Prosodic Trapping in Catalan*

Teresa Cabré (UAB) & Michael Kenstowicz (MIT)

January 1995

Mester (1994) points to the phenomenon of cretic shortening in Preclassical Latin as evidence for bimoraic metrical parsing: a HLH sequence (H=heavy syllable, L=light) shortens to HLL while LLH remains unchanged (cf. dēsinō > dēsino 'cease' 1sg. vs. studeō 'strive' 1sg.). According to Mester the force driving the asymmetry is exhaustive metrical parsing with an overriding condition of bimoraic feet: LLH parses directly as (LL)(H) while HLH parses initially as (H)L(H) and then improves to (H)(LL) by grouping the trapped medial syllable with the following H; the cost of parsing the trapped L is shortening the final H. Prince & Smolensky (1993) offer a constraints-based Optimality Theory (OT) analysis of the phenomenon that operates on a single level with no intermediate repair stage. In this squib we document a similar HLH vs. LLH contrast in Catalan hypocoristics and briefly compare a two-stage "parse and repair" analysis with a single-level OT account.

Following McCarthy & Prince (1986, 1991) we assume that hypocoristics arise from minimization of an input that like all other lexical categories must project up the hierarchy: Mora→Syllable→Foot→Prosodic Word. As we shall see, Catalan imposes minimal parsing at a precise point in the hierarchy--at the level of the Prosodic Word. The paradigms in (1) sample the phenomenon; see Cabré (1993a,b) for extensive analysis and discussion; in our transcriptions the truncated material is enclosed in angled brackets (e.g. Ambròs truncates to Bròs) and the tick marks stress on the following syllable.

(1)

a. H'H	b. L'H	c. LL'H	d. HL'H
<Am>bròs	<Re>mei	<Isa>bel	<Bar>tomeu, Meu
<An>ton	<Mi>quel	<Segi>mon	<Sal>vador
<Ar>nau	<Ra>mon	<Joa>quim	<Mont>serrat, Rat

e. H'HL	f. L'HL	g. LL'HL	h. H'LL
<Al>fonso	<Do>mingo	<Eli>zenda	<Ig>nasi
<Fran>cisco	<Clo>tilde	<Leo>poldo	<Cris>tina
<Rai>munda	<Ja>cinto	<Cala>manda	<Al>fredo

i. L'LL	j. HL'LL	k. LL'LL
<I>sidro	<Enri>queta	<Teo>dora
<Gre>gori	<Magda>lena	<Jose>fina
<Hi>lari	<Marga>lida	<Geno>veva

We see that in each case the output is at least two moras long but never greater than two syllables. Prosodic trapping is evident in the striking contrast between (1c) vs. (1d): LLH Isabel truncates to <LL>H Bel while HLH Bartomeu truncates to <H>LH Tomeu and optionally to <HL>H Meu. Native speakers reject as "completely impossible" <L>LH truncations such as Isabel → *Sabel, Segimon → *Gimon, and Joaquim → *Aquim. Additional examples: Nicolau → Lau; Meritxell → Txell; Frederic → Quic; vs. Baltasar → Tasar, Xar; Narciset → Ciset; Alcover → Cover.[1]

Following Cabré (1993a,b), we might propose a two-stage analysis in which a right-to-left moraic trochee parse (Hayes 1994:69) imposes a (LL) or (H) footing: (LL)(H) vs. (H)L(H). Minimization then parses out the rightmost (stressed) foot. This analysis derives most of the truncations directly; but it fails to account for the (H)L(H) trapping structures

<Bar>tomeu, <Sal>vador, <Mont>serrat. For these cases we might posit a repair rule that expands the hypocoristic to take in an unparsed syllable on the left (2a). Under this "loose footing" option (McCarthy & Prince 1991), the stray L is attached at the level of the Prosodic Word, violating strict layering of the prosodic hierarchy. Obligatory loose footing of a following L would also account for cases such as LHL Domingo → Mingo and HHL Alfonso → Fonso (2b).

```
(2)                             P                           P
                                |                           |
      F    F        F           F       F  F       F        F
      |    |        |       /   |       |  |       |        |  \
  a.  Salvador →  va dor →  va dor    b.  Alfonso →  fonso →  fonso
```

However, the loose footing option must be short-circuited by an overriding requirement that the output never exceed two syllables. Otherwise, <H>LLL *<En>riqueta might be expected beside <HL>LL <Enri>queta (1j). Also, if (H)(H)L Alfonso truncates to (H)L Fonso by suppressing the first foot and sister adjunction of L with the remaining foot (2b), we might expect proparoxytones such as Ang'elica and Pen'elope to truncate to G'elica and N'elope. But in fact proparoxytones have no legitimate hypocoristic output. (Thornton 1995 reports a similar finding for Italian.)

In their analysis of Japanese loanword truncation, Itô & Mester (1992) derive the loose footing option evident in (H)L maiku<rohoN> 'microphone' from a branching requirement that is allowed to range across different intervals of the prosodic hierarchy--a type of prosodic "relativized minimality". Japanese clippings impose branching at the Prosodic Word or Foot (but crucially not the Syllable) level. Foot-level branching ensures that the output is at least disyllabic. Branching at the

Prosodic Word correctly predicts that (H)L maiku<rohoN> will be accompanied by (LL)L terebi<zyoN> 'television' as well as F+F structures such as (LL)(LL) rihabiri<teesyoN> 'rehabilitation', (H)(LL) koNbini<ensu> 'convenience store', and (H)(H) baateN<daa> 'bartender'. But such longer structures is precisely what we do not find in Catalan: *<Flo>rentina, *<Er>nestina, *<Er>menegildo. Rather, these forms must truncate to disyllables: <Floren>tina, <Ernes>tina, <Ermene>gildo.

We might try to make better sense of the disyllabic upper bound as a direct reflection of Catalan foot structure. Under this scenario, all Catalan hypocoristics truncate to the minimal word qua quantity-sensitive foot. However, this analysis implicates a heterogeneous collection of foot types that do not fall neatly under any of the common footing algorithms: (H), (LH), (HL), (LL) but not (HH), (L), nor (LL)L#. From the OT perspective, such a motley inventory is expected if the outputs reflect the optimal choices among conflicting constraints. As McCarthy & Prince (1993a) emphasize, the notion of conflicting constraints allows the theory of prosodic morphology to come to terms with data that cannot be subsumed under a single template but still maintain the hypothesis that prosodic parsing arises from the imposition of a restricted set of UG constraints.

In the remainder of this squib we sketch such an OT analysis. The relevant constraints operate at two levels of the prosodic hierarchy: parsing by the Prosodic Word and parsing by the Foot. Although Catalan lacks secondary stresses, our analysis of the prosodic trapping crucially depends on the assumption that pretonic syllables are metrified into feet (see below). Three constraints play a major role in organizing syllables into feet: **Ft Binarity** (a foot must contain at least two moras and at most two syllables), **Parse-σ** (syllables are parsed into feet), and **Align-Ft** (feet are oriented with respect to the right edge of the word). Ft Binarity dominates Parse-σ blocking unbounded feet. Parse-σ dominates Align-Ft ensuring multiple footing. These constraints select a (LL)(H) analysis for Isabel, as shown in (3).

(3)

/LLH/	Ft-Bin	Parse-σ	Align-Ft
$ (LL)(H)			σ#, #
L(LH)		*!	#
(LLH)	*!		#
(L)(LH)	*!		σσ#, #

Three constraints operate at the level of the Prosodic Word. Given the LLH vs. HLH asymmetry, the output structure for <Isa>bel must be that in (4a) where the initial foot is unparsed at the level of the Prosodic Word.

```
          P              P                                P
           \            / \                              /
         F   F        F   F         F   F             F   F
        /\   |       /\   |        /\   |            /\   |
(4)   a.  Isa bel    b.  Isa bel    c.  Isa bel    d.  Isa bel
```

Since feet are normally parsed by Prosodic Word in virtue of the constraint **Parse-Ft**, candidate outputs with a full foot parsing (such as 4b) must be blocked by a higher ranking constraint specific to hypocoristics that imposes minimal structure at the level of the Prosodic Word. We informally abbreviate this constraint as **Min.PrWd** and count violations in terms of the number of daughters to the Prosodic Word node P. The best candidate would be one lacking any projection to P. Such a "null parse" is blocked by a superordinate constraint requiring that a lexical category (noun, verb, adjective) project to the level of the Prosodic Word (Prince & Smolensky's 1993:43 **Lx≈Pr**). The crucial rankings are thus Lx≈Pr >> Min.PrWd >> Parse-Ft. The tableau in (5) shows how these Prosodic-Word level constraints sort among the candidates in (4). In order to eliminate candidate (4d) and force truncation at the left edge, we postulate an alignment constraint requiring the right edge of the Prosodic Word to coincide with the right edge of the hypocoristic stem: **Align-R (PrWd, Stem)**.

(5)

/Isabel/	Lx≈Pr	Al-R(PW,St)	Min-PrWd	Parse-Ft
$ <F>F			*	*
FF			**!	
<F><F>	*!			**
F<F>		*!	*	*

Returning to the foot level, more interesting are the HLH trapping cases, where the attested foot parsings in basically trochaic (left-headed) systems include (HL)H Bani-Hassan Arabic (Kenstowicz 1994:572; cf. (HL)L Manam (Buckley 1994:22) and Chi-Mwi:ni (Kenstowicz 1994:249)), (H)(LH) Catalan, and (H)L(H) Classical Latin (Mester 1994). Following Prince & Smolensky (1993:59), we see these parses as the resolution of three conflicting foot-level constraints: *HL, Parse-σ, and Trochee. ***HL** requires a heavy syllable to align with the right edge of the foot. This constraint passes the (LH) iamb and stars the unbalanced (HL) trochee. **Parse-σ** penalizes syllables that are not metrified by a foot and **Trochee** is a cover term for the constraints that impose a strict left-headed bimoraic (LL), (H) parse. Depending on which constraint is ranked lowest, one of the competing (HL)(H), (H)(LH), (H)L(H) parses emerges as optimal. The Catalan (H)(LH) outcome reflects a {*HL, Parse-σ} >> Trochee hierarchy, as shown by the tableau in (6a). (*HL must also dominate Align-Ft, at least if alignment is evaluated from the edge of the foot--as opposed to the head, as in Green 1993).

(6) a.

/HLH/	*HL	Parse-σ	Trochee	Align-Ft
$ (H)(LH)			*	σσ#, #
(HL)(H)	*!			σ#, #
(H)L(H)		*!		σσ#, #

```
                P
                |
            F   F
            |  / \
b.        Sal va  dor
```

The constraints of (6a) choose (6b) as the optimal hypocoristic. As mentioned earlier, Catalan has no discernible secondary stresses; indeed, vowels are reduced outside the stressed syllable. Nevertheless, our analysis of the contrasting LLH vs. HLH truncations crucially depends on a different metrical grouping of the hidden pretonic material: (LL)('H) vs. (H)(L'H).

The variation between the (H)(LH) _Tomeu_ and (H)L(H) _Meu_ parses of _Bartomeu_ suggests that the ranking between Parse-σ and Trochee is not rigid and is in part lexically determined.[2] As the tableau in (7a) demonstrates, ranking Trochee above Parse-σ derives the monosyllabic option Meu in (7b).

(7) a.

/HLH/	*HL	Trochee	Parse-σ	Align-Ft
(H)(LH)		*!		σσ#, #
(HL)(H)	*!			σ#, #
$ (H)L(H)			*!	σσ#, #

```
              P
               \
           F       F
           |       |
b.      Bar to  meu
```

LH cases (1b) such as Ramon only show the (H) option; this makes sense in the context of language use--the Parse-σ>> Trochee option selecting (LH) over <L>(H) derives an output that is no shorter than the input.

Our explanation for the (H)(LH) parse depends crucially on the *HL constraint knocking out the (HL)(H) candidate. It is therefore perhaps surprising that a (HL) footing must be systematically chosen for the cases with a heavy penult such as Domingo if we are to maintain our hypothesis that the Catalan hypocoristic is coincident with a foot. By the logic of the OT model, this must reflect a constraint with higher priority. We suggest another alignment constraint on the Prosodic Word requiring its right edge to coincide with a foot: **Align-R(PrWd, Ft)**. If ranked above *HL, Align-R(PrWd, Ft) discards a loose footing parse <L>(H)L in favor of <L>(HL) for <Do>mingo.

(8)

/LHL/	Al-R(PW, Ft)	*HL
$ L(HL)		*
L(H)L	*!	

Alignment of the Prosodic Word with a foot is a natural solution to the HL case. Itô & Mester (1992) observe a similar effect in Japanese where LHL rokeešoN truncates to (LL)LH roke<ešooN> rejecting a loose footing L(H) option *rokee at the cost of splitting a long vowel in order to ensure that a bimoraic trochee stands flush against the left edge of the word.

Our analysis of Catalan has invoked several constraints that align the categories of Foot, Prosodic Word, and Stem at their right edges. These are reviewed in (9).

(9) Align-R(Ft, PrWd) enforces rightward foot orientation

Align-R(PrWd, Stem) blocks right-edge truncation

Align-R(PrWd, Ft) blocks loose footing of ...HL#

Right-edge alignment also offers an attractive explanation for why proparoxytones such as Ang'elica and Pen'elope have no legitimate hypocoristic even though trisyllabic G'elica and N'elope would minimize the base by one syllable. In Catalan (as in Spanish, Harris 1994), antepenultimate stress is a marked option in comparison to penultimate stress. Consequently, proparoxytones such as Pen'elope require a lexical marking of their foot structure: /L(LL)L/. Given the OT premise that the input is contained in the output, the only possible foot that could align with the right word edge would be a monosyllable: L(LL)(L). Since there are no absolutely malformed structures in OT but only better and worse ones, some move must be made to ensure that for /L(LL)L/ inputs no output is a better analysis than some output--in particular better than <L(LL)>(L). Prince & Smolensky (1993:48) raise this question for Latin where minimality blocks CV monomoraic words. Their suggestion is that constraints on prosodic shape such as Ft Binarity can dominate Lx≈Pr which requires a lexical word to project a Prosodic Word. The best candidate is thus one where Gen adds no prosodic structure to the input--the "null parse". The candidate with a null parse is ruled out by a requirement that stands outside the optimality system itself: according to Prince & Smolensky (1993:48), it is "uniquely unsuited to life in the outside world" and receives no phonetic interpretation.

We suggest a similar analysis here. In (10) we list the relevant competitors. The null parse candidate (10a) must be the winner. Consequently, the other candidates must be eliminated by constraints that rank above Lx≈Pr.

```
(10)                          P                P
                               \              /\
          F                 F   F            F  \
         /\                /\   |           /\   \
   a.   Penelope   b.   Penelope   c.   Penelope

          P                 P                P
          |                 |                 \
          F                 F                F  F
         /\               / | \             /\/\
   d.   Penelope   e.   Penelope   f.   Penelope
```

In (11) we review the constraints and crucial rankings developed to this point in the analysis. (10a) will emerge the winner if Lx≈Pr is ranked below the other three undominated constraints in (11).

```
(11)  Lx≈Pr        Al-R(PrWd, Stem)       Ft-Bin      Al-R(PrWd, Ft)
        |                                    |           |
      Min.PrWd                               |         *HL
        |                                    |        /  |
      Parse-Ft                            Parse-σ        |
                                        /    |           |
                                  Trochee   Al-R(Ft, PrWd)
```

If Ft-Bin is allowed to dominate Lx≈Pr then candidates (10b) and (10e) are eliminated.[3] Al-R(PrWd, Ft) that blocks loose footing will eliminate (10c). This constraint is already high-ranking in virtue of crucially dominating *HL; nothing prevents us from ranking it above Lx≈Pr too. (10d) can be excluded by embedding Lx≈Pr under the hitherto unranked Al-R(PrWd, Stem). The null parse candidate (10a) will pass both of these alignment constraints since it lacks a PrWd and hence fails to meet the antecedent clause of the constraint ("if there is a PrWd, then its right edge must align with the right edge of the stem/foot"). The final candidate (10f) can be excluded by a constraint banning the assignment of a syllable to two different feet. The tableau in (12) reviews how these constraints select the null parse as the optimal member of the candidate set in (10). (To save clutter, elements parsed by the Prosodic Word are in bold.)

(12)

/L(LL)L/	*Ft-Overlap	Ft-Bin	Al-R(PW,St)	Al-R(PW,Ft)	Lx≈Pr
$ L(LL)L					*
L(LL)(**L**)		*!			
L(**LL**)**L**				*!	
L(**LL**)L			*!		
L(**LLL**)		*!			
L(L(**L**)**L**)	*!				

In sum, the OT analysis of Catalan hypocoristics proposed here arises from the reranking of three constraints that are masked in the language's regular morphology: minimization of the prosodic word to a single foot (Min.PrWd) and alignment of these two prosodic categories at their right edges with the right edge of the stem: Align-R(PrWd, Stem) and Align-R(PrWd, Ft). Our analysis also crucially depends on taking the truncation of <Sal>vador as direct evidence for (H)(LH) foot parsing and consequently implies that Catalan has a hybrid inventory of feet in its output that combines the (L'H) canonical iamb with the ('LL) canonical trochee. Finding additional evidence to (dis)confirm these structures thus becomes a high priority.

Notes

*We thank Sylvain Bromberger, Tony Bures, Tom Green, Morris Halle, Jim Harris, and Jay Keyser for comments and criticism.

[1]Many words submit to an alternative analysis with truncation at the right edge: e.g. 'Isa<bel>. Such right-edge truncations constitute a different system (introduced more recently) that is essentially the same

as the truncation found in Castilian Spanish. See Cabré 1993a:24 for discussion and Prieto (1992) for analysis of Castilian Spanish hypocoristics.

[2]Alternatively, the variation might reflect the level at which minimality is imposed: Tomeu minimizes to a nonbranching PrWd (given that Tomeu constitutes a (LH) foot) while Meu minimizes at the foot level as well (but still projects to Prwd in virtue of its heavy syllable). Cf. Prince & Smolensky's analysis of Latin (1993:52) where Nonfinality is imposed at various levels of the prosodic hierarchy.

[3]Oxytones with a final light syllable such as Salomé and Bernabè also underlie no legitimate hypocoristics: *Mé, *Lomé, *Bè, *Nabè. Given that the exceptional stress of Salomé is recorded in the input LL(L), we have independent confirmation for the Binarity >> Lx≈Px ranking.

References

Buckley, Eugene. 1994. Alignment in Manam stress. Ms. University of Pennsylvania, Philadelphia, Pa.

Cabré i Monné, Teresa. 1993a. Estructura gramatical i lexicó: el mot mínim català. Tesi doctoral, Universitat Autònoma de Barcelona.

_____. 1993b. Sensitivity of Catalan to syllable weight. Ms., Universitat Autònoma de Barcelona

Green, Thomas. 1993. Configurational constraints in a generalized theory of phonological representation. To appear, MIT Ph.D. dissertation, MIT, Cambridge, Mass.

Harris, James. 1994. Spanish stress. In Handbook of phonological theory, ed. John Goldsmith, xxx-xxx. Cambridge, Mass. :Blackwell.

Hayes, Bruce. 1994. Metrical stress theory: principles and case studies. Chicago: University of Chicago Press.

Itô, Junko & Armin Mester. 1992. Weak layering and word binarity. Ms. Santa Cruz: University of California.

Kenstowicz, Michael. 1994. Phonology in generative grammar. Cambridge, Mass. Blackwell.

McCarthy, John & Alan Prince. 1986. Prosodic morphology. Ms. Brandeis University, Waltham, Mass.

______. 1991. "Prosodic morphology", lecture notes from LSA Summer Institute, Santa Cruz, University of California.

______ 1993a. Prosodic morphology I: constraint interaction and satisfaction. Ms., Rutgers University and University of Massachusetts, Amherst, Mass. To appear, MIT Press.

______ 1993b. Generalized alignment. Yearbook of Morphology 1993: 79-153.

Mester, Armin. 1994. The quantitative trochee in Latin. Natural Language & Linguistic Theory 12: 1-61.

Prieto, Pilar. 1992. Truncation processes in Spanish. Studies in the Linguistic Sciences 22,1: 143-158.

Prince, Alan & Paul Smolensky. 1993. Optimality theory: constraint interaction in generative grammar. Ms., Rutgers University and

University of Massachusetts. To appear, MIT Press.

Thornton, Anna. 1995. Italian prosodic morphology. Ms., Università dell'Aquila and Istituto di Psicologia del CNR, Roma.

Teresa Cabré
Departament de Filologia Catalana
Universitat Autònoma de Barcelona
08193 Bellaterra
Spain
ilft3@ccuab1.uab.es

Michael Kenstowicz
Department of Linguistics
20D-219
MIT
Cambridge, Ma. 02139
Kenstow@mit.edu