Title: Interspeaker Variation in Copular Agreement with Disjoined Subjects: an Optimality-Theoretic Account

Introduction: In English, copulas must agree with the subject in number (e.g., (1)). Agreement with disjoined subjects (i.e., *(either) A or B*) presents a puzzle. When the two disjuncts mismatch in number, it is unclear whether the whole disjoined DP is singular or plural. Previous work noted a slight preference for agreement with the second disjunct [1-3]. Furthermore, when the two disjuncts are both singular, there is an unexpected preference for plural agreement (e.g., (2)) [2,3]. In this study, we propose an Optimality-Theoretic (OT) account for copular agreement with disjoined subjects in English, supported by experimental evidence for systematic interspeaker variation in agreement patterns overlooked by previous studies.

Proposal: [4] proposed that subject-copula agreement in English can be captured by two OT constraints, *Sg and $\textsc{Agree}_{\textsc{subject}}$ (Table 1), and a ranking of $\textsc{Agree}_{\textsc{subject}} \gg$ *Sg. In the lexicon, there is a 3SG copula "is", a 1SG copula "am", and a null copula "are" without person/number features. Extending [4]'s analysis to disjoined subjects, we add two constraints: $\textsc{Agree}_{\textsc{close}}$ and $\textsc{Agree}_{\textsc{each}}$ (Tables 1 and 2). Maintaining $\textsc{Agree}_{\textsc{subject}} \gg$ *Sg, the constraints predict the factorial typology in (3). Assuming speakers are otherwise free to acquire any ranking while having $\textsc{Agree}_{\textsc{subject}} \gg$ *Sg, we expect three distinct patterns shown in (3) among the population.

Exp.1 (N=200) examines whether the predicted number agreement patterns in (3) are attested among American English speakers. Participants were tasked with completing 72 (24 critical, 48 filler) sentences with third-person disjoined subjects as in (4), presented in a random order. Only "is" or "are" could be used to complete the sentences. K-means clustering [5] and Silhouette analysis [6] (Figure 1) were used to identify distinct response patterns among the participants; three clusters emerged. Figure 2 shows the response patterns of the three clusters. The three clusters exactly match the three predicted agreement patterns in (3).

Exp.2 (N=298) corroborates our account by extending to person feature agreement. Exp.2 adopts a similar design as Exp.1, but with extra conditions with first-person disjuncts (as in (5)) and the addition of the possible answer "am". Assuming that $\textsc{Agree}_{\textsc{close}}$ and $\textsc{Agree}_{\textsc{each}}$ apply to person and number features alike, the three speaker clusters identified in Exp.1 are predicted to show the agreement patterns in (6). Figure 3 shows the responses collected in Exp.2. Using the same k-means clustering algorithm as in Exp.1, participants were first grouped into three clusters based on their responses to the third-person conditions. Response preferences are less clear-cut than in Exp.1, potentially due to fatigue since Exp.2 is much longer. But crucially, for all three clusters, the most preferred copula choices for all conditions exactly match the predictions in (6).

Conclusion: In sum, this study presents novel empirical generalizations about interspeaker variation in subject-copula agreement, and shows that [4]'s OT analysis for simple copula agreement with minimal extension can capture the variation pattern. This study not only contributes to our understanding of the mechanisms of subject-copula agreement in English but also highlights the need for attention to interspeaker variation in morphosyntactic research.

Constraint	Assign 1 violation when…
*Sg	…the copula is singular.
$\textsc{Agree}_{\textsc{subject}}$	…the copula does not agree with the subject. (Disjoined subjects are null in number/person.)
$\textsc{Agree}_{\textsc{close}}$	…the copula does not agree with the closer disjunct when the subject is disjoined.
$\textsc{Agree}_{\textsc{each}}$	…the copula does not agree with both disjuncts when the subject is disjoined.

Table 1 (left). Constraints for copula agreement with disjoined subjects.

Table 2 (right). OT tableau for agreement with third-person subjects. "am" is harmonically bounded and thus not shown. Constraints are unranked in the table.

		$\textsc{Agree}_{\textsc{close}}$	$\textsc{Agree}_{\textsc{each}}$	$\textsc{Agree}_{\textsc{subject}}$	*Sg
Either 3SG or 3SG	is (3SG)			*	*
	are (null)	*	*		
Either 3SG or 3PL	is (3SG)	*	*	*	*
	are (null)	*	*		
Either 3PL or 3SG	is (3SG)		*	*	*
	are (null)	*	*		
Either 3PL or 3PL	is (3SG)	*	*	*	*
	are (null)	*	*		
3SG	is (3SG)				*
	are (null)			*	
3PL	is (3SG)			*	*
	are (null)			*	

References: [1] Haskell & MacDonald (2005), *Journal of Experimental Psychology* [2] Foppolo & Staub (2020), *Cognition* [3] Petersen (1986), *Australian Journal of Linguistics* [4] Bresnan (2001), chapter in *Handbook of Contemporary Syntactic Theory* [5] MacQueen (1967), *Proceedings of BSMSP 5* [6] Rousseeuw (1987), *Computational and Applied Mathematics*

(1) The students *is/are in the room.

(2) Either the student or the teacher are in the room.

(3) Possible number agreement patterns based on constraints in Table 2 (all disjuncts are in third person):

a. If we have rankings: $\textsc{Agree}_{\textsc{close}} \gg \{\textsc{Agree}_{\textsc{subject}}, *\textsc{Sg}\}$, then we have:

Either SG or SG is… Either SG or PL are… Either PL or SG is… Either PL or PL are…

b. If we have rankings: $\textsc{Agree}_{\textsc{each}} \gg \{\textsc{Agree}_{\textsc{subject}}, *\textsc{Sg}\} \gg \textsc{Agree}_{\textsc{close}}$, then we have:

Either SG or SG is… Either SG or PL are… Either PL or SG are… Either PL or PL are…

c. If we have rankings: $\textsc{Agree}_{\textsc{subject}} \gg \{\textsc{Agree}_{\textsc{each}}, \textsc{Agree}_{\textsc{close}}\}$ or $*\textsc{Sg} \gg \{\textsc{Agree}_{\textsc{each}}, \textsc{Agree}_{\textsc{close}}\}$, then we have:

Either SG or SG are… Either SG or PL are… Either PL or SG are… Either PL or PL are…

(4) Either [those art students/Hossein] or [those illustrators/Elias] __is,are__ sketching the concept art.

(5) Either [I/we/the clown/the wizards] or [I/we/the sailor/the zoologists] __is,are,am__ giving away free donuts.

(6) Predicted response patterns for Exp.2

a. Cluster 1: Either 3SG or 3SG is… Either 3SG or 3PL are… Either 3PL or 3SG is… Either 3PL or 3PL are…
Either 1SG or 3SG is… Either 1SG or 3PL are… Either 1PL or 3SG is… Either 1PL or 3PL are…
Either 3SG or 1SG am… Either 3SG or 1PL are… Either 3PL or 1SG am… Either 3PL or 1PL are…
Either 1PL or 1SG am… Either 1SG or 1PL are…

b. Cluster 2: Either 3SG or 3SG is… Either 3SG or 3PL are… Either 3PL or 3SG are… Either 3PL or 3PL are…
For all other conditions, the preferred copula form is always “are”.

c. Cluster 3:Either 3SG or 3SG are… Either 3SG or 3PL are… Either 3PL or 3SG is… Either 3PL or 3PL are…
For all other conditions, the preferred copula form is always “are”.

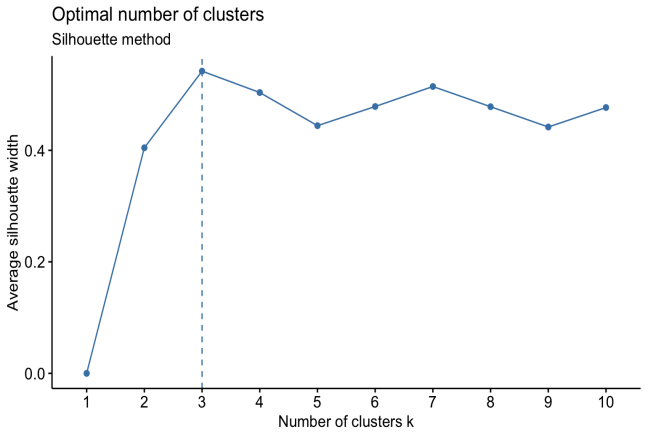

Fig.1 (upper left). Silhouette analysis results of Exp.1 responses, suggesting responses form three clusters.

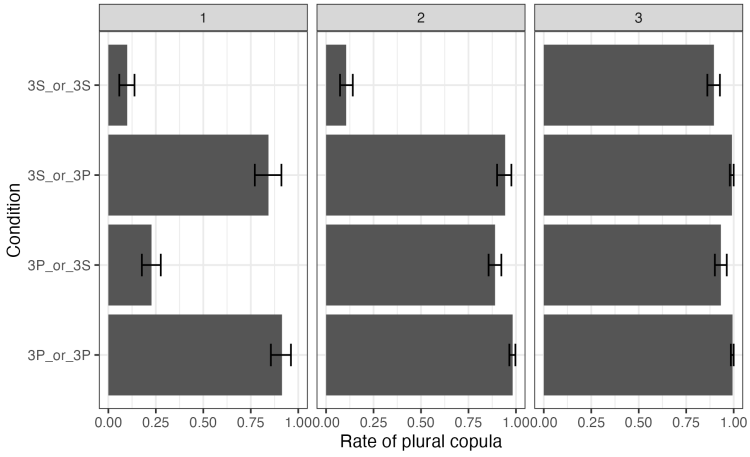

Fig.2 (upper right). Mean “are” response rate in Experiment 1, by condition and cluster.
3 = 3rd person, S = singular, P = plural

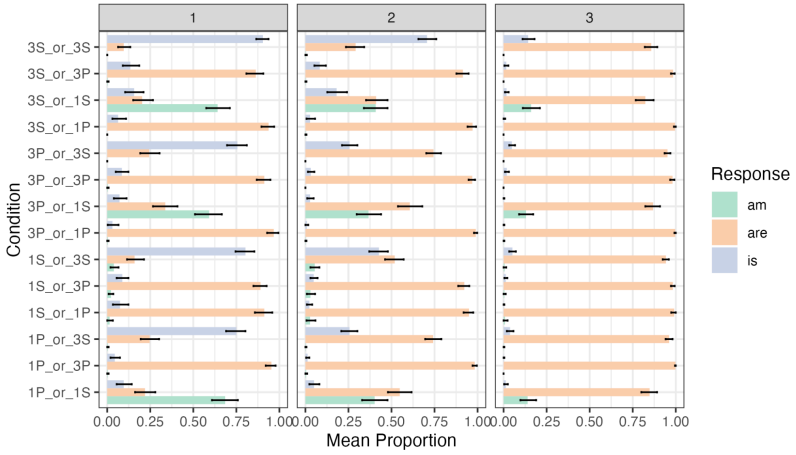

Fig.3 (lower right). Mean proportion of responses in Exp.2, by condition and cluster. 1 = 1st person, 3 = 3rd person, S = singular, P = plural