A Unified Model of Gradient Acceptability for Agreement with Coordinate Structures

Aixiu An [1], Anne Abeillé [1] ([1] University of Paris, LLF, CNRS), Roger Levy(MIT)
aixiu0806@gmail.com

Experimental acceptability data are very important empirical resources for current linguistic and psycholinguistic theories. Most of these studies test the role of certain factors in one/two experiments (significant or not). This present work, on the one hand, provides a detailed comparison of a rich set of human acceptability experiments, using cross-validated/held-out train/test methodology; on the other hand we show that gradient acceptability captures the typological tendencies proposed by Corbett (1991) ((1) p.3). The results suggest that violation of agreement with the closest conjunct (CCA, *The boy or the girl **is** coming*) has a bigger penalty on the acceptability in the nominal domain than in the verbal domain, for gender than for number.

Taking inspiration from Harmonic grammar (Legendre et al. 1990) and linear OT (Keller2000), we propose a new sum-weighted model, to handle the gradient grammaticality of agreement with a conjoined NP. A growing number of experimental studies begin to explore this question (e.g. Willer-Gold et al. 2017), but they only rely on a few coordination patterns. Given the sparsity of data obtained through human experiments, we propose a flexible model which can learn the constraints' weights through a limited number of structures and generalize to other structures. Agreement with a coordinate NP is a pervasive phenomenon which allows for several strategies (Resolution (RA), Closest Conjunct agreement (CCA), First conjunct agreement (FCA); see (2) p.3 in French). Contrary to previous studies which only concentrate on one feature/domain combination (for example number in the nominal domain in Le Bruyn and de Swart (2014), or gender in verbal domain in Nevins and Weisser (2019)), our model unifies constraints from four main properties of agreement: controller, target, domain, feature in order to i) identify the linguistic regularity that underlies this linguistic phenomenon; ii) generalize to unseen patterns.

The acceptability data consists of 54 conditions in French from 10 experiments: 5 for number agreement (2 attributive, 3 predicative), and 5 for gender agreement (3 attributive, 2 predicative).

In our model, each constraint consists of one agreement strategy (CCA,RA, FCA), one feature in the agreement domain inventory (Attributive/Predicative), and one agreement feature (gender/number) (e.g. CCA [Att][Gen]). Every sentence x is assigned a score $h(x)$ to represent the weighted sum of the form's constraint violations (eq. (1) on p.4). In order to estimate the weight of each constraint, we use package *fuser* (eq. (2)) in R which treats gender and number as two related sub-groups and force the coefficient vectors from different subgroups to be similar.

The model is trained with 28 observations from *and-coordination* which leads to the following findings: (i) CCA is possible for gender and number, and more acceptable for gender. Since grammatical gender is arbitrary for non humans, gender resolution is more arbitrary than number resolution, which expresses a plural meaning; another hypothesis is that gender is a recessive feature and thus more sensitive to linear proximity. (ii) CCA is possible in both nominal and verbal domains, but its violation yields a stronger penalty in the nominal domain, which is in line with the agreement hierarchy proposed by Corbett (1991). This suggests that structural proximity plays a role since the target and the conjunct are structurally closer in the nominal domain.

We firstly use this model to predict some unseen patterns in *and-coordination* (fig. 4). We also generalize the constraints learned from *and* to *or*. The model trained with *and* captures most disjunctive data except the N1sg-or-N2g pattern (fig. 5).

This work presents an explainable model that maps the gradient grammatical knowledge characterized by agreement to structural feature based constraints (controller, target, domain, feature). Through the acceptability data, it captures typological tendencies into a quantitative predictive framework, and learns linguistic generalizations from a limited input: linear proximity is more important in the nominal domain and for gender agreement.

Selected References: Corbett (1991). Gender. Cambridge University Press; Keller, F. (2000).Gradience in grammar: experimental and computational aspects of degrees of grammaticality. PhD thesis, U. of Edinburgh; Willer-Gold, J. et al. (2017). When linearity prevails over hierarchy in syntax. PNAS, pp 712-729.

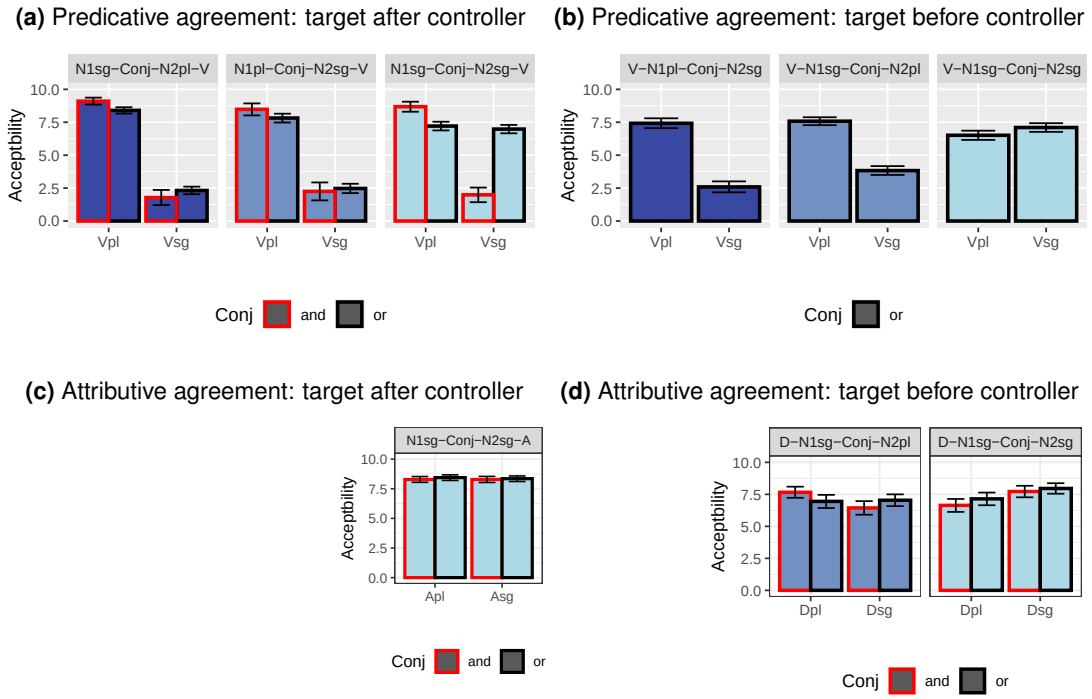

Figure 1: Experimental results on number agreement. Dark blue means closest noun plural, less dark closest noun singular, light blue both singular. For predicative agreement only *or* in VS order.

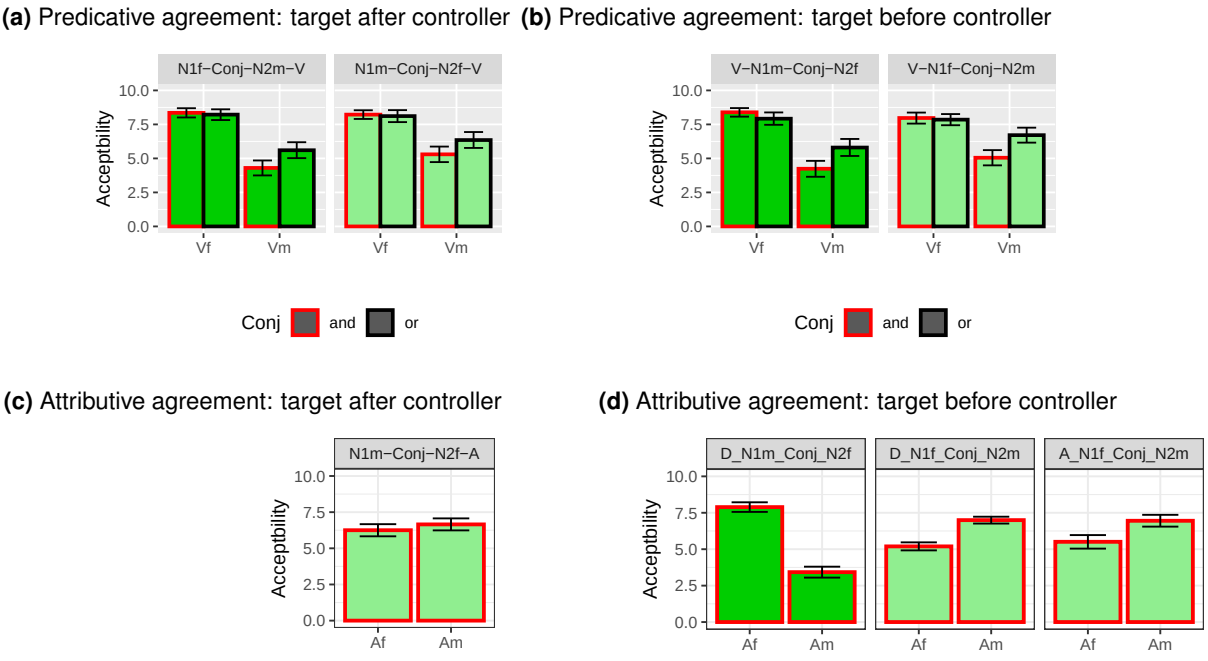

Figure 2: Experimental results for gender agreement. Dark green when the closest noun is masculine, light green when it is feminine. We only tested *and* for attributive agreement.

(1) attributive > predicate > relative pronoun > personal pronoun
⟵ non-resolution – – resolution ⟶

(2) Merci d' indiquer votre/vos nom et prénom.
thanks for indicate your.SG/PL name and first name
'Please indicate your first and last name'

Domain	Condition	Sentences
Attributive	D-N1sg-Conj-N2sg	Il faudrait pouvoir prévenir le/les directeur et/ou sous-directeur de l'établissement. it should warn the.MSG/PL director and/or assistant director of the establishment
	N1sg-Conj-N2sg-A	Cette formation gratuite vous prépare au mieux à la fonction de directeur et/ou sous-directeur administratif/administratifs. This training free you prepare at best for the position of director and/or assistant director administrative.MSG/PL
Predicative	N1sg-Conj-N2sg-V	Je me demande où le maire et/ou l'adjoint va/vont aller. I wonder where the mayor and/or the deputy is/are going
	V-N1sg-Conj-N2sg	Je me demande où va/vont aller le maire et/ou l'adjoint. I wonder where is/are going the mayor and/or the deputy
Attributive	D-N1f-Conj-N2m	Certains/certaines étudiantes et étudiants sont déjà en stage. Some.MPL/FPL student.FPL and student.MPL
	A-N1f-Conj-N2m	De nouveaux/nouvelles étudiantes et étudiants are already on internship. some new.MPL/FPL student.FPL and student.MPL are already on internship.
	N1m-Conj-N2f-A	Des étudiants et étudiantes nouveaux/nouvelles sont déjà en stage. some student.MPL and student.FPL new.MPL/FPL are already on internship.
Predicative	N1f-Conj-N2m-V	Je me demande où les étudiantes et/ou les étudiants seront conduits/conduites. I wonder where the student.FPL and/or the student.MPL will be taken.MPL/FPL
	V-N1f-Conj-N2m	Je me demande où seront conduits/conduites les étudiantes et/ou les étudiants. I wonder where will be taken.MPL/FPL the student.FPL and/or the student.MPL

Table 1: Conditions for the 10 experiments, including variations between conjuncts

$$h(x) = \sum_{i=1}^{N} w_i C_i(x) \tag{1}$$

$h(x)$ denotes the weighted sum of the form's constraint violations. If form x violates constraint i, $C_i(x)$ returns -1, otherwise 0. The weight associated to each constraint is penalized to be positive.

The model uses the penalized formula:

$$\hat{B} = \underset{B=[\beta_1 ... \beta_K]}{\operatorname{argmin}} \sum_{k=1}^{K} \{ \frac{1}{n_k} ||y_k - X_k \beta_k||_2^2 + \lambda ||\beta_k||_1 + \gamma \sum_{k`>k}^{K} \tau_{k,k`} ||\beta_k - \beta_{k`}||_2^2 \} \tag{2}$$

Where λ is sparsity penalty hyperparameter and γ is Fusion penalty hyperparameter.

	number	gender
Intercept	8.9323147	8.4508600
CCA:Att	0.7150191	1.9760858
CCA:Pred	0.3253251	0.3348568
RA:Att	1.4191908	1.5706473
RA:Pred	6.5755310	3.5291744
FCA:Att	0.9513824	1.2410600
FCA:Pred	0.1691036	0.1303033

Table 2: Coefficients of each constraint learned by Model $fuser$ when the residual sum of squares (RSS) is minimum. In such condition, $\lambda = 0.000005$, $\gamma = 0.00002$, RSS=0.3366905.

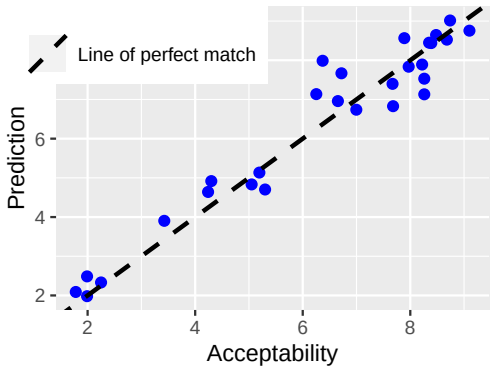

Figure 3: Leave one out cross-validation for each condition in and-coordination. The mean squared error in the leave-one out cross validation is 0.34

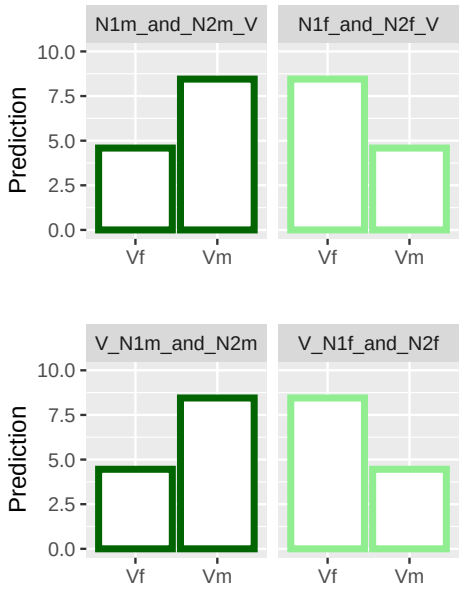

Figure 4: Predictions on unseen patterns of gender agreement with *and*

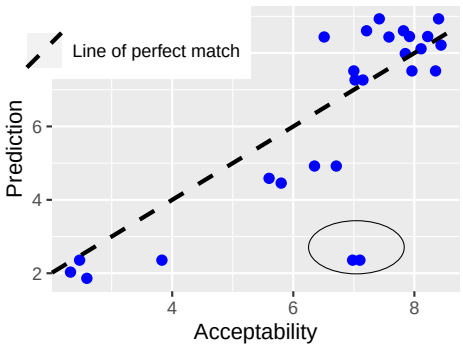

Figure 5: Predictions on 26 conditions with *or* using models in Table 2 trained with *and*. The model captures most generalizations except two conditions (circled): V-N1sg-or-N2sg, N1sg-or-N2sg-V