioncito			
lápiz	>	lapicito			
pez	>	pececito			

Classical vs. morphological epenthesis

- Classical epenthesis: motivated by phonology
- Non-canonical epenthesis: morphologically influenced

Inserted Elements	classical epenthesis	non-canonical epenthesis
semantic function	-	-
presence is phonologically motivated	+	+
distribution is influenced by morphology	-	+
quality is influenced by morphology	-	+

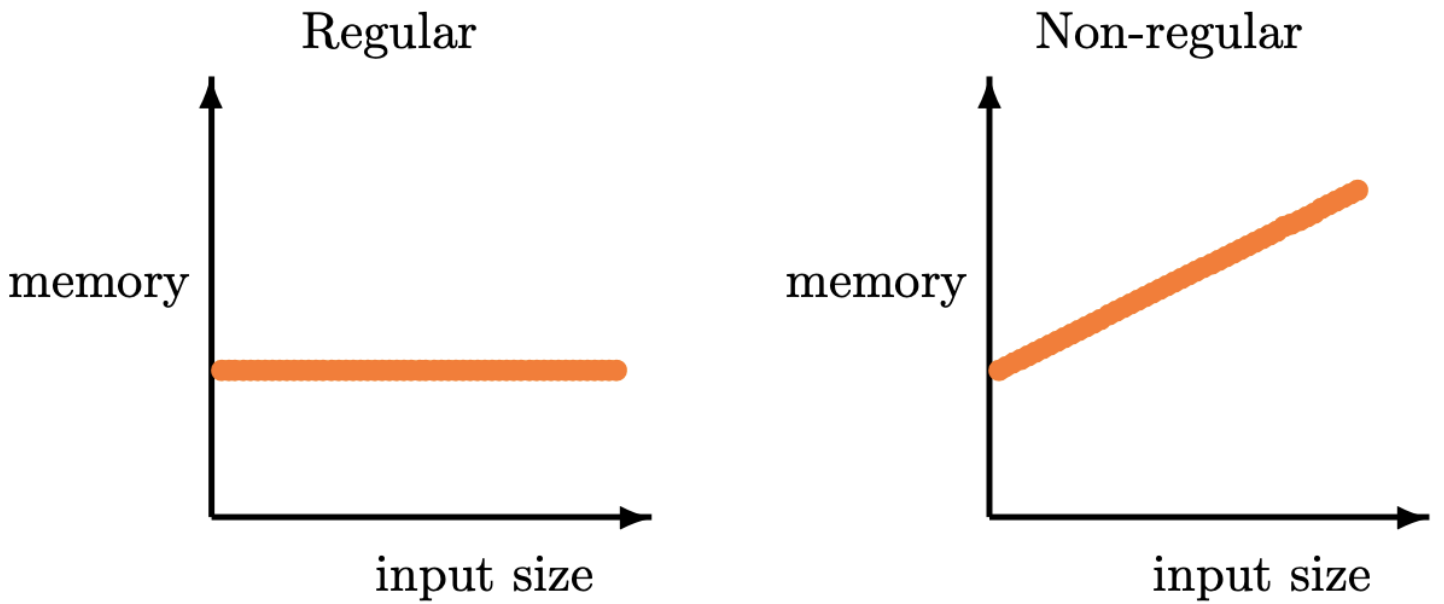
(Repetti, Moradi & Aronoff 2018)

The nature of morphological processes

- morphological processes are regular (Karttunen et al. 1992), most often subregular (Chandlee 2017)

The nature of morphological processes

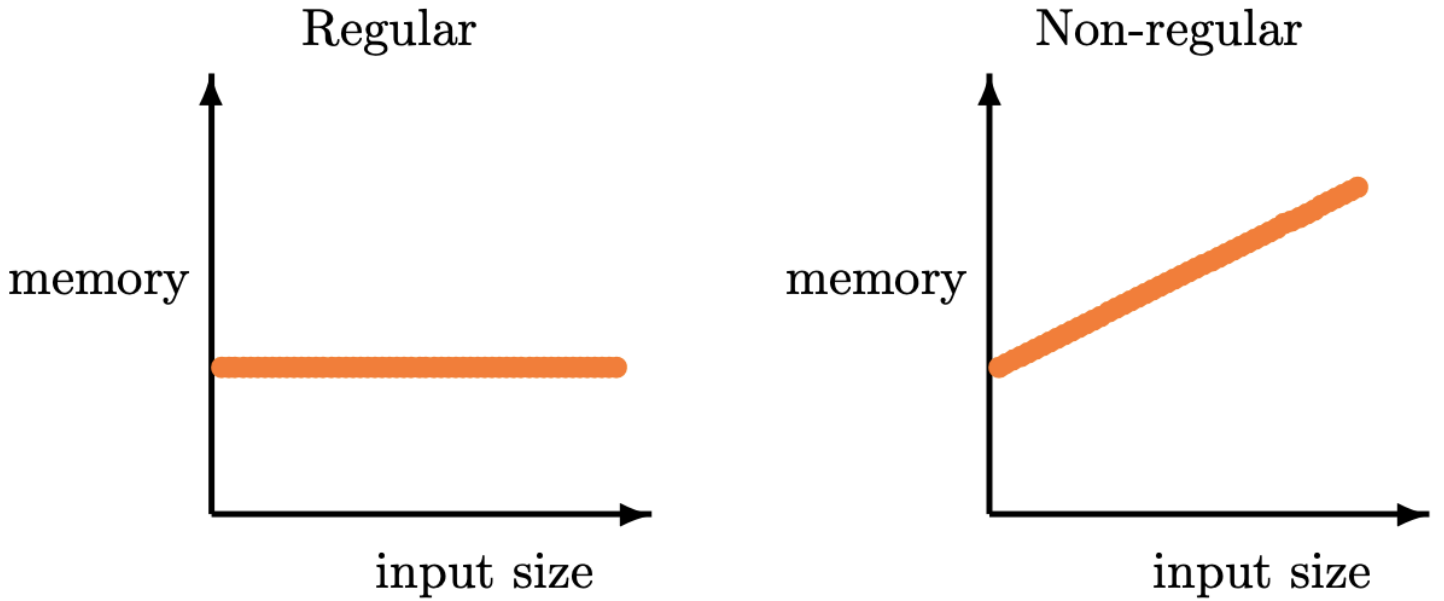
- morphological processes are regular (Karttunen et al. 1992), most often subregular (Chandlee 2017)



(Heinz 2018)

The nature of morphological processes

- A function is regular provided the memory required for the computation is bounded by a constant, regardless of the size of the input.



(Heinz 2018)

The nature of morphological processes

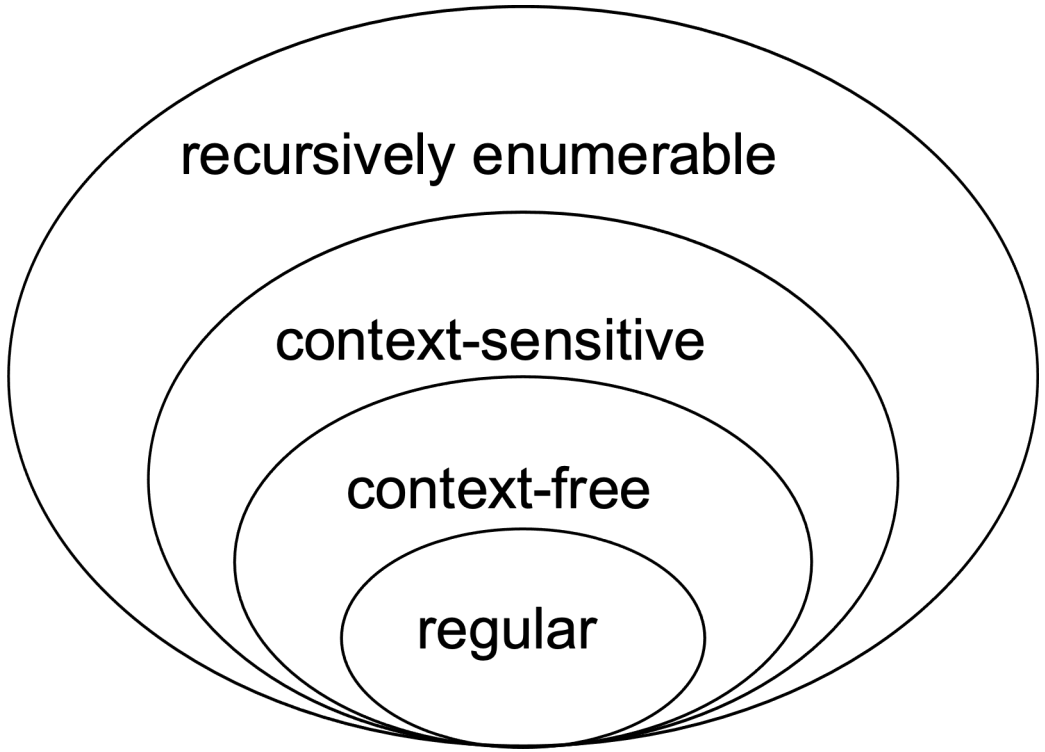
- Regular relations can be specified using (ordered sets of) *rewrite rules* (SPE-style)

In phonology and morphology:

- non-regular patterns are unattested (e.g. ‘majority rule’ in Bakovic 2000)
- human subjects fail to learn non-regular patterns in artificial grammar learning experiments (e.g. Finley 2008)

The nature of morphological processes

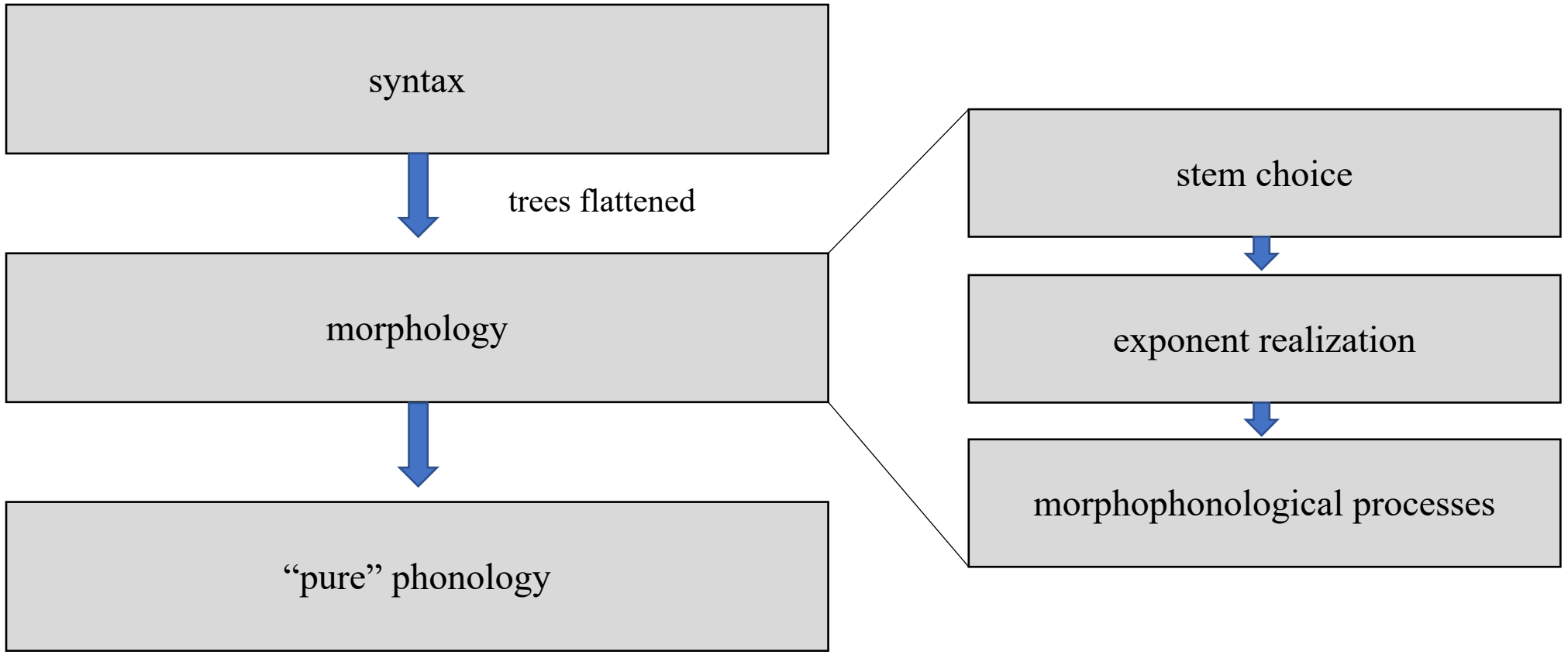
- The Chomsky hierarchy



The nature of morphological processes

- Syntax is context-free ☹
- DM is a tree-based theory, but can be formalized over strings (Ermolaeva & Edmiston 2018)
  - yield function flattens trees, retaining necessary boundary symbols
- In terms of computational power, W&P morphology is equivalent to a collection of regular relations (Karttunen 2003)

Proposed model



Analysis

- *Boolean Monadic Recursive Schemes* (BMRS)
  - incorporates the observation that morphology is regular (Karttunen et al. 1992, Chandlee 2017)
  - unlike FSTs, captures linguistically significant generalizations – we can use representations that are common in linguistics (Chandlee & Jardine 2020)
    - phonological, morphosyntactic features
    - Pāṇini's Principle (= Elsewhere condition) has no obvious finite-state implementation (Karttunen 2003), whereas it is directly captured with BMRS's 'if...then...else' syntax

Analysis

- primitives:
  - *boolean values* ⊤ and ⊥
  - *monadic predicates* $P(t)$ – take a *single* argument *t,* and return ⊤ or ⊥

$\mathbb{V}$ = {a, e, i, o, u, …}
$\mathbb{C}$ = {b, c, d, …}
$\mathbb{M}$ = {[hyp]…}

$\Sigma = \mathbb{V} \cup \mathbb{C} \cup \mathbb{M}$ $\quad\quad$ $\Sigma_{\bowtie} = \Sigma \cup \{\rtimes, \ltimes, +, \#\ldots\}$

Analysis

- primitives:
  - *boolean values* $\top$ and $\bot$
  - *monadic predicates* $P(t)$ – take a *single* argument t, and return $\top$ or $\bot$

$\mathbb{V} = \{\texttt{a}, \texttt{e}, \texttt{i}, \texttt{o}, \texttt{u}, \ldots\}$
$\mathbb{C} = \{\texttt{b}, \texttt{c}, \texttt{d}, \ldots\}$
$\mathbb{M} = \{\texttt{[hyp]}\ldots\}$

$\Sigma = \mathbb{V} \cup \mathbb{C} \cup \mathbb{M} \qquad \Sigma_{\bowtie} = \Sigma \cup \{\rtimes, \ltimes, \texttt{+}, \texttt{\#}\ldots\}$

$\mathcal{I} = \{\texttt{a}(t), \ldots, \texttt{z}(t), \texttt{[hyp]}(t), \ldots, \rtimes(t), \ltimes(t), \texttt{+}(t)\ldots\}$
$O = \{\texttt{a}_o(t), \ldots, \texttt{z}_o(t), \texttt{[hyp]}_o(t), \ldots, \rtimes_o(t), \ltimes_o(t), \texttt{+}_o(t)\ldots\}$

Analysis

$\mathcal{I}$ = {a(*t*), …, z(*t*), [hyp](*t*), ⋊(*t*), ⋉(*t*), +(*t*)…}

	1	2	3	4	5	6	…
	⋊	t	e	l	e	+	…
e(*x*)	⊥	⊥	⊤	⊥	⊤	⊥	
⋊(*x*)	⊤	⊥	⊥	⊥	⊥	⊥	
l(*p*(*x*))	⊥	⊥	⊥	⊥	⊤	⊥	
l(*s*(*s*(*x*)))	⊥	⊤	⊥	⊥	⊥	⊥	

Analysis

- Northwestern Catalan:

`l#pare[m]` → `lopare`

$\circ_{\circ}(x)$ = if `#`(x) then
if `C`(p(x)) then `C`(s(x)) else ⊥
else $\circ(x)$

`out`(x) = if `#`(x) then ⊥ else
if `feat`(x) then ⊥ else ⊤

Analysis

- Valencian:

`(agus)tina[f]+[hyp]` → `tina` `(isa)bel[f]+[hyp]` → `bela`

$a_{\circ}(x)$ = if +(x) then
if `[hyp]`(s(x)) then
if `[f]`(p(x)) then `NOTa`(p(p(x))) else ⊥
else ⊥
else `a`(x)

`out`(x) = if #(x) then ⊥ else
if `feat`(x) then ⊥ else ⊤

Analysis

- Valencian:

`(ma)nel[m]+[hyp]` → `nelo`

$\circ_{\circ}(x)$ = if $+(x)$ then
if `[hyp]`(s(x)) then
if `[m]`(p(x)) then `NOT`$\circ$(p(p(x))) else $\bot$
else $\bot$
else $\circ(x)$

`out`(x) = if `#`(x) then $\bot$ else
if `feat`(x) then $\bot$ else $\top$

Analysis

- Brazilian Portuguese:

`sofa+inho[dim]` → `sofazinho`

`[dim]`(*x*) = if `seg`(*x*) then `[dim]`(s(*x*)) else ⊥

$z_{\circ}$(*x*) = if +(*x*) then
if `[dim]`(s(*x*)) then
if `V`(s(*x*)) then `V`(p(*x*)) else ⊥
else ⊥
else `z`(*x*)

Analysis

- Spanish:

`#sofa#+ito[dim]` → `sofasito` `#pes#+ito[dim]` → `pesesito`

$s_{o}(x)$ = if +(x) then
if o(p(p(x))) then ⊥ else
if a(p(p(x))) then ⊥ else
if C-n,r(p(p(x))) then ⊥ else ⊤
else s(x)

$e_{o}(x)$ = if #(x) then MONO(p(x))

out(x) = if TERM(x) then ⊥ else
if feat(x) then ⊥ else
if symb(x) then ⊥ else ⊤

Conclusions

- Assuming morphological epenthesis enables us to express the generalizations explicitly and overtly
- A means to avoid stipulating listed allomorphy

- BMRSs can directly capture morphological and phonological generalizations, retaining the computationally restrictive nature of such processes
- Intuitive (? ☺), easily implementable, extendable to a wider range of phenomena
- We can account for the data in a very direct and parsimonious way

Team Epenthesis

Lori Repetti

Mark Aronoff

Ji Yea Kim

Veronica Miatto

Sedigheh Moradi

Thank you!

References

- Artés, E. (2013). Morphological Epenthesis in Romance: A Case for Lexical Conservatism. *Poster presented at the 39th Incontro di Grammatica Generativa*
- Artés, E. (2014, March). Valencian hypocoristics: when morphology meets phonology. In *Proceedings of the Annual Meetings on Phonology* (Vol. 1, No. 1).
- Bachrach, A., & Wagner, M. (2007). Syntactically driven cyclicity vs. output-output correspondence: the case of adjunction in diminutive morphology. *U. Penn Working Papers in Linguistics*, *10*(1).
- Bakovic, E. (2000). *Harmony, dominance and control* (Doctoral dissertation).
- Chandlee, J. (2017). Computational locality in morphological maps. *Morphology*, 27:599– 641.
- Chandlee, J., & Jardine, A. (2020). Recursive schemes for phonological analysis. *Unpublished ms., Haverford College and Rutgers University. Available at http://adamjardine.net/files/chandleejardineBMRSms.pdf*
- Crowhurst, M. J. (1992). Diminutives and augmentatives in Mexican Spanish: a prosodic analysis. *Phonology*, *9*(2), 221-253.
- Ermolaeva, M., & Edmiston, D. (2018). Distributed morphology as a regular relation. *Proceedings of the Society for Computation in Linguistics*, *1*(1), 178-181.
- Finley, S. (2008). *Formal and cognitive restrictions on vowel harmony* (Doctoral dissertation).
- Heinz, J. (2018). The computational nature of phonological generalizations. In Hyman, L. and Plank, F., editors, *Phonological Typology,* Phonetics and Phonology, chapter 5, pages 126–195. De Gruyter Mouton.
- Karttunen, L. (2003). Computing with realizational morphology. In *International Conference on Intelligent Text Processing and Computational Linguistics* (pp. 203-214). Springer, Berlin, Heidelberg.
- Karttunen, L., Kaplan, R. M., & Zaenen, A. (1992). Two-level morphology with composition. In *COLING 1992 Volume 1: The 15th International Conference on Computational Linguistics.*
- Norrmann-Vigil, I. (2012). Accounting for variation of diminutive formation in Porteño Spanish. *Mester*, *41*(1).
- Repetti, L., Moradi, S. & Aronoff, M. (2018). *Epenthesis and Morphology in Romance.* Presented at LSRL 48, York University, Toronto.