

PHONOTACTIC CONSTRAINTS AND WORD DEMARCATION IN ROMANCE¹

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

João Veloso & Pedro Tiago Martins
(Center of Linguistics of the University of Porto, Portugal)



¹Part of this research is funded by the Centro de Linguística da Universidade do Porto (Fundação para a Ciência e a Tecnologia, Portugal, FEDER/POCTI, U0022/2003).

Abstract

In many languages, some phonological segments or structures are allowed at word boundaries only. The importance of this fact is manifold: it shows that phonology is not “blind” to the word as a linguistic unit; it can explain some aspects of speech processing (how speech continua are split into words); it may help to build software aimed at identifying word boundaries within larger continua. For European Portuguese and for other Romance languages known to be very restrictive as far as the segmental coda-filling is concerned. However, these restrictions are somehow cancelled at word-endings, which admit phonotactic combinations that are not allowed elsewhere. A formalization of this phenomenon, inspired by the logical formalisms of Declarative Phonology, is also proposed here.

Some Basic Assumptions

1 – “Words” are **not universal** linguistic units; in some languages, they can be conveniently and completely replaced, for the sake of linguistic description, by units such as the “morphemes”.

2 – However, in languages that admit WFRs (inflectional languages, mainly), words do seem to behave as linguistic units/structures.

3 – Words as linguistic units are ruled by morphological, syntactical, phonological and semantic properties which interact language-specifically.

4 – **Phonotactic constraints act in many languages as word-boundary cues** (Jones 1931; Anderson 1965); this may be properly described by means of logic-based formalisms, as it is proposed by frameworks like Declarative Phonology (Scobbie et al. 1996; Angoujard 2006).

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.

Word-Sensitive Phonological Phenomena (some examples)

- Vowel Harmony
- Stress Assignment
- Prohibition of Segment(s)/Clusters in Word-Initial Position
- Restriction of Segment(s)/Clusters to Word-Initial Position Only
- Prohibition of Segment(s)/Clusters in Word-Final Position
- Restriction of Segment(s)/Clusters to word-Final Position Only
- Restriction of Occurrence of Certain Segments/Clusters at Word-Endings

Examples and sources in the *handout*.

Romance Codas

- tendency towards empty codas;
- *[ComplexCoda];
- “unmarked” Coda: [+son], [+cor].
(see, e.g., *Glessgen 2007: 142*)

Word-final codas: diachronical evidence of their special status.
(see, e.g., *Lausberg 1963*)

But:

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

Peninsular Spanish: admits /d/-filled codas word-finally:
«*sed*» ‘thirst’,
«*huésped*» ‘host’,
«*césped*» ‘grass’,
«*juventud*» ‘youth’.

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

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/ ([H]), /r/ and /s/ are admitted as coda-fillers
(see, e.g., *Mateus & Andrade 2000*)

But:

Word-finally,

- (1) ‘extra-heavy rhymes’
- (2) codas filled by consonants different from /l/, /r/ or /s/.
(see, e.g., *Veloso 2008, 2009, forthcoming*)

(1) Extra-heavy rhymes admitted word-finally only in European Portuguese:

/VGN/#
«*pão*» ‘bread’ [pãw̃]
«*ontem*» ‘yesterday’ [õntẽj]
«*homem*» ‘man’ [õmẽj]
«*ruim*» ‘bad’ [rũj]

/VGN_{pl}/##
«*mãos*» ‘hands’ [mãw̃j]
«*irmãos*» ‘brothers’ [irãmẽw̃j]
«*alemães*» ‘German [plural]’ [elãmẽj]

/VGN_{lex}/#
«*Guimarães*» (place-name) [gĩmẽrãmẽj]
«*Simões*» (person-name) [sĩmõj]

Unattested in non-final position: *[pãw̃.tu], *[mãw̃j.tru]

(2) Word-final codas filled by consonants different from /l/, /r/ and /s/

Segmental /n/: «*germen*» ‘germ’, [ʒerment];
«*plâncton*» ‘plankton’, [plãktõn]

Unattested in non-final position: *[kẽn.tu], *[ʒir̃men.gu]

/ks/ and /ps/: «*tórax*» ‘thorax’ [tõrẽks]
«*fórceps*» ‘forceps’ [fõrẽps]
«*córtex*» ‘cortex’ [kõrẽtks]
«*bíceps*» ‘biceps’ [bĩsẽps]

Unattested in non-final position: *[tẽks.tu], *[bĩsẽps]

“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/, /p/ and /t/ in European Portuguese

[Seg = (k ∨ p ∨ t)] → ~[#_{ini} ∅ Seg]

P.2: PTWRB in European Portuguese

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

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

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

C.1: Word-final /dʒ/-codas in Catalan

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

Key in the *handout*.

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:

- it can explain **how hearers may identify word-boundaries** in speech processing tasks;
- 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 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.
- Glessgen, M.-D. 2007. *Linguistique romane, Domaines et méthodes en linguistique française et romane*. Paris: Armand Colin.
- Jones, D. 1931. The ‘word’ as a phonetic entity. *Le Maître Phonétique*. 36: 60-65.
- Lausberg, H. 1963. *Romanisches Sprachwissenschaft*. Berlin: Walter de Gruyter.
- Mateus, M. H. M.; Andrade, E. 2000. *The Phonology of Portuguese*. Oxford: Oxford University Press.
- 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: ESRI/University of Salford, II. 685-709.
- Veloso, J. (forthcoming). Rimes /VGNs/ en position finale de mot en português: 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.uprr.edu/romanitas>.
- Veloso, J. 2009. Découpage de continuum phonétiques en mots: Critères formels vs. indices substantiels. In: O. Crouzet, A. Tifrit & J.-P. Angoujard (Eds.), *Proceedings of JEL’2009/IEG’2009. 6^{èmes} Journées d’Etudes Linguistiques*. Nantes: Université de Nantes, 85-90.