An Evolutionary Model-Based Approach to the Missing O-Corner

Maike Züfle[1], Roni Katzir[2]
[1]University of Heidelberg, [2]Tel Aviv University

Introduction. The typology of lexicalized logical operators is highly skewed (Horn 1972; Barwise and Cooper 1981; Fintel and Matthewson 2008). Focusing on the Aristotelian Square of Opposition, Horn observes that the A (all) and I (some) corners are frequently lexicalized across languages and categories and appear morphologically simplex; E (no) is often absent and, when lexicalized, is morphologically marked; and O (not all) is quite generally absent. The correct account of this pattern is still debated, but one idea (going back to Horn and developed further by Katzir and Singh 2013 and Uegaki 2020; H/KS/U) is that it relates to three factors: communication, economy, and cost. Specifically, I and O can both communicate $\exists \wedge \neg\forall$ (via scalar implicature), so lexicalizing both is in some sense redundant. Consequently, given economy, the 3-element inventories $\{A, I, E\}$ and $\{A, E, O\}$ block the full $\{A, I, E, O\}$. As to the choice between the two 3-corner inventories, H/KS/U posit that O is costlier than I (perhaps because O can be decomposed into A+neg), so economy chooses $\{A, I, E\}$ over $\{A, E, O\}$.

Goals. The present work accepts the general idea above and proposes to examine three issues left open in the literature on this typological pattern. (Q1) Why should inventories be economized in this way (selecting only small inventories whose elements have low costs)? After all, the savings seem very minimal, while the typological pattern is very robust. (Q2) What costs matter exactly? Is the ranking of $A, I \ll E, O$ assumed by H/KS/U warranted? (Q3) Could the culprit for the typological asymmetry be not grammatical costs (as in H/KS/U) but rather the usage likelihood of the different meanings (as suggested by Enguehard and Spector 2019)? We explore these three questions using evolutionary dynamics, specifically using the Replicator-Mutator Dynamic (Hofbauer 1985) as adapted for semantics/pragmatics by Brochhagen, Franke, and Rooij (2018), where communication and costs have a central role: communication affects fitness, and costs affect learnability.

Model. The model simulates evolution over generations of different *types*, each representing a lexical inventory (that may include any combination of the four corners $\{A, I, E, O\}$, as well as other operators). Types interact in two ways. First, they communicate with each other within each generation and can compute scalar implicatures (here modeled using the Rational Speech Act model; Frank and Goodman 2012, Bergen, Levy, and Goodman 2016). And second, they interact through language learning in each generation from the previous generation (here using Bayesian learning). If a type supports successful communication and faithful learning, its proportion grows: $x_i' = \sum_j Q_{ji} \frac{x_j f_j}{\sum_h x_h f_h}$ where x_i' is the frequency of type i at the next time step, x_j is the frequency of type j at the current time step, Q_{ji} is the learnability (probability that observing type j's linguistic

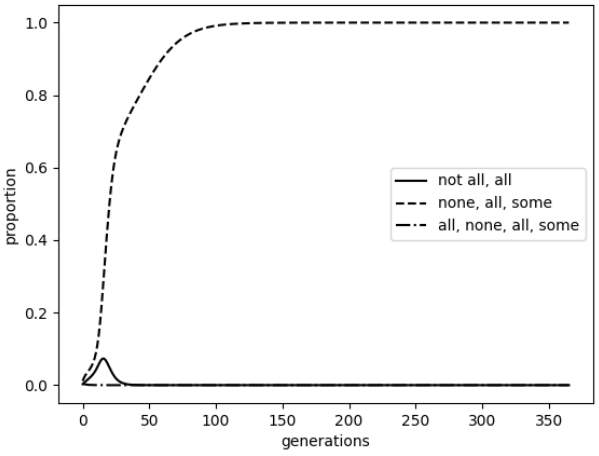

Figure 1: Proportion of the best language types of each lexicon size over multiple generations

output leads to the acquisition of type i), and f_j is the communicative success of type j.

Results. Regarding (Q1) above, we find that even very minimal differences in inventory cost can lead to a robust typological skew as long as the inventory is still communicatively useful (as in the attested pattern). Fig. 1, where the correct 3-corner lexicon outperforms a 4-corner type, illustrates. With respect to (Q2), we find that the ranking $A, I \ll E, O$ is indeed warranted: it leads to the attested pattern, while other cost assignments do not. A schematic summary of our results is presented in table 1, with $A, I \ll E, O$ represented in rows 5 and 7. (Note that $A, I, E \ll O$ gives the correct result for the 3-corner case, as shown in row 4 of table 1, but it fails for the 2-corner case, as in row 6.)

Rel. costs	lex. of WT	prop. WT	runs won
A, I, E, O	$A, I(/O), E$	50.0	51.0 %
$A, I, O \ll E$	$A, I(/O), E$	49.15	50.6 %
$E, I, O \ll A$	$A, I(/O), E$	50.01	49.52 %
$A, I, E \ll O$	**A, I, E**	100.0	99.7 %
$A, I \ll E, O$	**A, I, E**	100.0	100.0 %
$A, I, E \ll O$	I, E	100.0	84.2 %
$A, I \ll E, O$	**A, I**	100.0	100.0 %

Table 1: Relative costs for the operators, winning types, their proportions, and percentage of runs won under different initializations. Each row corresponds to a range of concrete numerical cost assignments. $I(/O)$ (rows 1–3): The winning types in these cases were split $\sim 50/50$ between lexica with I and with O.

Priors over states	lex. of WT	prop. WT	runs won
$P(\neg\exists) = P(\forall)$	$A, I(/O), E$	50.02	50.00 %
$P(\neg\exists) < P(\forall)$	A, O, E	100.0	87.47 %
$P(\forall) < P(\neg\exists)$	**A, I, E**	100.0	87.44 %
$P(\neg\exists) = P(\forall)$	A, E	100.0	40.44 %
$P(\neg\exists) < P(\forall)$	A, O	100.0	83.29 %
$P(\forall) < P(\neg\exists)$	I, E	100.0	83.73 %
$P(\forall) \ll P(\neg\exists)$	I, E	100.0	71.49 %

Table 2: Relative priors over states, winning types, their proportions, and percentage of runs won under different initializations. Each row corresponds to a range of concrete numerical priors. Row 7 corresponds to highly biased priors. $I/(O)$ = as in table 1.

As to (Q3), we find that biased priors over states can explain the absence of O in the 3-corner case (row 3 in table 2; in line with Enguehard and Spector 2019) but fail to derive either the relative markedness of E or the possibility of the empirically common $\{A, I\}$ inventories (rows 4–7). When both costs and priors over states are applied, the costs alone determine the outcomes while the priors make no contribution. Taken together, our results support H/KS/U, including their specific cost assumptions, and allow their economy assumption to be explained as the result of a general evolutionary dynamic. More generally, our results show how such an evolutionary dynamic makes it possible to reason both about the division of labor between grammatical and functional factors affecting the typology and about fine-grained grammatical properties (such as costs of primitives of semantic representation) that can be difficult to probe otherwise.

References. [**1**] J. Barwise and R. Cooper. "Gen. Quant. and Nat. Lang." (1981). [**2**] L. Bergen, R. Levy, and N. Goodman. "Prag. Reasoning through Sem. Inference" (2016). [**3**] T. Brochhagen, M. Franke, and R. v. Rooij. "Coevolution of Lex. Meaning and Prag. Use" (2018). [**4**] É. Enguehard and B. Spector. "Explaining Gaps in the Logical Lex. of Nat. Lang." Ms. 2019. [**5**] K. von Fintel and L. Matthewson. "Universals in Sem." (2008). [**6**] M. C. Frank and N. D. Goodman. "Predicting Prag.

Reasoning in Lang. Games" (2012). [**7**] J. Hofbauer. "The Selection Mutation Equation" (1985). [**8**] L. Horn. "On the Sem. Properties of the Logical Ops. in Eng." PhD thesis. 1972. [**9**] R. Katzir and R. Singh. "Constraints on the Lex. of Logical Ops." (2013). [**10**] W. Uegaki. "*NAND and the Comm. Effic. Model". Ms. 2020.