Early conditioning of Spanish variable subject expression

Rena Torres Cacoullos, Grant Berry, Chris Champi, Lauren Perrotti, Miguel Ramos
(Penn State University)

Grammatical change has been argued to be gradual rather than abrupt. Romance third person singular (3sg) pronouns grammaticalized from the distal demonstrative 'that one' of Latin, which originally had no separate category of third person pronouns. In present-day Spanish, 3sg subject pronouns with human referents are expressed at rates anywhere from 20% to 40% (e.g., Holmquist 2012:211; Erker & Guy 2012:540). In this study of 13th-16th century texts, we find an expression rate of only 9% (N=1322). This lower aggregate rate is consistent with a diachronic increase in expressed subjects (though rate differences among the texts are due to genre rather than chronology). More revealing of change, however, is the linguistic conditioning of the variation in 3sg subject expression.
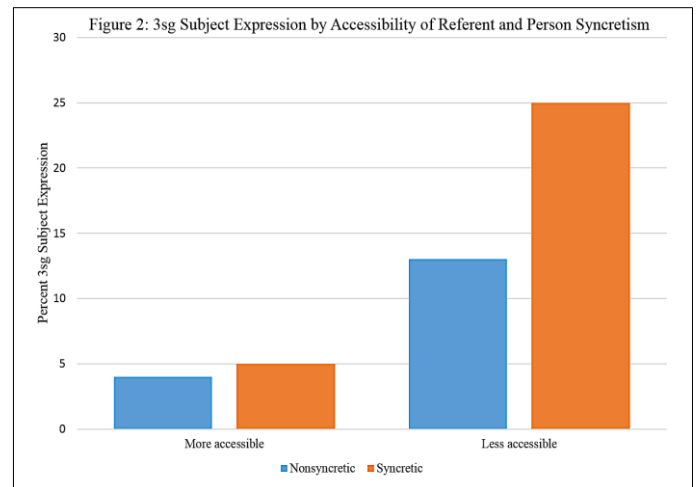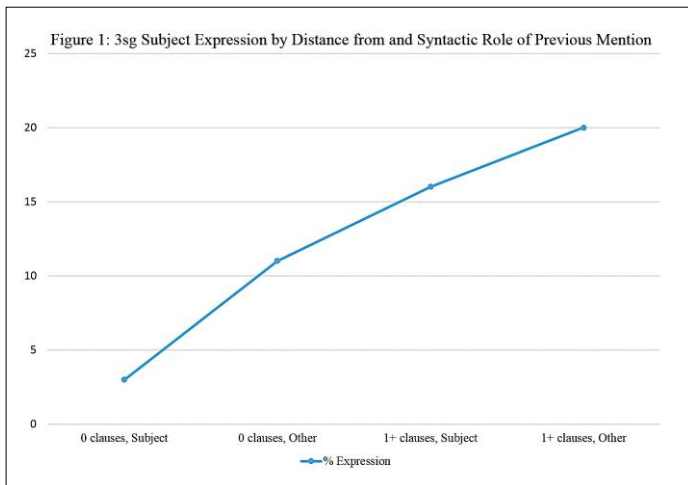
On the one hand, many of today's well-known effects were already operative more than 500 years ago, and with the same fine details in conditioning. Expression rates are lowest when the previous mention is the subject of the immediately preceding clause (e.g., Silva-Corvalán 1994:157) (Figure 1). Person syncretism in verb tense morphology interacts with this accessibility effect, applying with less accessible referents, i.e. when the previous mention is at a distance of 1 or more clauses, but being "nullified" in contexts of continuity of reference (Cameron 1993:317) (Figure 2). Pronoun-to-pronoun priming is also present, and this effect is stronger with more accessible than with less accessible referents (cf. Cameron 1994:40).

On the other hand, an effect which is to our knowledge unattested in studies of present-day variation is the gender of the referent, with feminine favoring subject expression. Furthermore, *ella* 'she' and *él* 'he' have distinct linguistic conditioning, as revealed in both independent variable-rule analyses and an interaction term between verbal morphology syncreticity and gender in a generalized linear model (Wald Z=2.891, p<.01).

(1a) *quel' Ø faría cuanto **ella** quisiesse*
       'that (he) would do <u>for her</u> whatever **she** wanted' (Lucanor, L)

(1b) *quel' Ø pidría que non le **Ø** fablasse más*
       'that (she) would ask <u>him</u> that (**he**) not speak to her' (Lucanor, L)

To interpret the gender effect, we look to the distribution of the data. In these data, 3sg subjects referring to females are not only less frequent than those referring to males (at a ratio of 1:4) but, when they do occur, they are less likely to have been previously mentioned as a subject: the proportion of previous mentions as object or oblique, as in (1), is 29% for feminine (N=251) but 17% for masculine (N=977) (p < .001). Given the association between subject and topic (e.g., Givón 1979:209, 298), these distributions suggest that female referents are less topical in the texts overall. As to the conditioning of variant choice, we find a stronger effect for *ella* than for *él* of favoring by previous mentions in other than subject role. Together the data distribution and effects, then, suggest that subject expression in these early texts is sensitive not only to accessibility of the referent (as measured by distance from the previous mention) but also to socio-pragmatic considerations, marking unexpected—or less likely—subjects.

Figure 1: 3sg Subject Expression by Distance from and Syntactic Role of Previous Mention

25
20
15
10
5
0

0 clauses, Subject   0 clauses, Other   1+ clauses, Subject   1+ clauses, Other

—♦—% Expression

Figure 2: 3sg Subject Expression by Accessibility of Referent and Person Syncretism

30
25
20
15
10
5
0

Percent 3sg Subject Expression

More accessible   Less accessible

■ Nonsyncretic ■ Syncretic

References

Cameron, Richard. 1993. Ambiguous agreement, functional compensation, and nonspecific *tú* in the Spanish of San Juan, Puerto Rico, and Madrid, Spain." *Language Variation and Change* 5(3): 305-334.

Cameron, Richard. 1994. Switch reference, verb class and priming in a variable syntax. *Papers from the Regional Meeting of the Chicago Linguistic Society: Parasession on variation in linguistic theory* 30(2): 27-45.

Erker, Daniel and Gregory R. Guy. 2012. The role of lexical frequency in syntactic variability: Variation subject personal pronoun expression in Spanish." *Language* 88(3): 526-557.

Givón, Talmy. 1979. *On understanding grammar.* New York: Academic Press.

Holmquist, Jonathan. 2012. Frequency rates and constraints on subject personal pronoun expression: Findings from the Puerto Rican highlands. *Language Variation and Change* 24(2): 203-220.

Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.

Texts and editions

Calila (1250) = Anonymous, *Calila e Dimna*, Juan Manuel Cacho Blecua & María Jesús Lacarra (eds.). Madrid: Castalia, 1984.

Lucanor (1335) = Don Juan Manuel, *El conde Lucanor o Libro de los enxiemplos del conde Lucanor et de Petronio*, José Manuel Blecua (ed.). Madrid: Castalia, 1971.

Celestina (1499) = Fernando de Rojas, Fernando. *La Celestina*, Dorothy S. Severin (ed.), Madrid: Cátedra, 1987.

Lazarillo (1554) = Anónimo, *Tri-linear edition of* Lazarillo de Tormes *of 1554 (Burgos, Alcalá de Henares, Amberes)*, Joseph V. Ricapito (ed.). Madison: The Hispanic Seminary of Medieval Studies, 1987