Abstract

Modeling individual variation in prosody: the case of Spanish Clitic Left-Dislocations

Ingo Feldhausen
(ingo.feldhausen@gmx.de)

Labex EFL, UMR 7018-LPP, Université Paris-Sorbonne Nouvelle
UMR 7110-LLF, Université Paris-Diderot
Goethe-Universität Frankfurt am Main

Poster, presented at
The Fifth European Conference on Tone and Intonation (TIE5)
Oxford, UK, September 2012
&
Talk, presented at the
'Syntax-Phonology Interface from a Cross-linguistic Perspective' workshop
ZAS, Berlin, November 2012

This paper deals with individual differences in the prosodic phrasing of clitic left-dislocations (CLLDs) in embedded and non-embedded clauses of Spanish. It proposes an optimality-theoretic analysis of the attested inter-speaker variation. This variation, which previously has been typically ignored or attributed to performance, is considered to be a part of the linguistic competence.

Background: Although it is well known that prosodic aspects such as (a) intonational phrasing, (b) boundary realization, and (c) pitch accent realization differ between languages (Jun 2005), a growing amount of evidence questions whether these aspects are used homogenously within a given language (e.g. Féry 2004, Feldhausen 2010, 2011, Myrberg 2010, Niebuhr et al. 2011). Most of these studies apply the Stochastic Optimality Theory (SOT, Boersma & Hayes 2001) to account for frequency-dependent variation and, as a consequence, constitute an improvement over the categorical, non-probabilistic analyses of different prosodic aspects which are typically proposed. In SOT, free variation arises due to overlapping constraints, as these can generate multiple output forms from a single underlying form. SOT differs from classical OT (Prince & Smolensky 1993/2004) by assuming a continuous ranking scale (↤↦in Fig.1) and a stochastic candidate evaluation: Constraints have a certain ranking value (e.g. 88, 86, 75 in Fig.1), and the distance between the constraints may vary; constraints are not single points, but rather act as if they are associated with ranges of values (grey boxes in Fig.1). As a consequence, when the ranges of two close constraints overlap (Fig.1a), it is possible that – instead of a 'normal' ranking' – a reverse ranking results and another candidate wins. The distance between the constraints is determined for a given language by SOT. Despite this progress, the SOT approaches suffer a certain shortcoming: They ignore variation between speakers by proposing a "grammar of the average speaker" (see, e.g., the pattern in Fig.1a for all speakers). According to Pierrehumbert (2001:201), variation is an intrinsic part of linguistic competence; the frequency with which a given unit appears is an important factor in how it behaves in the system. If this assumption is correct, a grammar should not abstract away systematic individual differences.

Experiment: Based on data from a production experiment (scripted speech) in which a homogenous group of four native speakers of Peninsular Spanish uttered 144 sentences with non-embedded (***El águila**, la vendió mi hermano* 'The eagle, my brother sold.') and embedded CLLDs (*Bárbara supone que **el águila**, la vendió su hermano* 'Barbara assumes that the eagle, his brother sold'), it is shown here that CLLDs have an obligatory right boundary. In addition, the embedded clause is obligatorily separated from the matrix clause. Inter-speaker variation appears in the phrasing of the matrix clause: the boundary separating the matrix subject (here: *Bárbara*) from the matrix verb (here: *supone*) is optional, and clear frequency-dependent variation exists across the speakers (Realizations of (SV) groupings: Speaker A, 100%; Speaker B, 66%; Speaker C, 72%; Speaker D, 61%).

Proposal: In order to account for individual differences, I propose that the distance between two constraints is not fixed for the grammar of a given language G_L, but rather differs between individual speakers of a given language. The constraint hierarchy remains the same for all speakers. Differences in the frequency of output forms between speakers thus arise due to the different degrees of overlap of the given constraints between these speakers (Fig.1a vs. Fig.1b).

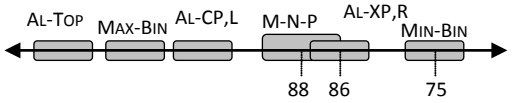


Fig.1a: Overlapping constraints (e.g. speaker D)

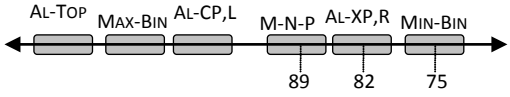


Fig.1b: Non-overlapping constraints (e.g. speaker A)

To account for the findings, I propose the constraint hierarchy ALIGN-TOP >> MAX-BIN$_{(IP\ HEAD)}$ >> ALIGN-CP,L >> MIN-N-PHRASES >> ALIGN-XP,R >> MIN-BIN, in which the two constraints MIN-N-PHRASES and ALIGN-XP,R overlap, thus guaranteeing a reverse ranking in order to account for the variation found in the data. Only ALIGN-TOP and ALIGN-CP,L are new (for the other constraints see Prieto 2006). ALIGN-TOP guarantees the boundary at the right edge of CLLD constituents, while ALIGN-CP,L secures the boundary preceding the embedded clause. Based on the modification of the SOT, I propose that speaker A shows no overlap between MIN-N-PHRASES and ALIGN-XP,R, (since he does not utter any (S)(V) groupings), while speaker D has the highest degree of overlap (with (S)(V) amounting to 39%). Speakers B and C exhibit intermediate degrees of overlap.

References:

Boersma, Paul & Hayes, Bruce. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32: 45–86.

Feldhausen, Ingo. 2010. *Sentential form and prosodic structure of Catalan* [Linguistik Aktuell / Linguistics Today 168]. Amsterdam: John Benjamins.

Feldhausen, Ingo. 2011. The prosodic phrasing of sentential objects. *Lingua* 121 (13): 1934-1964.

Féry, Caroline. 2004. Gradient prosodic correlates of phrasing in French. In: Meisenburg, Trudel & Selig, Maria (eds.): *Nouveaux départs en phonologie : Les conceptions sub- et suprasegmentales.* Tübingen: Narr, pp. 161-182.

Jun, Sun-Ah. 2005. *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: OUP.

Myrberg, Sara. 2010. *The Intonational Phonology of Stockholm Swedish* [Acta Universitatis Stockholmiensis. Stockholm Studies in Scandinavian Philology. New Series 53]. PhD Dissertation, University of Stockholm.

Niebuhr, Oliver, D'Imperio, Mariapaola, Gili Fivela, Barbara & Cangemi, Francesco. 2011. Are there "shapers" and "aligners"? Individual differences in signalling pitch accent category. *Proceedings of the 17th ICPhS, Hong Kong, China*, 120-123.

Pierrehumbert, Janet. 2001. Stochastic phonology. *Glot International* 5(6): 195–207.

Prieto, Pilar. 2006. Phonological phrasing in Spanish. In *Optimality-Theoretic Advances in Spanish Phonology* [Linguistik Aktuell/Linguistics Today 99], Sonia Colina & Fernando Martínez-Gil (eds), 39–60. Amsterdam: John Benjamins.

Prince, Alan & Smolensky, Paul. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar* [Rutgers University Center for Cognitive Science Technical Report 2]. (Published in 2004: Oxford: Blackwell).

Modeling individual variation in prosody: the case of Spanish Clitic Left-Dislocations

Ingo Feldhausen
(ingo.feldhausen@gmx.de)
TIE5, Oxford (UK), September 2012

Abstract

This paper deals with individual differences in the prosodic phrasing of *clitic left-dislocations* (CLLDs) in Spanish. It proposes a stochastic optimality-theoretic (SOT) analysis of the attested inter-speaker variation. This variation, which previously has been typically ignored or attributed to performance, is considered to be a part of the linguistic competence.

1. Introduction

- Prosodic aspects differ between languages (Jun 2005; e.g. (a) intonational phrasing, (b) boundary realization, and (c) pitch accent realization)
- Growing amount of evidence questions whether these aspects are used homogenously within a given language (Féry 2004, Feldhausen 2010, 2011, Myrberg 2010, Niebuhr et al. 2011 etc.)

> Variation is an intrinsic part of linguistic competence (Pierrehumbert 2001)

- Models, such as *Stochastic Optimality Theory* (SOT, Boersma & Hayes 2001), account for free variation within a population
- Approaches using SOT typically propose a "grammar of the average speaker", and ignore variation between speakers of that population

> "Variation is idiosyncratic and inherent in individual grammars" (Bresnan et al. 2007:340)

- A grammar should not abstract away systematic individual differences

2. CLLD in Spanish

- is "characterized by the presence of a phrase in the first position of the clause which is connected with that clause through the intermediary of some anaphoric element" (Alexiadou 2006)
- expresses the informational status of givenness (López 2009)

CLLD in simple clauses

(1) **El águila de Málaga**, *la* vendió mi hermano
'The eagle of Málaga, my brother sold (it)'

CLLD in embedded clauses

(2) Bárbara supone que **el águila**, *la* vendió su hermano
'Barbara assumes that the eagle, her brother sold (it)'

3. Prosodic Aspects of CLLD in Spanish

- 1st detailed study of Sp. CLLD (Feldhausen 2012)
- Production experiment (scripted speech)
  - Four subjects (3f,1m), 27-31 years of age (monolingual, subjects held a university degree)
  - Native speakers of Spanish (Murcia)
- 144 sentences (12 sent. x 3 repetitions x 4 speakers)
- Analysis based on Sp_ToBI (Aguilar et al. 2009)

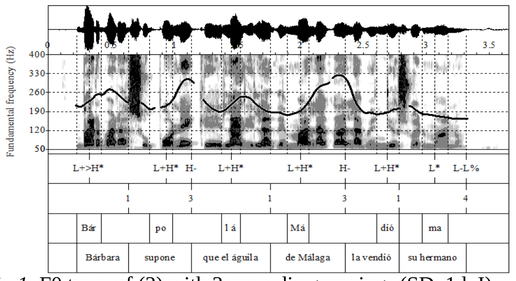


Fig.1: F0 trace of (2) with 3 prosodic groupings (SD_1d_I)

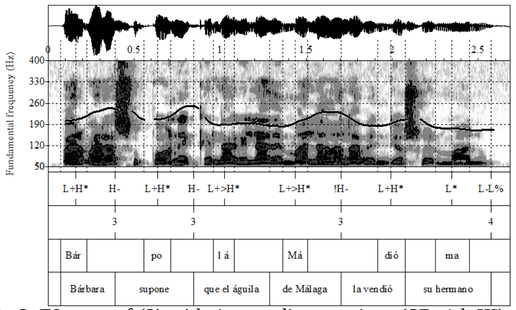


Fig.2: F0 trace of (2) with 4 prosodic groupings (SB_1d_III)

- **Results**:
  - Obligatory boundary after: CLLD (97%), matrix verb (99%)
  - Variation in phrasing & boundary realization

4. Individual Differences

- Frequency-dependent variation in realizing a boundary in a given position:

25% ~100% ~100%
( )( )( )
(3) [Bárbara supone [que **el águila** *la* vendió su hermano]$_{CP2}$]$_{CP1}$
S | V | q CLLD | V S
'Barbara assumes that the eagle, her/his brother sold (it)'

S)(V	Speaker
0%	A
33%	B
28%	C
39%	D
25%	Total

Table 1: Percentages of individual realizations of the 'S)(V'-boundary across target sentences with embedded CLLD

4. Individual Differences (cont.)

- (3) shows that only some boundaries are optional, while others are obligatory
- The degree of optionality differs across speakers & can hardly be caught by the average %

5. OT analysis - Constraints

Established constraints (cf. Prieto 2006):

Align-XP,R: For each XP there is a P such that the right edge of XP coincides with the right edge of P (P = prosodic phrase)
Min-N-Phrases: Minimize the number of prosodic phrases
Min-Bin: P-phrases should consist of minimally two ω
Max-Bin (IP Head): A phonological phrase which is the head of an IP constituent must be binary (at the ω level)

Additional constraints:

Align-Top,R: Align the right edge of a topic constituent [= CLLD] to the right edge of a prosodic phrase (Feldhausen 2010)
Align-CP,L: Align the left edge of a CP to the left edge of a prosodic phrase (Feldhausen 2010, 2011)

S V q CLLD V O	AlignTop,R	Max-Bin (IP Head)	Align-CP,L	Min-N-Phrases	Align-XP,R	Min-Bin
(S V q CLLD V O)	*!	*	*	*	**	
(S V q CLLD) (V O)			*!	**	*	
☞ (S V) (q CLLD) (V O)				***	*	*
(S)(V) (q CLLD) (V O)				****!		***
(S V q CLLD) (V)(O)			*!	***	*	**
(S V) (q CLLD) (V)(O)				****!	*	***

Table 2: Ranking for grouping of Fig.1 ('normal ranking')

S V q CLLD V O	AlignTop,R	Max-Bin (IP Head)	Align-CP,L	Align-XP,R	Min-N-Phrases	Min-Bin
(S V q CLLD V O)	*!	*	*	**	*	
(S V q CLLD)(V O)			*!	*	**	
(S V)(q CLLD)(V O)				*!	***	*
☞ (S)(V)(q CLLD)(V O)					****	***
(S V q CLLD)(V)(O)			*!	*	***	**
(S V)(q CLLD)(V)(O)				*!	****	***

Table 3: Ranking for grouping of Fig.2 ('reverse ranking')

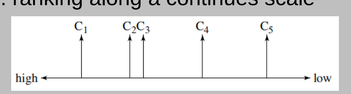
6. Stochastic Optimality Theory

- SOT (Boersma & Hayes 2001) accounts for frequency-dependent variation in data

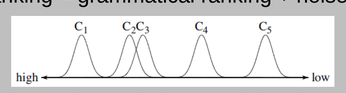
- OT: a grammar is a set of strictly ranked constraints
- Non-stochastic OT: ordinal ranking

$C_1 >> C_2 >> C_3 >> C_4 >> C_5$

- Stochastic OT: ranking along a continues scale



- Evaluation time:
  evaluation ranking = grammatical ranking + noise



strict ranking determines winning candidate form
90% of the time: $C_2 >> C_3$
10% of the time: $C_3 >> C_2$
(taken from Boersma 2003)

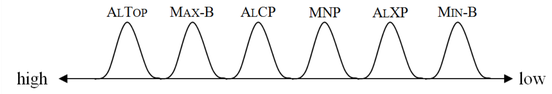
7. Modeling Individual Differences

Underlying Constraint Hierarchy

Align-Top,R >> Max-Bin$_{(IP\ Head)}$ >> Align-CP,L >> Min-N-Phrases >> Align-XP,R >> Min-Bin

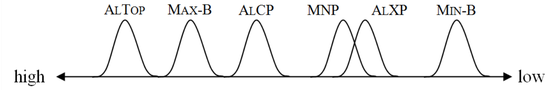
- ***Proposal***:
  - The constraint hierarchy of a grammar is invariant (as in other OT accounts)
  - But the distance between the two constraints M-N-P & AL-XP,R is not fixed and can differ between speakers
  - Thus, differences in the frequency of output forms between speakers arise due to different degrees of overlap of M-N-P & AL-XP,R

- **Grammar of speaker A:**



strict ranking determines winning candidate form
100% of the time: Min-N-P >> AL-XP,R (= Table 2)
0% of the time: AL-XP,R >> Min-N-P

- **Grammar of speaker D:**



strict ranking determines winning candidate form
61% of the time: Min-N-P >> AL-XP,R (= Table 2)
39% of the time: AL-XP,R >> Min-N-P (= Table 3)

- Individual grammars are built based on the grammar of a given language L

References

Aguilar, L, **et al**. 2009. Sp_ToBI. Training materials. <http.//prosodia.upf.edu/sp_tobi/en/> // **Alexiadou**, A. 2006. Left dislocation (including CLLD). In *The Blackwell Companion to Syntax*. Everaert, M. et al. (eds), 668-699. Oxford: Blackwell. // **Boersma**, P. 2003. SOT. Talk at LSA. // **Boersma**, P, **Hayes**, B. 2001. Empirical tests of the gradual learning algorithm. *LI* 32: 45–86. // **Bresnan**, J. **et al**. 2007. Typology in Variation. *EL&L* 11:301-346. // **Feldhausen**, I. 2010. *Sentential form and prosodic structure of Catalan*. Amsterdam: Benjamins. // **Feldhausen**, I. 2011. The prosodic phrasing of sentential objects. *Lingua* 121: 1934-1964. // **Feldhausen**, I. 2012. Prosodic Aspects of Clitic Left-Dislocations in Spanish. Ms. U Frankfurt & Labex EFL. // **Féry**, C. 2004. Gradient prosodic correlates of phrasing in French. In: Meisenburg, T, Selig, M. (eds.): *Nouveaux départs en phonologie.* Tübingen: Narr,161-182. // **Jun**, S.-A. 2005. *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: OUP. // **López**, L. 2009. *A Derivational Syntax for Information Structure*. Oxford: OUP. // **Myrberg**, S. 2010. *The Intonational Phonology of Stockholm Swedish*. PhD Thesis, U Stockholm // **Niebuhr**, O. **et al**. 2011. Are there "shapers" and "aligners"? *17th ICPhS*, 120-123. // **Pierrehumbert**, J. 2001. Stochastic phonology. *Glot International* 5: 195–207. // **Prieto**, P. 2006. Phonological phrasing in Spanish. In: *OT Advances in Spanish Phonology*, Colina, S., Martínez-Gil, F. (eds) Amsterdam: Benjamins, pp. 39-60.

Individual variation and the prosody of Spanish Clitic Left-Dislocations

Ingo Feldhausen
Goethe-Universität Frankfurt am Main
Laboratoire de Phonétique et Phonologie (LPP), Sorbonne-Nouvelle – Paris 3

Workshop 'Syntax-Phonology Interface from a Cross-linguistic Perspective'
November 2-3, 2012 - ZAS, Berlin

feldhausen@em.uni-frankfurt.de

1. Introduction

- Goal of the talk:
  To highlight the individual differences in the production of the prosodic phrasing of Clitic Left-Dislocations (CLLD) in Spanish and to propose a modified version of the Stochastic Optimality Theory (SOT) to account for the attested inter-speaker variation
- Variation and linguistic competence
- Production experiment
  - Embedded (and non-embedded) CLLD in Spanish
  - There are obligatory and optional boundaries
  - Optional boundaries shows inter-speaker variation
- Analysis
  - Modified version of Stochastic Optimality Theory (Boersma & Hayes 2001)
  - Basic constraints for Spanish SVO phrasing (Prieto 2006)
  - New constraint: ALIGN-TOP,R and ALIGN-CP,L
- Conclusion

1. Introduction

- Prosodic aspects differ between languages (e.g. Jun 2005, 2012)
  - intonational phrasing
  - boundary realization
  - pitch accent realization
- Growing amount of evidence questions whether these aspects are used homogenously within a given language (e.g. Grabe 2002, Féry 2004, Feldhausen 2010, 2011, Myrberg 2010, Niebuhr et al. 2011)
- Only few studies addressing **free variation in prosody** and especially in intonational phrasing within a population

Free variation: Variant A and variant B appear in the same context without changing the meaning of the utterance, i.e. they are optional realizations. The occurrences of the variants cannot be related to factors such as style, genre, syntactic or prosodic complexity

1. Introduction

- Many studies investigate aspects of determining boundary placement (e.g. Cooper & Paccia-Cooper 1980, Gee & Grosjean 1983, Selkirk 1984 and subsequent work, Nespor & Vogel 1986/2007, Ghini 1993, Sandalo & Truckenbrodt 2002, D'Imperio et al. 2005, Elordieta et al. 2003, 2005)
- But: variation within a population is typically ignored

WHY?

- Pierrhumbert (2001:195):
  - In generative phonology, accounts modeling linguistic competence are non-probabilistic, and for this reason any given sequence is **either grammatical or completely impossible**
  - As a consequence, **variation has largely been ignored or eliminated**:
    Variation in observed data typically has not been taken to reflect linguistic competence but rather has been attributed to differences in performance.

1. Introduction

Variation is an intrinsic part of linguistic competence (Pierrehumbert 2001)

- Pierrehumbert (2001:195) claims that the learner develops a cognitive system in which frequency information plays a central role.
- **Cognitive system remains grammatical**: It establishes the "well-formedness" of complex forms and has the power to create and process completely novel forms.
- **Grammar is probabilistic**: It maintains frequency distributions.
- The frequency of a given unit is an important factor in how it behaves in the system (see also Frisch 2000, Goldinger 2000 or Kirchner 2002).

1. Introduction

- Several (Optimality Theoretic) models exist for capturing variation in Generative Phonology:
  - Reynolds & Nagy's (1994) *floating constraints,*
  - Anttila's (1997, 2002) *partial ordered grammar* or *stratified grammar,*
  - Truckenbrodt's (2002) *output-to-output faithfulness,* and

Boersma & Hayes' (2001) ***Stochastic Optimality Theory (SOT)***

- In SOT, constraints do not need to be categorically ranked
- Grammar allows for a certain amount of overlapping (which is to be determined experimentally)
- Variation in output results from the different rankings produced by the probabilistic grammar
- SOT easily accounts for variation in a homogenous group

1. Introduction

- But what about individual differences?
- If there is variation within a homogenous group, there must be variation between speakers => individual variation / inter-speaker variation

"Variation is idiosyncratic and inherent in individual grammars" (Bresnan et al. 2007:340)

- A grammar should not abstract away systematic individual differences
- Present study:
  Individual variation in the prosodic phrasing of sentences with (embedded) clitic left-dislocations in Spanish

2. Experiment

- **Clitic Left-Dislocations (CLLD)** are "characterized by the presence of a phrase in the first position of the clause which is connected with that clause through the intermediary of some anaphoric element". (Alexiadou 2006)
- CLLD expresses the informational status of givenness
  (López 2009, see p. 39 for details on the informational structural status)

CLLD in simple clauses

(2) [**El águila de Málaga**, *la* vendió mi hermano]$_{CP1}$
'The eagle of Málaga, my brother sold (it)'

CLLD in embedded clauses

(3) [Bárbara supone [que **el águila**, *la* vendió su hermano]$_{CP2}$]$_{CP1}$
'Barbara assumes that the eagle, her brother sold (it)'

2. Experiment

Why Clitic Left-Dislocations in Spanish?

- Much work has been devoted to the syntax and the information structure of CLLD in Spanish (Rivero 1980, Silva Corvolán 1984, Casielles-Suárez 2003, Suñer 2006, López 2009, among others)
  Up to now, no detailed study on the prosody of CLLD in Spanish
  (but see Feldhausen 2012 for the study presented here)
- Combination of "obligatory" and "less obligatory" prosodic boundaries;
  Prosody of CLLD known for other Romance languages:
  - French (Delais-Roussarie et al. 2004, Avanzi 2012)
  - Italian (Gili-Fivela 1999, Frascarelli 2000)
  - Catalan (Feldhausen 2010)

  => Obligatory boundary at right edge of CLLD constituents, but less obligatory boundaries in other positions.

2. Experiment

- **Production experiment based on scripted speech**
- Locus:
  - Murcia (Spain)
- Subjects:
  - Four subjects (three female, one male), 27 to 31 yoa
  - (Monolingual) Native speakers of Peninsular Spanish from Murcia city
  - subjects held a university degree
  - Totally naïve to the purpose of the experiment

(Figure from: Wikipedia „Murcia (Region)", 10.28.12

2. Experiment

- Material
  - 144 sentences (12 sentences x 3 repetitions x 4 speakers)
    - Embedded and non-embedded CLLD
    - CLLD : one ω and two ω ω
  - 72 filler sentences (6 sentences X 3 repetitions x 4 speakers)
  - Context question to guarantee givenness of CLLD constituent:

- ***Context***

  (4) Qué pasó con el águila que me compré en Málaga? Dónde esta?
  'What happened to the eagle I bought in Malaga? Where is it?'

- ***Target sentence***

  ω ω
  (5) [Bárbara supone [que **el águila de Málaga** la vendió su hermano]]
  B. assume.3SG that the eagle of M. CL sold his/her brother
  'Barbara assumes that the eagle of Málaga, his brother sold.'

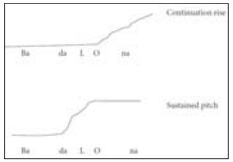
2. Experiment

- Procedure
  - Subjects were recorded at their homes
  - Stimuli were presented in a pseudo-randomized order on sheets of paper (ca. eight Q/A pairs per sheet)
  - Three repetitions; pauses between each block
  - Subjects were asked to read the stimuli out loud at a normal rate of speech only after having silently read the Q/A pair
  - Sentences lacked orthographically typical commas to avoid an induced comma reading
  - Small practice session at the beginning
- Prosodic analysis is based on the ToBI system for Spanish - Sp_ToBI
  (Aguilar et al. 2009; Prieto & Roseano 2010)

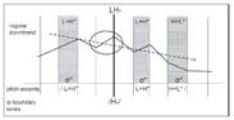
2. Experiment

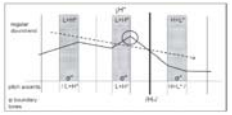
1. Introduction / 3. Analysis 4. Conclusion

- Boundary cues
  (based on Frota et al. 2007, Feldhausen et al. 2010, Gabriel et al. 2011)
  a) Continuation Rise, CR (H-)
  b) Sustained Pitch, SP (!H-)
  c) Pitch Reset
  d) Preboundary Upstep, PU
  e) Low tones
  f) Pauses
  g) Complex boundary tone, CBT (L-H%)



CR & SP (Frota et al. 2007: 135)


CBT (Gabriel et al. 2011: 166)


PU (Gabriel et al. 2011: 166)

13

2. Experiment

1. Introduction / 3. Analysis 4. Conclusion

- Material preceding CLLD (including the complementizer *que*) is separated by a prosodic break at the ip level (99%)

- Spanish CLLD are almost always separated from the subsequent material by a prosodic break (97%)
- High boundary tone marking the edge of an intermediate phrase (~90%)
- Break can be accompanied by a pause (if so, IP boundary; 23%)
- Hardly any reconstruction with following material (if so, only with non-branching CLLD constituents)

25% ~100% ~100%
( )( )( )( )
(6) [Bárbara supone [que **el águila de M.** *la* vendió su hermano]$_{CP2}$]$_{CP1}$
S V q CLLD V S
'Barbara assumes that the eagle of Málaga, her/his brother sold (it)'

14

2. Experiment

1. Introduction / 3. Analysis 4. Conclusion

- Typically, matrix subject and matrix verb constitute one single prosodic unit (75%); in contrast to sentences without CLLD (D'Imperio et al. 2005; Elordieta et al. 2003, 2005; Feldhausen et al. 2010)
- Separated matrix subject only in 25% of the cases
- Great variation between speakers

S)(V	Speaker
0%	A
33%	B
28%	C
39%	D
25%	Total

Table 1: Percentages of individual realizations of the 'S)(V'-boundary across target sentences with embedded CLLD

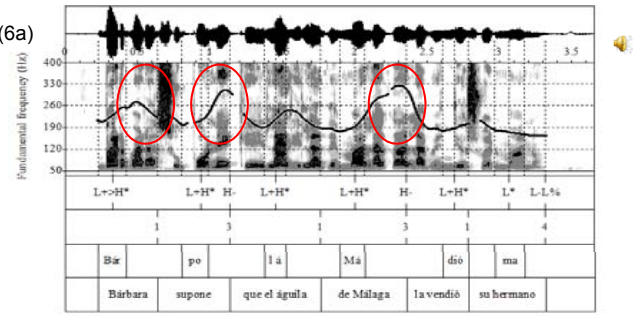
25% ~100% ~100%
( )( )( )( )
(6) [Bárbara supone [que **el águila de M.** *la* vendió su hermano]$_{CP2}$]$_{CP1}$
S V q CLLD V S
'Barbara assumes that the eagle of Málaga, her/his brother sold (it)'

15

2. Experiment

1. Introduction / 3. Analysis 4. Conclusion

(6a)



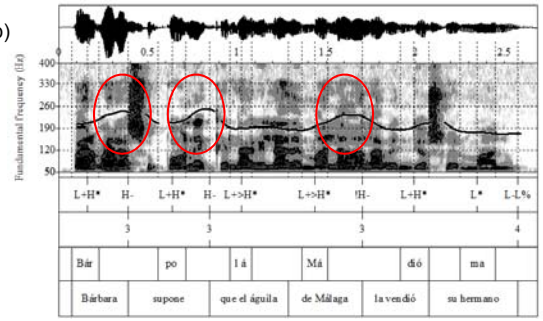
( )( )( )( )
[Bárbara supone [que **el águila de M.** *la* vendió su hermano]$_{CP2}$]$_{CP1}$
S V q CLLD V S
'Barbara assumes that the eagle of Málaga, her/his brother sold (it)'

16

2. Experiment

1. Introduction / 3. Analysis 4. Conclusion

(6b)



( )( )( )( )

[Bárbara supone [que **el águila de M.** *la* vendió su hermano]$_{CP2}$]$_{CP1}$

S V q CLLD V S

'Barbara assumes that the eagle of Málaga, her/his brother sold (it)'

17

2. Experiment

1. Introduction / 3. Analysis 4. Conclusion

- (6a,b) show that only some boundaries are optional, while others are obligatory
  - Obligatory boundaries around CLLD constituent
  - Optional boundary after matrix subject in complex clauses (with CLLD)
  - Simple SVO sentences in Spanish: Obligatory boundary after the subject (D'Imperio et al. 2005; Elordieta et al. 2003, 2005; Feldhausen et al. 2010)
  - "Impact of prosodic boundaries depends on the other prosodic choices a speaker has made" (Frazier et al. 2006:244)
    => CLLD boundary is grammatically required, while boundary after S is less important
- The degree of optionality differs across speakers and can hardly be caught by the average %
- Nevertheless, the frequency-dependent variation is part of the individual's linguistic competence

18

3. Analysis

1. Introduction 2. Experiment / 4. Conclusion

- Proposal of a formal analysis in the framework of Optimality Theory (OT, Prince & Smolensky 1993/2004).
- More exactly, in one of its extensions: the Stochastic Optimality Theory (SOT, Boersma & Hayes 2001)
- Basic idea of OT:
  - Universal Grammar is an ensemble of contraints for the well-formedness of representations
  - OT supposes that the constraints are universal, but that they are in conflict with one another (since satisfying one constraint generally means violating another constraint)
  - A grammar consists of both constraints and a general method to solve the conflicts
  - The attested representations are thus those which satisfy the conflicting constraints in the best possible way

19

3. Analysis

1. Introduction 2. Experiment / 4. Conclusion

- Classical OT
  (Prince & Smolensky 1993/2004)

Strict order of constraints
(the highest constraints in the hierarchy precede the lower constraints, which in turn are located at the right hand side of the table)

		Constraint C_1	Constraint C_2
a.	Candidate 1	*!	
b. ☞	Candidate 2		*

The optimal candidate
(corresponds to the surface form; it is indicated by ☞)

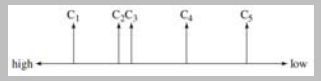
Violation of a constraint: *
(fatal violation, i.e. a violation which eliminates a candidate from the competition is marked by: !)

20

3. Analysis

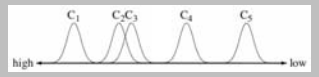
- **OT**: a grammar is a set of strictly ranked constraints
- **Non-stochastic OT**: ordinal ranking

  $C_1 >> C_2 >> C_3 >> C_4 >> C_5$
- **Stochastic OT**: ranking along a continues scale



- Evaluation time:

  evaluation ranking = grammatical ranking + noise



strict ranking determines winning candidate form

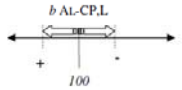
90% of the time: $C_2 >> C_3$

10% of the time: $C_3 >> C_2$

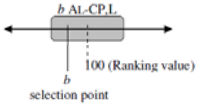
(taken from Boersma 2003)

3. Analysis

(35) Pertubation of constraint by a random positive or negative value



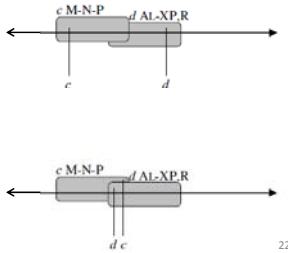
(36) Range of value and selection point



(37) Overlapping constraints

Upper panel: normal ranking

Lower panel: reverse ranking



3. Analysis

- **ALIGN-XP,R**: For each XP there is a P such that the right edge of XP coincides with the right edge of P (P = prosodic phrase)

(S)(V O)

[[Javier]$_{NP}$ [visitó [Galicia]$_{NP}$]$_{VP}$]$_{IP/CP}$	ALIGN-XP,R	MAX-BIN	MIN-BIN
a. ( )φ	*!	*	
☞ b. ( )φ ( )φ			*
c. ( )φ ( )φ	*!		*

(taken from Prieto 2006: 50)

– Selkirk's (1986) classical end-based theory, which is generalized to the format of Generalized Alignment (McCarthy & Prince 1993) in Optimality Theory in the 1990s (see Selkirk 1995)

3. Analysis

- **MIN-N-PHRASES**: Minimize the number of prosodic phrases
  - The constraint is part of the constraint family *STRUC that ensures that structure is constructed minimal (Prince & Smolensky (1993:25, fn.13)
  - It seeks to avoid prosodic phrases altogether (cf. also Truckenbrodt 1999:228; Féry 2007)
- **MIN-BIN**: P-phrases should consist of minimally two ω (Prieto 2006: 45)

  => $((\omega\ \omega)_{PhP})_{IP}$ is better than $((\omega)_{PhP}\ (\omega)_{PhP})_{IP}$
  - Minimality size effects

3. Analysis

- **MAX-BIN (IP Head)**: A phonological phrase which is the head of an IP constituent must be binary (at the ω level). (Prieto 2006: 52)

(ω ω)

[Compraba [mapas [de Barcelona] $_{PP}$] $_{NP}$] $_{VP}$	MAX-BIN (IP Head)	ALIGN-XP,R	MIN-BIN
a. ( )φ	*!		
☞ b. ( )φ ( )φ			*
☞ c. ( )φ ( )φ			*

(taken from Prieto 2006: 56)

- Size constraint
- Based on Sandalo & Truckenbrodt (2002:295) who introduce size constraints for prosodic constituents in Brazilian Portuguese (BP) [their constraint is based on Ghini (1993)]

3. Analysis

- **ALIGN-CP,L** (Feldhausen 2010: 113; 2011: 1954)

  ALIGN(CP,Left;PrP,Left), or align the left edge of a CP to the left edge of a prosodic phrase (PrP).

  => accounts for the prosodic break preceding the embedded clause

  - Universality of constraints (de Lacy 2003): in all grammars all constraints are present.
    A grammar cannot chose between, for example, ALIGN (XP,Left; ϕ,Left) and ALIGN (XP,Right; ϕ,Right)
    => both constraints are present in every grammar (de Lacy 2003:60)
  - He gives evidence from the Polynesian language Māori: There is both right and left alignment in one single language.

3. Analysis

- **ALIGN-TOP,R** (Feldhausen 2010: 174)

  Align the right edge of a topic constituent to the right edge of a prosodic phrase (PrP).

  => accounts for the prosodic break immediately after the CLLD constituent

  - Constraint of the alignment family (McCarthy & Prince 1993)
  - Advantages over Frascarelli's (2000: 63) *Topic Prosodic Domain*, since ALIGN-TOP,R does not predict a left edge
    (In Catalan, embedded CLLD constituents phrase with the preceding matrix clause (unlike Spanish), see Feldhausen 2010:ch.5)
  - See also Féry (2011: 1910) for an OT constraint rather similar to Frascarelli's (2000) rule, predicting a left and a right edge

3. Analysis

(1) Normal ranking

ALIGN-TOP,R >> MAX-BIN (IP Head) >> AL-CP,L >> MIN-N-P >> AL-XP,R >> MIN-BIN

S V q CLLD V O	ALIGNTOP,R	MAX-BIN (IP Head)	ALIGN-CP,L	MIN-N-PHRASES	ALIGN-XP,R	MIN-BIN
(S V q CLLD V O)	*!	*	*	*	**	
(S V q CLLD) (V O)			*!	**	*	
☞ (S V) (q CLLD) (V O)				***	*	*
(S)(V) (q CLLD) (V O)				****!		***
(S V q CLLD) (V)(O)			*!	***	*	**
(S V) (q CLLD) (V)(O)				****!	*	***

75% ~100% ~100%

( )( )( )

(6) [Bárbara supone [que **el águila de M.** *la* vendió su hermano]$_{CP2}$]$_{CP1}$

S V q CLLD V S

'Barbara assumes that the eagle of Málaga, her/his brother sold (it)'

3. Analysis

(2) Reverse ranking

ALIGN-TOP,R >> MAX-BIN (IP HEAD) >> AL-CP,L >> AL-XP,R >> MIN-N-P >> MIN-BIN

S V q CLLD V O	ALIGNTOP,R	MAX-BIN (IP Head)	ALIGN-CP,L	ALIGN-XP,R	MIN-N-PHRASES	MIN-BIN
(S V q CLLD V O)	*!	*	*	**	*	
(S V q CLLD)(V O)			*!	*	**	
(S V)(q CLLD)(V O)				*!	***	*
☞ (S)(V)(q CLLD)(V O)					****	***
(S V q CLLD)(V)(O)			*!	*	***	**
(S V)(q CLLD)(V)(O)				*!	****	***

25% ~100% ~100%

( )( )( )( )

(6) [Bárbara supone [que **el águila de M.** *la* vendió su hermano]$_{CP2}$]$_{CP1}$

S V q CLLD V S

'Barbara assumes that the eagle of Málaga, her/his brother sold (it)'

3. Analysis

- Up to now:

UNDERLYING CONSTRAINT HIERARCHY
ALIGN-TOP,R >> MAX-BIN$_{(IP\ HEAD)}$ >> ALIGN-CP,L >> MIN-N-PHRASES >> ALIGN-XP,R >> MIN-BIN

  - Stochastic OT
  - Two overlapping constraints: MIN-N-PHRASES & ALIGN-XP,R
  - They have a certain distance on the continuous ranking scale (which is experimentally defined, see the *Gradual Learning Algorithm* by Boersma & Hayes 2001)
  - This distance / amount of overlap guarantees the frequency distribution of the whole group of speakers (75% vs. 25%)

- But what about the individual differences?

3. Analysis

Modeling individual variation in prosody

- The constraint hierarchy of a grammar is invariant (as in other OT accounts)
- The degree of overlap is not fixed for the whole population
- The distance between the two constraints MIN-N-PHRASES & ALIGN-XP,R can differ between speakers:

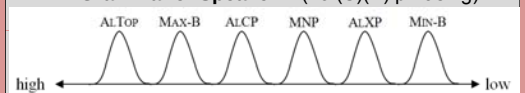
Each speaker has a different degree of overlap

- Thus, differences in the frequency of output forms between speakers arise due to different degrees of overlap of MIN-N-PHRASES & ALIGN-XP,R, while keeping the underlying constraint hierarchy unchanged

3. Analysis

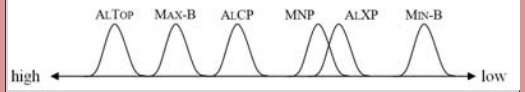
Grammar of Spanish (underlying constraint hierarchy)

Grammar of Speaker A (no (S)(V) phrasing):



strict ranking determines winning candidate form
100% of the time: MIN-N-P >> AL-XP,R (= Table 1) 0% of the time: AL-XP,R >> MIN-N-P

Grammar of speaker D:



strict ranking determines winning candidate form
61% of the time: MIN-N-P >> AL-XP,R (= Table 1)
39% of the time: AL-XP,R >> MIN-N-P (= Table 2)

4. Conclusion

- Spanish CLLDs have an obligatory right boundary (as the other Romance languages)
- In embedded contexts, there is a boundary at the left edge of the embedded clause (in contrast to Catalan)
- There are optional sentence-internal boundaries
- This optionality allows for inter-speaker differences
- "Variation is idiosyncratic and inherent in individual grammars" (Bresnan et al. 2007:340)
- Slightly modifying the basic assumptions of Stochastic OT allows for accounting for inter-speaker variation
- Individual grammars are built based on the grammar of a given language L

4. Conclusion

- Prediction:
  Underlying order of constraints is fixed for all speakers of a population (i.e. there is no speaker who realizes the "reversed order" more often than the "normal order")
  - Conduct studies with more speakers
- What does it mean for a grammar of a given language if a speaker of that homogenous group realizes the "reversed ranking" more often?
  - Can there be two grammars of a given language?
  - If so, how does a grammar of a language with all its variability differs from the grammar of another language?
  - Or: Do we really deal with free variation? Aren't there rather influencing factors which have not been considered yet?

¡Muchas gracias!

Thank you for your attention

Special thanks:

The Volkswagen Foundation
The Andrew W. Mellon Foundation

References

Aguilar, Lourdes, de la Mota, Carme & Prieto, Pilar. 2009. Sp_ToBI. Training materials. <http.//prosodia.upf.edu/sp_tobi/en/> (26 January 2012).

Alexiadou, Artemis. 2006. Left dislocation (including CLLD). In *The Blackwell Companion to Syntax* - Vols. I-V, Marti Everaert, Henk van Riemsdijk, Rob Goedemans & Bart Hollebrandse (eds), 668-699. Oxford: Blackwell.

Anttila, Arto. 1997. Deriving variation from grammar: A study of Finnish genitives. In *Variation, Change and Phonological Theory* [Current Issues in Linguistic Theory 146], Frans Hinskens, Roeland van Hout & Leo Wetzels (eds), 35–68. Amsterdam: John Benjamins.

Anttila, Arto. 2002. Variation in phonological theory. In *The Handbook of Language Variation and Change,* Jack K. Chambers, Peter Trudgill & Nathalie Schilling-Estes (eds), 206–243. Oxford: Blackwell.

Avanzi, Mathieu. 2012. *L'interface prosodie/syntaxe en français*. Peter Lang.

Boersma, Paul. 2003. Stochastic Optimality Theory. Talk at LSA.

Boersma, Paul & Hayes, Bruce. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32: 45–86.

Bresnan, Joan et al. 2007. Typology in Variation. *English Language & Literature* 11: 301-346.

References

Casielles-Suárez, Eugenia. 2003. Left-Dislocated Structures in Spanish. *Hispania* 86(2): 326-338.

Cooper, William E. & Paccia-Cooper, Jeanne. 1980. *Syntax and Speech*. Cambridge, MA: Harvard University Press.

de Lacy, P., 2003. Constraint universality and prosodic phrasing in Māori. In: Carpenter, A., Coetzee, A., de Lacy, P. (Eds.), *Papers in Optimality Theory II* [UMOP 26]. GLSA, Amherst MA, pp. 59–79.

Delais-Roussarie, Elisabeth, Doetjes, Jenny & Sleeman, Petra. 2004. Dislocations in French. In *Handbook of French semantics*, Corblin Francis & de Swart, Henriëtte (eds), pp.501-528. Stanford: CSLI.

D'Imperio, Mariapaola, Elordieta, Gorka, Frota, Sonia, Prieto, Pilar & Vigario, Marina. 2005. Intonational phrasing in Romance: The role of syntactic and prosodic structure. In: *Prosodies: With Special Reference to Iberian Languages,* Sonia Frota, Marina Vigario & Maria Joao Freitas (eds), 59–97. Berlin: Walter de Gruyter.

Elordieta, Gorka, Frota, Sonia, Prieto, Pilar & Vigario, Marina. 2003. Effects of constituent weight and syntactic branching on intonational phrasing in Ibero-Romance. In: *Proceedings of the 15th International Congress of Phonetic Sciences*, Maria-Josep Sole, Daniel Recasens & Joaquin Romero (eds), 487–490. Barcelona.

References

Elordieta, Gorka, Frota, Sonia & Vigario, Marina. 2005. Subjects, objects and intonational phrasing in Spanish and Portuguese. *Studia Linguistica* 59(2–3): 110–143.

Feldhausen, Ingo. 2010. *Sentential form and prosodic structure of Catalan* [Linguistik Aktuell / Linguistics Today 168]. Amsterdam: John Benjamins.

Feldhausen, Ingo. 2011. The prosodic phrasing of sentential objects. *Lingua* 121 (13): 1934-1964.

Feldhausen, Ingo. 2012. Prosodic Aspects of Clitic Left-Dislocations in Spanish. Submitted to *Probus* (Ms. Goethe-Universität Frankfurt, LPP – Paris 3).

Feldhausen, Ingo; Christoph Gabriel & Andrea Pešková. 2010. Prosodic Phrasing in Argentinean Spanish: Buenos Aires and Neuquén. *Speech Prosody* 2010, Chicago.

Féry, Caroline. 2004. Gradient prosodic correlates of phrasing in French. In: Meisenburg, Trudel & Selig, Maria (eds.): *Nouveaux départs en phonologie : Les conceptions sub- et suprasegmentales.* Tübingen: Narr, pp. 161-182.

Féry, Caroline. 2007. The prosody of topicalization. In: Schwabe, K., Winkler, S. (Eds.), *On Information Structure, Meaning and Form: Generalizations Across Languages* [Linguistik Aktuell/Linguistics Today 100]. John Benjamins, Amsterdam, pp. 69–86.

Féry, Caroline. 2011. German sentence accents and embedded prosodic phrases. *Lingua* 121(13): 1906-1922.

References

Frascarelli, Mara. 2000. *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian*. Kluwer: Dordrecht.

Frazier, Lyn, Carlson, Katy & Clifton, Charles Jr.. 2006. Prosodic phrasing is central to language comprehension. *Trends in Cognitive Science* 10(6):244-249.

Frisch, Stefan. 2000. Temporally organized lexical representations as phonological units. In Broe & Pierrehumbert (Eds,) Papers in Laboratory Phonology III. Cambridge: Cambridge University Press, pp. 283-298.

Frota, Sonia, D'Imperio, Mariapaola, Elordieta, Gorka, Prieto, Pilar & Vigario, Marina. 2007. The phonetics and phonology of intonational phrasing in Romance. In *Segmental and Prosodic Issues in Romance Phonology*, Pilar Prieto, Joan Mascaro & Maria-Josep Sole (eds), 131–153. Amsterdam: John Benjamins.

Gabriel, Christoph, Feldhausen, Ingo & Pešková, Andrea. 2011. Prosodic phrasing in Porteño Spanish. In *Intonational phrasing in Romance and Germanic* [Hamburg Studies on Multilingualism 10], C. Gabriel & C. Lleó (eds), 153−182. Amsterdam: John Benjamins.

Gee, James P. & Grosjean, François. 1983. Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, 15: 411-458.

Ghini, Mirco. 1993. ϕ-Formation in Italian: A new proposal. *Toronto Working Papers in Linguistics* 12: 41–79.

References

Gili Fivela, Barbara. 1999. The Prosody of Left-Dislocated Topic Constituents in Italian Read Speech. *Eurospeech* 1999: 531-534.

Goldinger, Stephen D. 2000. The role of perceptual episodes in lexical processing. In Cutler, A., McQueen, J. M. and Zondervan, R. (Eds.) *Proceedings of SWAP (Spoken Word Access Processes)*, Nijmegen, Max Planck Institute for Psycholinguistics, pp. 155-159.

Grabe, Esther. 2002. Variation adds to prosodic typology. *Speech Prosody*, Aix-en-Provence.

Jun, Sun-Ah. 2005. *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: OUP.

Jun, Sun-Ah. 2012. Prosodic Typology Revisited: Adding Macro-rhythm. In: *Proceedings of Speech Prosody*, 6th international conference, Shanghai, China.

Kirchner, Robert. 2002. Preliminary thoughts on "phonologisation" within an exemplar-based speech processing model. *UCLA Working papers*. Volume 6.

López, Luis. 2009. *A Derivational Syntax for Information Structure*. Oxford: OUP.

McCarthy, John & Prince, Alan. 1993. Generalized alignment. In G. Booij and J. van Marle, eds.. *Yearbook of Morphology*, 79-153. Dordrecht: Kluwer.

Myrberg, Sara. 2010. *The Intonational Phonology of Stockholm Swedish* [Acta Universitatis Stockholmiensis]. PhD Dissertation, University of Stockholm.

References

Nespor, Marina & Vogel, Irene. 1986/2007. *Prosodic Phonology*. Berlin: Mouton de Gruyter (1986: Dordrecht: Foris).

Niebuhr, Oliver, D'Imperio, Mariapaola, Gili Fivela, Barbara & Cangemi, Francesco. 2011. Are there "shapers" and "aligners"? Individual differences in signalling pitch accent category. *Proceedings of the 17th ICPhS, Hong Kong, China*, 120-123.

Pierrehumbert, Janet. 2001. Stochastic phonology. *Glot International* 5(6): 195–207.

Prieto, Pilar. 2006. Phonological phrasing in Spanish. In *Optimality-Theoretic Advances in Spanish Phonology* [Linguistik Aktuell/Linguistics Today 99], Sonia Colina & Fernando Martínez-Gil (eds), 39–60. Amsterdam: John Benjamins.

Prieto, Pilar & Roseano, Paolo (Eds.). 2010. *Transcription of Intonation of the Spanish Language*. München: Lincom.

Prince, Alan & Smolensky, Paul. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar* [Rutgers University Center for Cognitive Science Technical Report 2]. (Published in 2004: Oxford: Blackwell).

Reynolds, William & Nagy, Naomi. 1994. Phonological variation in Faetar: An optimality account. *Chicago Linguistic Society (Variation and Linguistic Theory)* 30(2): 277–292.

Rivero, María-Luisa 1980. On Left-Dislocation and Topicalization in Spanish. *Linguistic Inquiry* 11(2): 363-393.

References

Sandalo, Filomena & Truckenbrodt, Hubert. 2002. Some notes on phonological phrasing in Brazilian Portuguese. *MIT Working Papers in Linguistics* 42: 285–310.

Selkirk, Elisabeth. 1984. *Phonology and Syntax. The Relation between Sound and Structure.* Cambridge MA: The MIT Press.

Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology* 3, 371–405.

Selkirk, Elisabeth. 1995. The prosodic structure of function words. In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, James L. Morgan & Katherine Demuth (eds), 187–214. Mahwah NJ: Lawrence Erlbaum Associates.

Silva Corvolán, Carmen. 1984. Topicalización y pragmática en español. *Revista Española de Lingüística* 14(1): 1-20.

Suñer, Margarita. 2006. Left dislocations with and without epithets. *Probus* 18(1): 127-158.

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30: 219–255.

Truckenbrodt, Hubert. 2002. Variation in P-phrasing in Bengali. In: *Linguistic Variation Yearbook* 2, Pierre Pica (ed.), 259–303. Amsterdam: John Benjamins.