Variable position of 1sg subject pronoun in New Mexico Spanish-English Bilingual Corpus (NMSEB)

Amelia J. Dietrich & Nicole M. Benevento
Department of Spanish, Italian, and Portuguese

The Linguistic Problem

Much of the previous research that has focused on subject-verb word order has not specifically highlighted the 1sg *yo*

If convergence is occurring, we might expect to see its effects in SV word order

Especially given the high rate of code switching that occurs in this community

English "I" is near-categorically preverbal

The Data

The data on which this study is based were extracted from the New Mexico Spanish / English Bilingual (NMSEB) corpus (Torres Cacoullos & Travis, in preparation) housed at the Evaluating Convergence via Code Switching project office at the Pennsylvania State University.

Dependent Variable: Position of *yo*

All tokens of expressed *yo* were extracted

(Note: We checked for English "I" + Spanish verb, but there were no tokens of this in the data analyzed here)

NOT included in the dependent variable those cases when *yo* is unexpressed

168 tokens of "I" + English verb were coded for position – in English, position of "I" is categorically preverbal

Total: 337 tokens

Pre-verbal (297 tokens / 91.6%)

.. yo no sabía cómo bailar. (01.El_Abuelo, 1144)

.. I   no knew how to dance

"I didn't know how to dance"

Post-verbal (40 tokens / 8.4%)

y le sacaba yo el papelito, (03.Dos_Comadres, 530)

and from him was taking I the paper (diminutive)

"and I was taking the paper from him"

Position of preceding 1sg

- Preceding pre-verbal *yo* disfavors postposition
- Preceding post-verbal *yo* favors postposition
- Surprisingly, preceding pre-verbal English *I* does not disfavor postposition of *yo* as one might expect: This is evidence AGAINST convergence

Semantic class

- Cognitive verbs disfavor postposition very strongly as a class
- The cognitive class is accounted for largely by a handful of highly frequent verbs: CREER (n=80/149), SABER (n=44/149)
- Other verbs in Cognitive class: ACORDARSE, CONOCER, APRENDER, HACER PAY ATTENTION, ENTENDER, PENSAR, RECORDAR – also disfavor postposition

Lexical effects

- Lexical ID was not included in the multivariate analysis
- We can also look at the most frequent verbs and note that we might be seeing lexical effects due to chunking or prefabs
- The verbs we examined more closer SABER (n=44), CREER (n=80), and DECIR (n=24 out of 25 enunciative tokens): *Corpus del español* (CdE), Davies, 1900s, ORAL, Accessed December 2/6, 2011

Factors contributing to the post-position of 1sg pronoun *yo* in NMSEB (N = 328)

Overall rate = 11.9% Input = 0.070	Factor Weight	% Post-verbal yo	Total N	% Data
Factor Group				
Semantic class				
Cognitive	0.21	20.1%	3	0.9%
Other	0.75	2.0%	36	11.0%
Range:	54			
Position of previous expressed 1sg pronoun				
Pre-verbal yo	0.40	7.6%	14	4.3%
Post-verbal yo	0.71	26.3%	5	1.5%
Pre-verbal I	0.63	18.0%	20	6.1%
Range:	31			
Subject realization of previous finite verb				
Pre-verbal Span SUBJ	[.40]	6.8%	5	1.5%
Post-verbal Span SUBJ	[.45]	12.5%	3	0.9%
Pre-verbal Eng SUBJ	[.55]	17.2%	10	3.0%
Post-verbal Eng SUBJ (questions only)	[.62]	16.7%	1	0.3%
Unexpressed	[.53]	12.1%	20	6.1%
Range:	n.s.			
Distance to preceding other language material (in IUs)				
0-50 Intonational Units	[.48]	11.7%	33	10.1%
51+ Intonational Units	[.60]	13.0%	6	1.8%
Range:	n.s.			
Log likelihood: -100.299				

Conclusions

- There does not seem to be any indicator of convergence in NM Spanish
- NM Spanish-English bilinguals appear to be able to maintain two separate and independent grammars, despite frequent code-switching

Questions still remain about why *yo* gets post-posed in Spanish:

- Position of preceding expressed *yo* is one factor
- Some particular lexical verbs may strongly favor or disfavor post-position due to their storage as prefabs or chunks
- Other factors to investigate might include pragmatic factors such as emphasis or contrast