Predictor	**Estimate**	**SE**	***z* value**	***p* value**
(Intercept)	-0.12	0.57	-0.2	0.833
constr=person-marking	-0.49	0.05	-9.2	<0.001*
group=West&South	-2.52	0.66	-3.8	<0.001*
constr=person-marking x group=West&South	0.79	0.08	10.1	<0.001*

Table 1: Summary of the logistic-regression model: presence of a pronominal subject as predicted by construction and language group with by-VERB and by-LANGUAGE random effects; within clauses in simple sentences and first conjuncts only. Asterisks denote significance at the 0.05 level.

In the article, we report a regression model that is fit to data from all clauses (Table 3, Figure 4). It is, however, reasonable to expect that the means of encoding subject will strongly vary across clause types (main, simple, subordinate, coordinate). However, adding CLAUSE TYPE as a predictor leads to severe convergence problems, rendering the models unusable. Instead, we opt for a simpler way of controlling for a potential effect of CLAUSE TYPE. We fit a separate model with exactly the same specification as the model reported in the paper, but using only those clauses that are neither subordinate clauses nor main clauses that have a subordinate clause nor non-first conjuncts in coordinated constructions. What remains are clauses in simple sentences and clauses that are first conjuncts in coordinated constructions. The model yields similar results, see Table 1. The absolute values of the coefficients are different, but the general picture is the same.

$C = 0.91$, Somers' $D_{xy} = 0.82$.