

**ANALYZING INTERACTING
PHENOMENA
WORD ORDER AND NEGATION IN
BASQUE**

J Crowgey and Emily M. Bender

University of Washington

25 August 2011

18th International Conference on HPSG

TABLE OF CONTENTS

introduction

overview

Basque [eus]

who,what,where

syntax

word order

overview

aux-first types

verb-first types

negation

conclusions and outlook

NEGATION, BASQUE AND GRAMMAR ENGINEERING

OVERVIEW

- negation in Basque:
 - ordering of major constituents is quite free
 - but negation constrains possible word orders
- we have negation:
 - Kim (2000) examines negation lit, proposes types for HPSG
 - morphological marking
 - syntactic marking
- we have free word order:
 - Fokkens (2010)

NEGATION, BASQUE AND GRAMMAR ENGINEERING

OVERVIEW

- will existing analyses of negation and free word order interact correctly to capture the natural language patterns of Basque?
- the methodology:
 - grammar engineering: implement your analysis, test it
 - open source tools:
 - LKB (Copestake 2002)
 - [incr tsdb()] (Oepen & Flickenger 1998) grammar development platform
 - Grammar Matrix customization system (Bender et al. 2002; 2010)
- we find: **construction types motivated to account for word order in Basque provide the proper analytical division to account for word order under negation patterns**

SYNTACTIC FACTS OF BASQUE

- ergative-absolutive (S=O)
- rich system of agreement markers expressed on the finite element of the clause
- most lexical verbs in Basque cannot be *finite*
- typical (minimal) clause has as least three elements: subject, lexical verb (LV), auxiliary verb (Aux)

Miren ibilli da
 Mary.ABS walk.PERF 3.SG.S.PRES ¹
Mary has walked. [eus]

¹data here and below adapted from (Manandise 1988)

WORD ORDER

- major constituent order is **nearly** free
 - a pragmatic constraint:
 - element in preverbal (LV) position is in focus
 - focused element traditionally termed *galdegaia* “object of inquiry”

- a. Liburu bat nork irakurri du?
book one.ABS.SG who.ERG.SG.FOC read.PERF 3.SG.O.PRES.3.SG.A
Who has read one book? [eus]
- b. Liburu bat Mirenek irakurri du.
book one.ABS.SG Mary.ERG.SG.FOC read.PERF 3.SG.O.PRES.3.SG.A
Mary has read one book. [eus]
- c. Mirenek liburu bat irakurri du.
Mary.ERG.SG book one.ABS.SG.FOC read.PERF 3.SG.O.PRES.3.SG.A
Mary has read one book. [eus]

WORD ORDER

- major constituent order is **nearly** free
 - a pragmatic constraint:
 - element in preverbal (LV) position is in focus
 - focused element traditionally termed *galdegaia* “object of inquiry”
 - only (b) is an acceptable answer to (a)
- a. Liburu bat nork irakurri du?
book one.ABS.SG who.ERG.SG.FOC read.PERF 3.SG.O.PRES.3.SG.A
Who has read one book? [eus]
- b. Liburu bat Mirenek irakurri du.
book one.ABS.SG **Mary.erg.sg.foc** read.PERF 3.SG.O.PRES.3.SG.A
Mary has read one book. [eus]
- c. Mirenek liburu bat irakurri du.
Mary.ERG.SG **book one.abs.sg.foc** read.PERF 3.SG.O.PRES.3.SG.A
Mary has read one book. [eus]

MANANDISE'S FILTER

a syntactic constraint on word order

If the lexical verb is to the left of the auxiliary, then the lexical verb must be left-adjacent to the auxiliary. (Manandise 1988, 15)

*Liburu irakurri Mirenek du.
 book.ABS.SG read.PERF Mary.ERG.SG AUX
Mary has read a book. [eus]

NP	NP	V	Aux	NP	NP	Aux	V
*NP	V	NP	Aux	NP	V	Aux	NP
*V	NP	NP	Aux	*V	NP	Aux	NP
NP	Aux	NP	V	Aux	NP	NP	V
NP	Aux	V	NP	Aux	NP	V	NP
V	Aux	NP	NP	Aux	V	NP	NP

TWO CLASSES OF POSSIBLE SENTENCES

- Manandise's filter suggests a bifurcation of *a priori* sentence types
 - aux-first types
 - free word order
 - verb-first types
 - no interveners

NP	NP	V	Aux	NP	NP	Aux	V
*NP	V	NP	Aux	NP	V	Aux	NP
*V	NP	NP	Aux	*V	NP	Aux	NP
NP	Aux	NP	V	Aux	NP	NP	V
NP	Aux	V	NP	Aux	NP	V	NP
V	Aux	NP	NP	Aux	V	NP	NP

TWO CLASSES OF POSSIBLE SENTENCES

- Manandise's filter suggests a bifurcation of *a priori* sentence types
 - aux-first types
 - free word order
 - verb-first types
 - no interveners

NP	NP	V	Aux	NP	NP	Aux	V
*NP	V	NP	Aux	NP	V	Aux	NP
*V	NP	NP	Aux	*V	NP	Aux	NP
NP	Aux	NP	V	Aux	NP	NP	V
NP	Aux	V	NP	Aux	NP	V	NP
V	Aux	NP	NP	Aux	V	NP	NP

TWO CLASSES OF POSSIBLE SENTENCES

- Manandise's filter suggests a bifurcation of *a priori* sentence types
 - aux-first types
 - free word order
 - verb-first types
 - no **interveners**

NP	NP	V	Aux	NP	NP	Aux	V
NP	V	Aux	NP	NP	Aux	NP	V
V	Aux	NP	NP	NP	Aux	V	NP
*NP	V	NP	Aux	Aux	NP	NP	V
*V	NP	Aux	NP	Aux	NP	V	NP
*V	NP	NP	Aux	Aux	V	NP	NP

AUX-FIRST TYPES

OVERVIEW

free word order

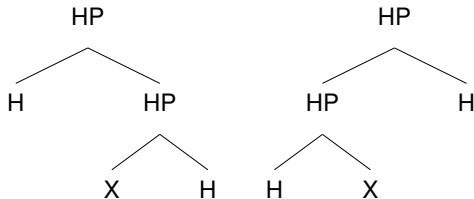
- allow free word order without creating spurious ambiguity
- no ID-LP split
- our approach starts with Fokkens (2010):
 - head-final and head-initial versions of head-nexus rules
 - apply any head-initial rules before (“lower”) than any head-final rules
- additional measures:
 - agreement and the checking off of valence lists (Aux, Verb, NP)

BINARY BRANCHING RULES

- rules must be specified for arity and order of daughters
- capturing all permutations of $\{NP_1, NP_2, V\}$ requires 6 rules; more generally, permutations of n elements will require at least $n!$ rules
- a binary branching analysis with a projecting headpath captures all permutations of $\{NP_1, NP_2, V\}$, with only 4 rules
 - subj-head
 - head-subj
 - comp-head
 - head-comp
- optimizing on the size of the grammar — this grammar is maintained by hand

FREE WORD ORDER

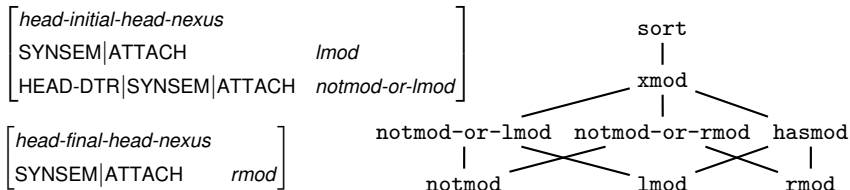
- simply providing head-init and head-final versions of combinatory rules leads to massive spurious ambiguity
- both these parses yield identical MRS structures:



HEAD-INITIAL AND HEAD-FINAL

xmod hierarchy

- phrasal rules annotated to pass [ATTACH *xmod*]



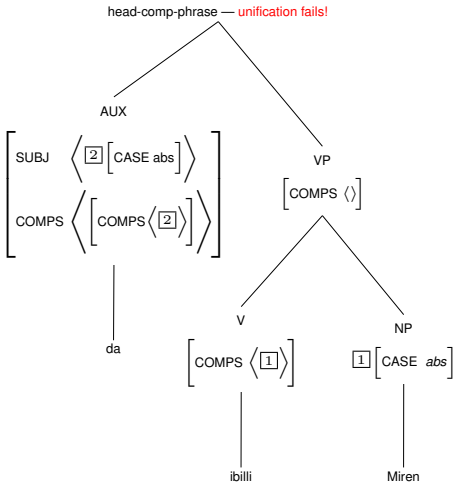
FREE WORD ORDER

- potential for spurious ambiguity on form-types like: *Aux, V, NP*
- using only head-initial types, there is a potential for two derivation trees with equivalent semantic representations
- our approach:
 - *Aux* must know about the case and PNG information of argument NPs
 - argument composition *Auxes*² and valence list cancellation is in effect
- but *Aux* requires its verbal complement to store case information in this position

$$\left[\begin{array}{l} \textit{transitive-absseg-aux-lex} \\ \text{SUBJ} \langle [\text{CASE } \textit{erg}] \rangle \\ \text{COMPS} \langle [\text{FORM } \textit{nonfinite} \\ \text{COMPS} \langle [\text{CASE } \textit{abs}] \rangle] \rangle \end{array} \right]$$

²(Hinrichs & Nazagawa 1990)

*(AUX (V NP))



VERB-FIRST TYPES

verbal complex analysis

- no interverters
- verbal complex rule added (Fokkens 2010):

$$\left[\begin{array}{l} \text{comp-aux-phrase} \\ \text{SYNSEM|LOCAL|CAT|HEAD} \left[\begin{array}{l} \text{verb} \\ \text{AUX} \quad + \end{array} \right] \\ \text{NON-HEAD-DTR|SYNSEM|LOCAL|CAT|HEAD} \left[\text{verb} \right] \\ \text{HEAD-DTR|SYNSEM|LIGHT} \quad + \end{array} \right]$$

- inherits from head-final, so it's only potentially available to verb-first data
- non-head daughter is a verb
- confront spurious ambiguity on sequences like: (V, Aux, NP) using LIGHT
 - *lex-synsems* [LIGHT +], *phr-synsem* are [LIGHT -] (matrix.tdl)

VERB-FIRST TYPES

the feature [VC *luk*]

- defined on both phrasal and lexical *synsems* (and lexical rules annotated to pass its value up)
- lexical verb types stipulated [VC +], auxiliaries [VC –]
- head-complement rules redefined to take their [VC] value from the non-head daughter
- value of VC on a phrase indicates whether the lexical verb is present in that phrase
- specify that in comp-head and subj-head rules, the head daughter must be [VC +]

VERB-FIRST TYPES

Example

*(V, NP, Aux)

- if $((V, NP)_{VP} \text{ Aux})$: case information inaccessible on daughter of VP, unification fails
- if $(V (NP, \text{Aux}))$: head daughter of potential comp(/subj)-head rule is [VC –], unification fails

WORD ORDER SUMMARY

- *a priori* possible data divided into two classes (by the condition of Manandise's Filter)
 - aux-first
 - verb-first
- head-nexus rules and valence list cancellation capture free word order in the aux-first data
- a verbal complex rule and LIGHT ensure no interveners on the verb-first data

NEGATION

morphological negation

- shape of negative morpheme: *ez-*
- bound:
 - nearly free permutation of syntactic elements, but *ez* is fixed to the aux
 - no intervention of adverbials, generally (some particular grammatical particles are possible between *ez* and aux, but these seem bound too)
 - by treating *ez* as bound, these facts follow from lexical integrity (and therefore don't need to be treated in the syntax)

NEGATION AND WORD ORDER

major constituent order interacts with negation

- generally, auxes can appear on either side of the lexical verb
- negated auxes can only appear on the left of the lexical verb
- under negation we have a narrowing of possible word orders

Miren	ez-da	ibilli		*Miren	ibilli	ez-da
Mary.abs	NEG-AUX	walk.PERF		Mary.abs	walk.PERF	NEG-AUX
<i>Mary has not walked.</i>		[eus]		<i>Mary has not walked.</i>		[eus]

NEGATION AND WORD ORDER

analysis

- recall that our word order analysis treated the data as belonging to two paradigms, with a construction specific rule that only (and always appears in one of the paradigms)
- we define the feature [NEGATED *luk*], and modify the lexical rule that carries out negation such that its result is [NEGATED +]
- add [NEGATED −] to the verbal cluster rule (*comp-aux-phrase*)
- the interaction of these components conspires to rule out any examples in which the lexical verb appears to the left of a negated auxiliary

CONCLUSIONS

existing (independently motivated) analyses working together

- constructional approach created a specific rule associated with a class of sentences
- the rule forms the locus upon which constraints about negation were placed

OUTLOOK

next steps: word order × focus × negation

- focus is configurationally marked
- when the negated auxiliary is in the focus position, Manandise treats this as sentential negation
- when an NP appears as *galdegaia* in a negated clause, constituent negation results
- issues which concern the interface between syntax, semantics and information structure
- extend the grammar presented here to cover interactions with focus