

CUNY CONFERENCE ON THE WORD IN PHONOLOGY
(14th-15th January, 2010, City University of New York)

PHONOTACTIC CONSTRAINTS
AND WORD DEMARCATION IN ROMANCE

João Veloso (jveloso@letras.upp.pt),

Pedro Tiago Martins (ptsgmartins@gmail.com)

University of Porto (Portugal)

Part of this research is funded by the
CENTRO DE LINGÜÍSTICA DA UNIVERSIDADE DO PORTO
(Fundação para a Ciência e a Tecnologia, Portugal).

Goals of this study

- 1 – To analyze how phonotactic constraints cue word-boundaries in Romance languages.
- 2 – To propose some formalizations of the role of phonotactic constraints as cues for word-demarcation in Romance languages.
- 3 – To underline the theoretical implications of such findings for a correct evaluation of the linguistic units admitted by linguistic descriptions and eventual applications in domains such as automatic processing.

One basic assumption (a.o.):

Phonotactic constraints act in many languages as word-boundary cues.

Early references:

"[...] I would go further, and say that a 'word' is a phonetic entity – that the blank spaces between written words do have phonetic significance." (Jones 1931 : 154)

"The demarcative function (DF) of phonic elements, consists of isolating each semantic unity in the spoken chain. Demarcative signals can be positive when they indicate the limit of a word, and negative when they indicate that at such point a word boundary does not exist [...]." (Anderson 1965 : 185)

Word-Sensitive Phonological Phenomena (Cross-linguistic evidence)

VOWEL HARMONY: One feature is mandatorily shared by all vowels of the word:

--- **Finnish, Hungarian** [FINNO-UGRIC], ...

(Van Der Hulst & Van De Weijer 1995 : 498-499)

J. Veloso, P. T. Martins. 2010. Phonotactic Constraints and Word-Demarcation in Romance. *CUNY Conference on the Word in Phonology*. City University of New York, January 2010. Handout.

STRESS ASSIGNMENT :

Stress saturates the word: no words with more than one stress are admitted:

~(σ'.σ')_ω

Stress falls on a fixed syllable of the word (Kager 1995: 368):

- last: **French** [ROMANCE], ...
- first: **Finnish, Hungarian** [FINNO-UGRIC],

...

- penultimate: **Indonesian** [AUSTRONESIAN] (Soderberg & Olson 2008), **Salasaca Quichua** [QUECHUAN] (Masquiza & Marlett 2008), ...

- penultimate if heavy, otherwise antepenultimate: **Classical Latin** [ITALIC].

Stress falls on a fixed morpheme of the word:

- **Portuguese** [ROMANCE] (Mateus et al. 2003): Nouns: on the last vowel of the root; Verbs: on the Class Marker (Past Tense forms), on the last vowel of the root (Present Tense forms), on the Tense Morpheme (Future Tense forms).
- **Seri** [HOKAN] (Marlett et al. 2005): on the first syllable of the root.

PROHIBITION OF SEGMENT(S)/CLUSTERS IN WORD-INITIAL POSITION:

("negative demarcative signals" – Anderson 1965)

A given segment/cluster can never occur at word beginnings:

[~(#Segment)] ∧ [~(#Cluster)]

Portuguese [ROMANCE]: disallows [ɲ], [ʎ] and [r]

word-initially: ~(#ɲ), ~(#ɲ), ~(#r).

Kabiye ([VOLTIC] (Padayodi 2008): disallows

voiced obstruents word-initially: ~(#VoicedObstr).

Nepali [INDO-ARYAN] (Khatiwada 2009): disallows

word-initial clusters, unless C2 is a rhotic or a glide:

{[C₁C₂] ∧ [C₂ ≠ (R ∨ G)]} → ~(#[C₁C₂]).

Tamil [DRAVIDIAN] (Kean 2004): disallows retroflex consonants word-initially: ~(#Retroflex).

RESTRICTION OF SEGMENT(S)/CLUSTERS TO WORD-INITIAL POSITION

ONLY:

("positive demarcative signals" – Anderson 1965)

A given segment/cluster occurs mandatorily at word beginnings:

(Segment ∨ Cluster) → [(#Segment) ∨ (#Cluster)]

Yakima Sahaptin [PENUITIAN] (Hargus & Beavert

2006): CCV is admitted word-initially only: (CCV) → (#CCV)

Salasaca Quichua [QUECHUAN] (Masquiza &

Marlett 2008): aspirated stops are allowed word-initially only: (AspStop) → (#AspStop)

PROHIBITION OF SEGMENT(S)/CLUSTERS IN WORD-FINAL POSITION:

("negative demarcative signals" – Anderson 1965)

A given segment/cluster can never occur at word endings:

$$[\sim(\text{Segment}\#\)] \wedge [\sim(\text{Cluster}\#)]$$

Indonesian [AUSTRONESIAN] (Soderberg & Olson

2008), **Ibibio** [NIGER] (Urua 2004), ...: disallow /b/, /d/, /g/ word-finally: [C=(b ∨ d ∨ g)]→[~(C#)]

RESTRICTION OF SEGMENT(S)/CLUSTERS TO WORD-FINAL POSITION

ONLY:

("positive demarcative signals" – Anderson 1965)

A given segment/cluster occurs mandatorily at

word endings:

$$(\text{Segment} \vee \text{Cluster}) \rightarrow [(\text{Segment} \vee \text{Cluster})\#]$$

Nepali [INDO-ARYAN] (Khatiwada 2009): the velar

nasal always occurs word-finally: (C=ŋ)→(C∅#)

RESTRICTION OF OCCURRENCE OF CERTAIN SEGMENTS/CLUSTERS AT

WORD-ENDINGS:

Segments/clusters that are the only admitted word-finally (though they can occur in other positions as well – see

"onset/coda asymmetry").

$$(\text{C}\#) \rightarrow (\text{C} \in \text{A})$$

(A=Subset of L consonant inventory)

Gayo [AUSTRONESIAN] (Eades & Hajek 2006):

Fricatives, nonpalatal nasals, rhotics, laterals and voiceless stops are the only consonants admitted word-finally:

[[C#]→(C ∈ A)]∧[A={Fric, NonPalNas, R, Lat, VlessStops}]

[.....]

ROMANCE CODAS

- tendency towards empty codas;

- *[ComplexCoda];

- "unmarked" Coda: [+son], [+cor].

(see, e.g., Glessgen 2007: 142)

But:

- some evidence of the existence of word-final codas different from word-medial codas:

Peninsular Spanish: admits /d/-filled codas word-finally

(Navarro Tomás 1926 : 99-100 ; Quilis 1993 : 204-205):

«*sed*» 'thirst',

«*huésped*» 'host',

«*césped*» 'grass',

«*juventud*» 'youth'.

Catalan: admits /dʒ/-filled codas word-finally (even though they are often phonetically deleted;

http://en.wikipedia.org/wiki/Catalan_phonology).

THE CASE OF EUROPEAN PORTUGUESE

European Portuguese phonology:

highly restrictive as far as coda-filling is concerned:

- filled codas are less frequent than empty codas
- no complex codas
- only /l/ ([ʎ]), /r/ and /s/ are admitted as coda-fillers (see, e.g. *Mateus & Andrade 2000*)

But:

Word-finally,

- 'extra-heavy rhymes'
- codas filled by consonants different from /l/, /r/ or /s/.

(see, e.g., *Veloso 2008, 2009, forthcoming*)

Extra-heavy rhymes admitted word-finally only in

European Portuguese:

/NGN/#	/NGNS _{pl} /#	/NGNS _{lev} /#
« <i>pão</i> » 'bread' [pãw̃]	« <i>mãos</i> » 'hands' [mãw̃]	« <i>Guimarães</i> » (place-name) [gimã'rẽj]
« <i>ontem</i> » 'yesterday' [õ'tẽj]	« <i>irmãos</i> » 'brothers' [ir'mãw̃]	« <i>Simões</i> » (person-name) [si'mõj]
« <i>homem</i> » 'man' [õ'mẽj]	« <i>alemães</i> » 'German [plural]' [alimãj]	
« <i>ruim</i> » 'bad' [rũj]		

Unattested word-medially : *['pãw̃.tu], *['mãw̃'.tru]

Word-final codas filled by consonants different from /l/,

/r/ and /s/

Segmental /n/

« *germen* » 'germ', ['ʒɛrmɛn]

« *plâncton* » 'plankton', ['plãkton]

Unattested word-medially: *['kɛn.tu], *['ʒir'mɛn.gu]

/ks/#

« *tórax* » 'thorax' ['tɔrɛks]

['fɔrɛps]

/ps/#

« *fórceps* » 'forceps'

« *bíceps* » 'biceps' ['bisɛps]

« *córtex* » 'cortex' ['kɔrɛks]

Unattested word-medially: *['tɛks.tu], *['laps.tu]

J. Veloso, P. T. Martins. 2010. Phonotactic Constraints and Word-Demarcation in Romance. *CUNY Conference on the Word in Phonology*: City University of New York, January 2010. Handout.

**“PROSODIC TOLERANCE
OF THE WORD RIGHT BOUNDARY”
(PTWRB)
(*Veloso forthcoming*)**

**Declarative Phonology-Based Formalizations
of Possible “PTWRB Phenomena”
and Other Word-Boundary- Sensitive
Phonotactic Constraints in Romance**

P.1: Prohibition of word-initial /k/, /ɲ/ and /r/ in European Portuguese

[Seg = (k ∨ ɲ ∨ r)] → ~[#_{init}∅Seg]

P.2: PTWRB in European Portuguese

{Seg = [(n) ∨ ((k ∨ p)s) ∨ (VGN) ∨ (VGNs)]} → [(Seq ∅#_{fin}) ∧ (. = #_{fin})]

S.1 : Word-final /d/-codas in Pensinsular Spanish

[Seq=(d.)] → [(Seq ∅#_{Fin}) ∧ (. = #_{Fin})]

C.1 : Word-final /dʒ/-codas in Catalan

[Seq=(dʒ.)] → [(Seq ∅#_{Fin}) ∧ (. = #_{Fin})]

Key

P = (European) Portuguese	Init = Initial
S = Spanish	Fin = Final
C = Catalan	∅ = precedes immediately
Seg = Segment	# = Word boundary
Seq = Segment combination	
. = Syllable boundary	
V, G, N, S = Vowel, Glide, Nasal, Palatal Fricative	
C = Consonant	

FINAL REMARKS

- 1 – Some phonotactic constraints can be accepted as word-boundary cues at least in some languages.
- 2 – Apart from the descriptive interest of this regularity, its relevance is twofold:
 - a) it can explain how hearers may identify word-boundaries in speech processing tasks;
 - b) it may be useful for the development of automatic tools for word demarcation within speech continua.

- 3 – It must be added, though, that in languages where PTRWB is found, it is a sufficient but unnecessary condition for word boundary identification.
- 4 – This kind of linguistic data offer us an extra amount of evidence favouring the word as a linguistic domain/unit and the necessity of including phonological aspects among the “wordhood criteria”.
- 5 – Declarative Phonology-based formalisms (Scobbie et al. 1996; Angoujard 2006) offer adequate descriptions of all relevant, surface-observable phonotactic regularities of the lexicon of a given language.

REFERENCES

- Anderson, J. M. 1965. The demarcative function. *Lingua*. 13: 185-188.
- Angoujard, J.-P. 2006. *Phonologie Déclarative*. Paris: CNRS.
- Eades, D.; Hájek, J. 2006. Gayo. *Journal of the International Phonetic Association*. 36(1): 107-115.
- Glessgen, M.-D. 2007. *Linguistique romane. Domaines et méthodes en linguistique française et romane*. Paris: Armand Colin.
- J. Veloso, P. T. Martins. 2010. Phonotactic Constraints and Word-Demarcation in Romance. *CUNY Conference on the Word in Phonology*: City University of New York, January 2010. Handout.
- Hargus, S.; Beavert, V. 2006. Word-initial clusters and minimality in Yakima Sahaptin. *Phonology*. 23: 21-58.
- Jones, D. 1931. The ‘word’ as a phonetic entity. *Le Maître Phonétique*. 36: 60-65. In: W. E. Jones, J. Laver (Eds.), *Phonetics in Linguistics. A Book of Readings*. London: Longman, 154-158.
- Kager, R. 1995. The metrical theory of word stress. In: J. A. Goldsmith (Ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell, 367-402.
- Keane, E. 2004. Tamil. *Journal of the International Phonetic Association*. 34(1): 111-116.
- Khatiwada, R. 2009. Nepali. *Journal of the International Phonetic Association*. 39(3): 373-380.
- Marlett, S. A. et al. 2005. Seri. *Journal of the International Phonetic Association*. 35(1): 117-121.
- Masquiza, F. C.; Marlett, S. A. 2008. Salasaca Quichua. *Journal of the International Phonetic Association*. 38(2): 223-227.
- Mateus, M. H. M. et al. 2003. *Gramática da Língua Portuguesa*. 5th ed. Lisboa: Garninho.
- Mateus, M. H. M.; Andrade, E. 2000. *The Phonology of Portuguese*. Oxford: Oxford University Press.
- Navarro Tomás, T. 1926. *Manual de pronunciación española*. Madrid: Junta para Ampliación de Estudios e Investigaciones Científicas / Centro de Estudios Históricos.
- Padayodi, C. M. 2008. Kabiye. *Journal of the International Phonetic Association*. 38(2): 215-221.
- Quilis, Antonio (1993): *Tratado de fonología y fonética españoles*. Madrid: Gredos.

- Scobbie, J. M.; Coleman, J. S. ; Bird, S. 1996. Key Aspects of Declarative Phonology. In: J. Durand & B. Laks (Eds.). *Current Trends in Phonology: Models and Methods*. Manchester: ESRU/University of Salford, II, 685-709.
- Soderberg, C. D.; Olson, K. S. 2008. Indonesian. *Journal of the International Phonetic Association*. 38(2):209-213.
- Urua, E. 2004. Ibibio. *Journal of the International Phonetic Association*. 34(1): 105-109
- Van Der Hulst, H. ; Van De Weijer, J. 1995. Vowel harmony. In: J. A. Goldsmith (Ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell, 495-534.
- Veloso, J. (forthcoming). Rimes /VGNs/ en position finale de mot en portugais: une contrainte «sensible au mot». *Proceedings of the XXV International Congress of the Société de Linguistique Romane* [Innsbruck, 2007].
- Veloso, J. 2008. Coda-avoiding: Some Evidence from Portuguese. *Romanitas*. 3. <http://humanidades.uprrp.edu/romanitas>.
- Veloso, J. 2009. Découpage de continuums phonétiques en mots: Critères formels vs. indices statutaires. In : O. Crouzet, A. Tifrit & J.-P. Angoujard (Eds). *Proceedings of JEL'2009/JEG'2009*. 6^{èmes Journées d'Études Linguistiques. Nantes: Université de Nantes, 85-90.}
- J. Veloso, P. T. Martins. 2010. Phonotactic Constraints and Word-Demarcation in Romance. *CUNY Conference on the Word in Phonology*: City University of New York, January 2010. Handout.