On the Optimality of Discrete Object Naming: a Kinship Case Study

Anonymous ACL submission

Abstract

The structure of naming systems in natural languages hinges on a trade-off between high informativeness and low complexity. Prior work capitalizes on information theory to formalize these notions; however, these studies generally rely on two assumptions: (i) optimal listeners, and (ii) universal communicative need across languages. Here, we address these limitations by introducing an information-theoretic framework for discrete object naming systems, and we use it to prove that an optimal trade-off is achievable only when the listener's decoder is equivalent to the Bayesian decoder of the speaker. Adopting a referential game setup from emergent communication, and focusing on the semantic domain of kinship, we show that our notion of optimality is not only theoretically achievable but also emerges empirically in learned communication systems.

1 Introduction

Languages across the world exhibit substantial variation in their lexical systems. On the surface level, lexical items that refer to equivalent meanings are expressed in different word forms (e.g., the word *aunt* in English is *tía* in Spanish), but also at the semantic level, meaning partitions vary cross-linguistically (in Vietnamese, there are different words for *aunt*, depending on whether she is the younger or elder sister of one's mother or father).

Despite the richness of this variation, it does not appear to be arbitrary. A growing body of work suggests that languages do not explore the space of possible semantic partitions freely, as evidenced by constrained and recurrent cross-linguistic patterns (Kemp and Regier, 2012; Regier et al., 2015; Zaslavsky et al., 2018; Kemp et al., 2018; Carr et al., 2020; Chaabouni et al., 2021). Instead, the structure of object naming systems appears to reflect pressures for communicative—and possibly cognitive—efficiency. These pressures are thought to be domain-general, as similar patterns have been observed across semantic domains such as kinship, color, and general object categorization. In particular, these studies suggest that languages tend to evolve toward object naming systems that approximate an (often near-)optimal trade-off between informativeness and complexity.

To formalize this trade-off, much of the prior work capitalizes on constructs from information theory (Shannon, 1948). Informativeness is typically quantified by the amount of information preserved in communication—often framed as the inverse of information loss—while complexity measures how concisely a language compresses meaning into words. These frameworks define optimal trade-off boundaries: curves along which no system can reduce complexity without increasing information loss, or decrease information loss without becoming more complex.

As an example, Kemp and Regier (2012) demonstrate that natural kinship systems lie near the optimal trade-off frontier. However, since their measure of complexity is based on the shortest kinship description in a language, it is not analytically tractable to derive a closed-form expression for the trade-off curve. As a workaround, they approximate the curve using a set of generated hypothetical systems. In contrast, the Information Bottleneck (IB) framework (Tishby et al., 2000) enables Zaslavsky et al. (2018, 2019); Chaabouni et al. (2021) to derive theoretical, closed-form *approximations* of the optimal trade-off frontier in domains such as color, container, and animal naming.

However, these studies generally rely on two simplifying assumptions: (i) that listeners are optimal in the Bayesian sense, and (ii) that a universal communicative need distribution, i.e., a distribution over the object space, applies uniformly across all languages. The first assumption overlooks the impact of listener suboptimality, which can arise from various factors in real-world settings (Gibson et al.,

2019). The second assumption clearly oversimplifies the cultural and linguistic diversity observed across natural communication systems.

In this work, we revisit the question of how to measure optimality in object naming systems, particularly in scenarios where the objects are inherently discrete. Building on prior work such as Kemp and Regier (2012) and Zaslavsky et al. (2018), our approach is grounded in information theory. Unlike these earlier studies, however, we derive an *exact*, closed-form expression for the optimal trade-off curve. This formulation enables us to analyze how both optimal and suboptimal listeners influence the informativeness–complexity trade-off in communication. Moreover, because the optimal curve is agnostic to the communicative need distribution, it supports meaningful comparisons of communication systems across varying distributions over the objects to be named.

Our contributions are as follows. First, we introduce an information-theoretic framework that formally characterizes information loss, complexity, and optimality in discrete object naming systems. Second, using this framework, we prove that object naming achieves an optimal trade-off under a specific and well-defined condition—namely, when the listener's decoder is equivalent to the Bayesian decoder the speaker. Third, adopting a referential game setup commonly used in emergent communication (Lazaridou and Baroni, 2020; Lazaridou et al., 2017; Havrylov and Titov, 2017; Chaabouni et al., 2022), and focusing on the semantic domain of kinship, we show that our notion of optimality is not only theoretically achievable but also emerges empirically in learned communication systems.

2 Background

The observed structure of object naming systems in natural languages appears to balance complexity and informativeness, often achieving a near-optimal trade-off. However, the definition of optimality is underpinned by the definitions of complexity and informativeness, and those are not unique. A relevant distinction lies in the fact that some of these domains are inherently continuous (e.g., color), while others are concerned with discrete 'objects', such as kinship. Here we focus on two representative cases of this distinction.

2.1 Optimality in Color Naming

Zaslavsky et al. (2018) introduce an information-theoretic framework for quantifying the trade-off between informativeness and complexity in lexical systems. Focusing on the continuous domain of color, they propose that natural color naming systems approximate optimal trade-offs by compressing perceptual meanings into words in a manner consistent with the Information Bottleneck (IB) principle (Tishby et al., 2000), provided that listeners are optimal in the Bayesian sense.

In this framework, meanings are modeled as probability distributions over perceptual states—in this case, color stimuli—while lexical items are treated as compressed representations of these distributions. The IB objective seeks an encoder $q(w|m)$ that maps meanings m to words w by minimizing the following functional:

$$\mathcal{F}_\beta[q(w|m)] = I_q(M;W) - \beta I_q(W;U),$$

where $I_q(M;W)$, the mutual information between meanings and words, quantifies the complexity of the lexicon, and $I_q(W;U)$ measures how much information about the environment is preserved through language. The trade-off parameter $\beta \geq 1$ controls the balance between compression and informativeness. By *approximately* minimizing this functional across a range of β values, the authors trace out the optimal trade-off curve in the two-dimensional space defined by informativeness and complexity. Using data from the World Color Survey (Cook et al., 2005), they demonstrate that natural languages approximate *near*-optimal solutions for color naming, close to the optimal trade-off frontier defined by the IB curve.

2.2 Optimality in Kinship Naming

Every society uses language to refer to family members, or *kin*, through a system of lexical items that categorize familial roles (e.g., father and sister). A relevant source of cross-linguistic variation lies in how kinship meaning space is partitioned, or in other words, which family members are considered part of the same semantic category. For example, in English, both maternal and paternal grandmothers fall under the same category (grandmother). In contrast, Vietnamese differentiates both lineage and age, employing distinct terms for maternal versus paternal grandparents, and for older versus younger siblings (Van Luong, 1989). At the other extreme, Tagalog collapses the gender distinction entirely, using a single term (kapatid) for both brother and sister (Murdock, 1970).

Kemp and Regier (2012) observe that kinship systems appear to reflect a trade-off between

informativeness—here, the ability to distinguish between kin members based on kinship names—and complexity. The latter relies on a symbolic rule system to characterize the meaning of kinship names through logical compositions of primitives (e.g., mother would be *PARENT & OLDER & FEMALE*). Complexity is then quantified as the minimal number of logical rules needed to generate the system. Informativeness is measured as the expected Kullback-Leibler (KL) divergence between intended and inferred referents, averaged over a communicative need distribution.

3 Framework

In this section, we introduce our information-theoretic framework that formalizes the trade-off in object naming. Specifically, we consider a scenario where there is a pool $\mathcal{U}$ of objects, associated with a communicative need distribution $p(\cdot)$. A Speaker aims to communicate about an object $u \sim p(u)$ by selecting a message (or name) $w \in \mathcal{W}$ with encoding probability $q_s(w|u)$. The Speaker sends this message to a Listener, who then attempts to infer the intended object using a decoder $q_l(u|w)$.

As illustrated in Figure 1, consider an English-speaking Speaker who wishes to refer to the object u = elder brother. Due to the structure of English kinship terminology, the most specific term available is w = "brother", and the Speaker deterministically selects it, i.e., $q_s(\text{"brother"}|\text{elder brother}) = 1$. Upon receiving the message, an English-speaking Listener infers that the referent could be either elder brother or younger brother, assigning equal probabilities: $q_l(\text{elder brother}|\text{"brother"}) = q_l(\text{younger brother}|\text{"brother"}) = 0.5$.

3.1 Complexity

Inspired by Zaslavsky et al. (2018), and viewing the framework through the lens of Shannon's communication model (Shannon, 1948), we quantify complexity as the amount of information the Speaker compresses through its encoder $q_s(w|u)$. This is measured by the mutual information between the object random variable U and the message random variable W:

$$C = I_{q_s}(U;W) = \sum_{u,w} p(u)\, q_s(w|u) \log \frac{q_s(w|u)}{p_s(w)}$$

where $p_s(w) = \sum_{u \in \mathcal{U}} p(u)\, q_s(w|u)$ is the marginal distribution over messages. Intuitively,

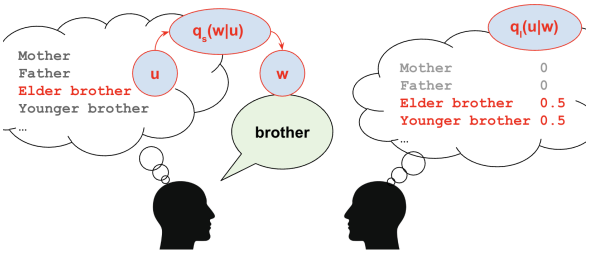

Figure 1: Illustration of two English-speaking agents playing the kinship naming game. The Speaker (left) selects a family member and produces a name. The Listener (right) receives the name and infers which member is being referred to.

this quantity captures the average amount of information that a message w conveys about the intended object u. If the same message is used for every meaning, mutual information is zero. Conversely, if each meaning is encoded with a distinct message, mutual information reaches its maximum value—the entropy $H(U)$. Thus, complexity is bounded by $0 \leq C \leq H(U)$.

Notably, this measure of complexity increases with lexical granularity: when the communication system encodes many fine-grained distinctions among meanings, the mutual information is high; when multiple meanings are collapsed onto a single message, the complexity is correspondingly low.

3.2 Information Loss

Information loss quantifies how much information is *not* preserved throughout the communication process, and is therefore directly related to the errors made by the Listener when picking a referent. We define information loss as the expected cross-entropy between the true referent and the Listener's prediction, i.e., the standard loss function in multi-class classification:

$$L = -\mathbb{E}_{u \sim p}\, \mathbb{E}_{w \sim q_s(\cdot|u)} \log q_l(u|w)$$

Intuitively, the more confident the Listener can be about the intended object u given the message w, the lower the information loss. Conversely, uncertainty in decoding leads to higher loss.

3.3 Optimality

Let $\tilde{q}_s(u|w) \propto q_s(w|u)\, p(u)$ be the Speaker's Bayesian decoder. We prove in Appendix A that:

$$\begin{aligned} L &= H(U) - C + \mathbb{E}_{w \sim p_s}\left[D_{\mathrm{KL}}(\tilde{q}_s \| q_l)\right] \\ &\geq H(U) - C \end{aligned} \tag{1}$$

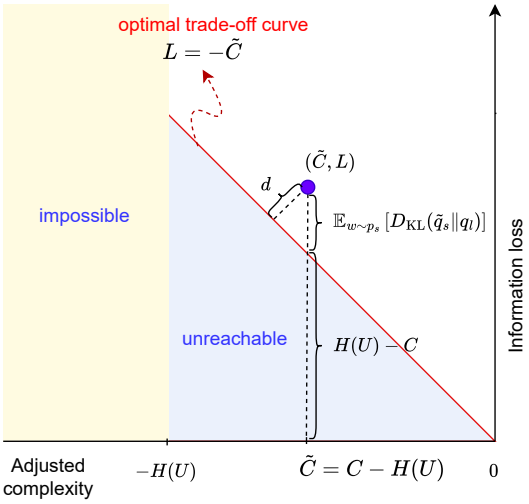

Figure 2: Visualization of the optimal trade-off curve (red line) and the feasible region (white area) encompassing all valid communication systems. The Euclidean distance from a system with complexity C and information loss L (the blue dot) to the curve is given by $d = \mathbb{E}_{w\sim p_s}\left[D_{\mathrm{KL}}(\tilde{q}_s\|q_l)\right]/\sqrt{2}$.

The inequality follows from the non-negativity of the KL divergence.

This proof is an important contribution due to the following reasons. First of all, a crucial result from this identity is that equality holds if and only if $q_l = \tilde{q}_s$. This implies that, for discrete object naming, the optimal trade-off is achieved when the Listener's decoder exactly matches the Bayesian decoder of the Speaker.

Second, unlike the frameworks proposed by Kemp and Regier (2012) and Zaslavsky et al. (2018), Equation 1 provides an *exact, closed-form* expression for the optimal trade-off curve. In addition, it offers a quantitative measure of deviation from this curve. Specifically, we can compute the Euclidean distance from a given communication system to the curve analytically by $d = \frac{L-H(U)+C}{\sqrt{2}}$.

3.4 Need-agnostic Optimal Curve

The curve $L = H(U) - C$ described above depends on the communicative need distribution $p(u)$, due to the inclusion of the entropy term $H(U)$. This dependency complicates cross-linguistic comparisons of optimality, as different need distributions induce different optimal curves in the complexity–information loss space. To address this issue, prior work—including Kemp and Regier (2012) and Zaslavsky et al. (2018)—has often assumed a universal, fixed need distribution across languages.

Our framework offers a more principled alternative: by defining an adjusted complexity $\tilde{C} = C - H(U) \leq 0$, we transform the optimal curve into the simplified form $L = -\tilde{C}$. This reformulation removes the dependence of the optimal trade-off curve on $p(u)$, while preserving Euclidean distance—thereby enabling meaningful comparisons across languages, regardless of their underlying communicative needs. Notably, the smaller $\tilde{C}$ is, the less complex the system is, and the more *capacity* it has to increase in complexity. When $\tilde{C} = 0$, the system reaches its maximum allowable complexity under the given need distribution.

Figure 2 illustrates the optimal trade-off curve, along with the feasible region in which all valid communication systems must lie. In Appendix B, we demonstrate that our framework is compatible with Zaslavsky et al. (2018)'s IB framework.

4 Kinship Case Study

We present a case study of object naming in the domain of kinship, based on the familial structure introduced by Kemp and Regier (2012). Illustrated in Figure 3, this structure includes 33 family members, spanning five generations, with a designated *ego* representing the speaker. Since kinship terms in some languages—such as Korean—depend on the (binarized) gender of the speaker, we consider two ego identities: Alice (female) and Bob (male).

We examine two types of kinship naming terminologies: human (i.e., based on a sample of natural languages) and neural network-based (i.e., emerging from neural-network (NN) agents simulations). In the former case, we assume a simple probabilistic model to formalize encoding and decoding of messages by Speaker and Listener, while in the latter, the encoder and decoder are learned with the same neural network agents, which develop their own kinship terminology while playing a referential game. We refer to these systems as HP (for Human-Probabilistic) and NN, respectively.

4.1 Human (HP) Kinship Systems

We investigate kinship naming across four natural languages—English, Dutch, Spanish, and Vietnamese. We estimate the communicative need distribution $p(u)$, the Speaker's encoder $q_s(w|u)$, and the corresponding Bayesian decoder $\tilde{q}_s(u|w)$ using frequency counts extracted from text corpora, as in Kemp and Regier (2012) (however, unlike this study, we estimate a separate need distribution for

each language).

Concretely, for each family member u, we compile a set of commonly used referring expressions $T(u)$, e.g., "mother," "mommy," "mom" for mother. For each term $w \in T(u)$, we estimate count(u, w)—the number of times w refers to u—by searching the corpora using possessive constructions with the first person singular pronoun in each language, such as "*my* mother". If a term is polysemous (e.g., "brother" can refer to either elder brother or younger brother), the count is evenly divided among all plausible referents u.

We then compute the total frequency of each referent u by summing the counts across all its corresponding terms

$$\text{count}(u) = \sum_{w \in T(u)} \text{count}(u, w)$$

Let $\mathcal{F}$ be the set of all family members excluding ego, we finally estimate the required distributions

$$p(u) = \frac{\text{count}(u)}{\sum_{v \in \mathcal{F}} \text{count}(v)} \; ; \quad q_s(w|u) = \frac{\text{count}(u, w)}{\text{count}(u)}$$

$$\tilde{q}_s(u|w) = \frac{q_s(w|u)p(u)}{\sum_{v \in \mathcal{F}} q_s(w|v)p(v)}$$

Further details regarding the corpora used and the counts are presented in Appendix G.

4.2 Emergent (NN) Kinship Systems

In order to prompt the emergence of a neural-network based kinship system, we frame the model task as a referential game in which the referents are family members (see Figure 1). The NN-Speaker is given access to the full family tree along with a randomly selected target individual. Based on this input, the Speaker generates a message intended to identify the target. Upon receiving the message and observing a candidate set that includes the target, the NN-Listener must infer which candidate the NN-Speaker is referring to.

4.2.1 Input encoding

Most of the literature in language emergence uses input representations based on images (Havrylov and Titov, 2017; Lazaridou et al., 2017; Evtimova et al., 2018; Bouchacourt and Baroni, 2018) or feature vectors (Kottur et al., 2017; Chaabouni et al., 2020). The tree-like structure of a family tree, however, motivates the use of a structure that is more akin to trees–such as graphs.

A kinship graph consists of 33 nodes, including a designated *ego* node, which can be either "Bob" (male) or "Alice" (female). The graph is adapted from the used family tree and is visualized in Figure 3. Each node represents an individual family member and is annotated with categorical features that encode key relational distinctions:

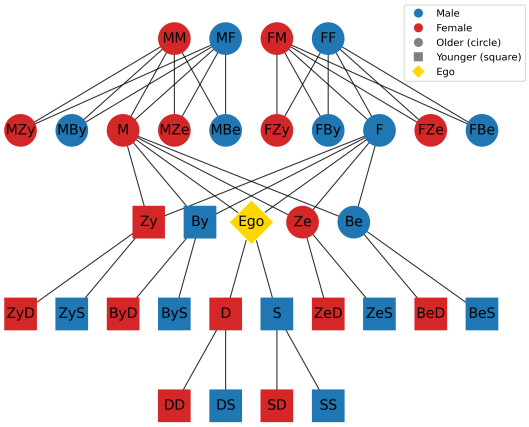

Figure 3: The kinship graph is adapted from the familial structures described by Kemp and Regier (2012). Nodes are labeled using abbreviations, where "F", "M", "B", "Z", "S", "D", "y", and "e" stand for "father", "mother", "brother", "sister", "son", "daughter", "younger", and "elder", respectively. For example, MBe denotes the "mother's elder brother". Each edge in the graph is bidirectional, labeled parent-of when traversing top-down and child-of when traversing bottom-up.

- *Gender* (male or female),[1]
- *Gender relative to ego* (equal or different),
- *Age relative to ego* (older or younger),
- *Age relative to parent* (older or younger)

All features are one-hot encoded and are designed to reflect the compositional kinship semantics used in Kemp and Regier (2012).

To connect nodes (i.e., family members), we diverge from Kemp and Regier (2012) by using only the two most primitive relationships—parent-of and child-of—which allow bidirectional traversal across generations. For example, the node F (father) connects to Bob (ego) via "F is parent of Bob," and to Be (Bob's elder brother) via "F is parent of Be"; correspondingly, Bob and Be each connect back to F via child-of edges. More complex relationships in the kinship trees of Kemp and Regier (2012), such as sibling-of, must instead be inferred compositionally from the primitive relations. This design encourages agents to discover

[1]We consider only binary gender distinctions, consistent with those typically encoded in human kinship systems.

and exploit relational structure rather than rely on shortcut or explicitly labeled edges.

We observe that the kinship graph contains redundant information. For instance, determining whether two individuals are brothers can be achieved by checking whether they share either the same mother or the same father. To eliminate such redundancy, we apply a pruning procedure that retains only the shortest paths from each node to the ego. This results in a best-first-search tree. We refer to this process as *graph pruning*.

4.2.2 Model architecture

Our agents are implemented as two neural networks. The NN-Speaker encodes the kinship graph G using a graph neural network GNN_s, producing node-level embeddings:

$$[\mathbf{h}_1, \ldots, \mathbf{h}_{33}] = \text{GNN}_s(G)$$

where $\mathbf{h}_i \in \mathbb{R}^d$ is the representation of the i-th node. To generate a message, the Speaker concatenates the embeddings of the ego and target nodes, transforms them via a two-layer network, and applies the Gumbel-Softmax (GS) to sample a *one-token* message w from a fixed vocabulary $\mathcal{V}$:

$$\begin{aligned} \text{score}_s(u) &= \mathbf{W}_{\text{lex}} \cdot \mathbf{W}_{\text{hid}} \cdot \text{cat}(\mathbf{h}_{\text{ego}}, \mathbf{h}_{\text{target}}) \\ w &\sim \text{GS}(\text{score}_s(u)) \end{aligned}$$

Here, $\mathbf{W}_{\text{hid}} \in \mathbb{R}^{d_h \times 2d}$ and $\mathbf{W}_{\text{lex}} \in \mathbb{R}^{|\mathcal{V}| \times d_h}$ are trainable weight matrices, and $\text{score}_s(u) \in \mathbb{R}^{|\mathcal{V}|}$ denotes the unnormalized scores for tokens. Unless stated otherwise, we use a vocabulary of size $|\mathcal{V}| = 128$ and a Gumbel-Softmax temperature of 1.5.

The NN-Listener receives the same graph along with the sampled message w and infers the target referent. It encodes the graph using another graph neural network GNN_l, identical in architecture to the Speaker's, and computes compatible scores between the message and family members:

$$\begin{aligned} [\mathbf{v}_1, \ldots, \mathbf{v}_{33}] &= \text{GNN}_l(G) \\ \text{score}_l(w, i) &= \mathbf{e}_w^\top \cdot \mathbf{W} \cdot \mathbf{v}_i \quad \forall i \in [1, \ldots, 33] \end{aligned}$$

where $\mathbf{v}_i \in \mathbb{R}^d$ is the embedding of node i, $\mathbf{e}_w \in \mathbb{R}^{d_h}$ is the embedding of token w, and $\mathbf{W} \in \mathbb{R}^{d_h \times d}$ is a trainable bilinear transformation. The Listener selects the node with the highest score as its prediction. In all experiments, we use three shared-parameter graph neural layers, set $d = 80$ and $d_h = 20$, and the NN-Speaker and NN-Listener do not share parameters.

Graph neural networks (GNNs) serve as the backbone of both agents, enabling the processing of structured relational data and facilitating the emergence of compositional communication. Given the inherently relational nature of kinship structures, we adopt RGCN (Schlichtkrull et al., 2018), which is specifically designed for multi-relational graphs. In addition, we explore alternative GNN architectures and hyperparameter configurations, as detailed in Appendix E.

4.2.3 Training

From the kinship graph described above, we generate a dataset of 10,000 data points, each corresponding to a single game turn. Each data point consists of: (i) the full pruned kinship graph with the ego node uniformly sampled from $\{\text{Bob}, \text{Alice}\}$; (ii) a target node u uniformly selected from the remaining 32 nodes; (iii) a distractor set D of five nodes, uniformly sampled from the remaining nodes (i.e., excluding both the ego and the target node).

We split the dataset into 80% for training and 20% for validation. The two agents are trained jointly to minimize the following loss:

$$L = -\sum_{(u,D)} \log \frac{\exp(\text{score}_l(w, u))}{\sum_{v \in \{u\} \cup D} \exp(\text{score}_l(w, v))}$$

where w is the message generated by the Speaker for target u.[2] We use the Adam optimizer with a learning rate of 1×10^{-3}, training for 500 epochs using mini-batches of size 50.

4.2.4 Evaluation

To assess the generalizability of the learned communication protocol, we evaluate the agents every 5 epochs on a more challenging setting than used during training or validation. Specifically, the NN-Listener must identify the target node from among *all* 32 possible family members (rather than from a limited set of six candidates, as done during training). The evaluation dataset thus consists of $2 \times 32 = 64$ data points, corresponding to each combination of $\text{ego} \in \{\text{Bob}, \text{Alice}\}$ and the 32 non-ego family members as targets.

Unlike during training, the NN-Speaker operates deterministically at evaluation time, producing the most likely token: $w^* = \arg\max_w \text{score}_s(u)[w]$; hence $q_s(w|u) = [[w = w^*]]$. The NN-Listener

[2] Note that during training, w sampled from the Gumbel-Softmax is a distribution over $\mathcal{V}$, rather than a discrete token as in evaluation.

computes a probability distribution over all candidates using a softmax over the score function: $q_l(u|w) \propto \exp(\text{score}_l(w, u))$.

5 Experiments

We conduct computational experiments to validate our theoretical framework on the optimality of kinship naming across four languages: English, Dutch, Spanish, and Vietnamese.

5.1 Systems

For each natural language, we evaluate kinship naming between an HP-Speaker and one of the following versions of the HP-Listener:

- *Optimal* HP-Listener, which employs a decoder identical to the Bayesian decoder used by the HP-Speaker;
- *Noise_*r_e HP-Listener ($r_e \in (0, 1]$), a variant of the optimal Listener that introduces random errors by misinterpreting a family member as another with probability r_e (i.e., with r_e error rate). The higher the value of r_e, the more the HP-Listener deviates from the optimal one.

Each noisy (or suboptimal) HP-Listener is simulated with a population of 10,000 listeners to ensure reliable performance estimates.

We also evaluate the communication performance of our neural network agents (NN) under each language-specific communicative need distribution.[3] The communication setup is ego-specific, with ego $\in \{\text{Bob}, \text{Alice}\}$. Following the configuration described in Section 4 and summarized in Appendix C, we repeat training and evaluation 50 times to account for variability across runs.

5.2 Metrics

Since our primary interest lies in the optimality of the complexity-information loss tradeoff, our primary performance metric is the Euclidean *distance* to the optimal curve, defined in Section 3.3.

In addition, under each language's communicative need distribution, we evaluate *accuracy*, defined as the expected probability that a Listener correctly identifies the intended referent family member: $\mathbb{E}_{u\sim p,\, w\sim q_s(\cdot|u)}\left[q_l(u|w)\right]$. This is a *relaxed* variant of traditional accuracy, since the precise predictions of (natural language) Listeners are not directly observable due to polysemy (e.g., the term "brother" may refer to either elder brother or younger brother).

[3] We use the EGG framework (Kharitonov et al., 2019) and the PyTorch Geometric library (Fey and Lenssen, 2019). Our source code is publicly available at abc.anonymized.xyz.

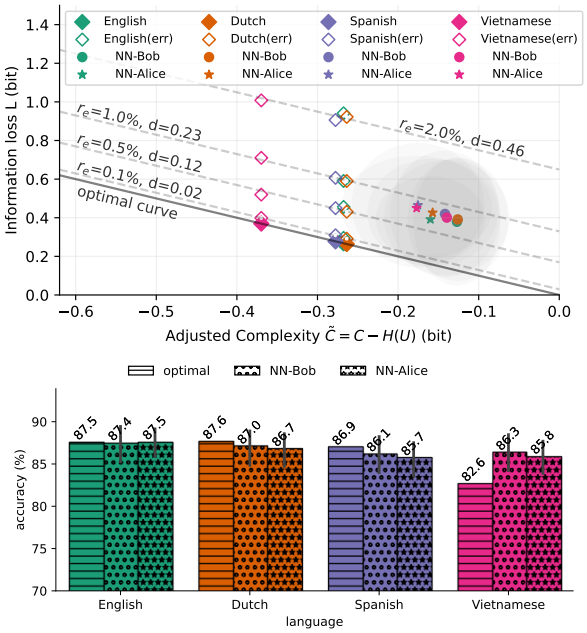

Figure 4: (Top) Trade-offs of HP and NN systems, averaged over 50 runs (standard deviation shown as light gray ellipses). Trade-offs under suboptimal HP-Listener conditions collectively form lines approximately parallel to the optimal curve. Lines are annotated with error rate r_e and distance d to the optimal curve. (Bottom) Accuracy of NN and human systems with optimal HP-Listener.

5.3 Results

The impact of sub-optimal HP-Listeners. We evaluate four suboptimal HP-Listeners with error rates $r_e \in \{0.1\%, 0.5\%, 1\%, 2\%\}$. Figure 4-top shows the trade-off under each of these HP-Listener conditions, plus the condition without noise (optimal HP-Listener). The figure reveals that, across all languages, communication with the *optimal* HP-Listener lies exactly on the optimal curve (i.e., zero distance), while communication with noisier Listeners deviates progressively further from it. Moreover, HP-Listeners with lower error rates consistently yield more favorable trade-offs than those with higher error rates. These findings align with the theoretical prediction outlined in Section 3.3, which posits that trade-off optimality depends on how closely the HP-Listener's decoder approximates the Bayesian decoder of the HP-Speaker.

Interestingly, we observe that human communication systems in the three indoeuropean languages—English, Dutch, and Spanish—exhibit

similar levels of adjusted complexity and information loss (and also accuracy as shown in Figure 4-bottom). In contrast, Vietnamese shows higher information loss but lower adjusted complexity, suggesting that it has more capacity to increase complexity in order to improve informativeness.

HP vs NN. Since the NN models are trained to minimize information loss, they are not inherently constrained by complexity. Fortunately, we find that early stopping serves as an effective mechanism for limiting complexity, therefore we select the checkpoint with accuracy closest to that of human communication (see Figure 4-bottom).

Figure 4-top shows that the NN communication system achieves trade-offs that are closer to the optimal curve than human communication with the *error_1%* HP-Listener. This suggests that communication systems emerging from object naming tasks can approach theoretical optimality without incurring exhaustive complexity—similar to human communication.[4]

NN naming system evolves towards optimality. Building on the previous results, we examine the evolution of NN communication over time. Figure 5 shows the trade-offs under the communicative need distribution for English (results for other languages are in Appendix D).

Initially, the emergent communication exhibits moderate complexity but high information loss, as agents rely on only a few vocabulary terms, resulting in a simple but inefficient language. As training progresses, information loss gradually decreases at the cost of increasing complexity, with the trade-off trajectory approaching the optimal curve. This trend suggests that agents progressively expand their vocabulary usage to enhance communicative success, thereby reducing information loss while incurring greater complexity.

6 Discussion and Conclusion

We have presented an information-theoretic framework to measure the trade-off between information loss and complexity in discrete object naming systems. We have shown that optimality is reached only when assuming a perfect HP-Listener. This has consequences for communication settings in which perfect message reception cannot be guaranteed, which is in fact common in noisy realistic scenarios; in fact, even human listeners with clinically normal hearing struggle with decoding in everyday communication (Ruggles et al., 2011). Our framework allows for the estimation of trade-offs that systems can achieve under realistic conditions, whether in human communication or in applied settings where message passing between agents may be subject to perturbations.

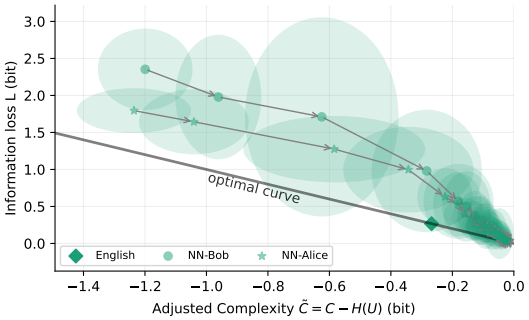

Figure 5: Trajectories of NN systems over time, recorded every 10 epochs, under the English communicative need distribution. Results are averaged over 50 runs, with standard deviation represented by ellipses.

The human kinship systems we have analyzed abide to the above: under the assumption of an optimal HP-Listener, these systems are exactly optimal. Interestingly, the trade-offs vary between language families, with Vietnamese affording lower complexity than the indoeuropean languages, at the cost of informativeness. The reason for cross-linguistic contrasts must lie either on the semantic partitions into kinship categories, or on differences in communicative need. We leave this analysis to future work, along with an expansion on the analyzed languages and linguistic families.

Finally, we have shown that NN models exhibit a trade-off comparable to that of HP systems with less than 1% noise. This suggests that optimizing only for communication accuracy (or information loss) and applying early stopping are sufficient mechanisms to trigger the learning and evolutionary dynamics that result in the observed trade-offs.

Overall, our framework allows us to characterize discrete naming systems in general, and analyze kinship naming systems in particular, both as found in human language and as emerging from communication games. Our results bode well for the use of emergent communication setups to develop efficient naming systems in any applied settings.

[4]Note that information loss is not an absolute indicator of accuracy—for example, in the case of Vietnamese. This is analogous to what is commonly observed between log-likelihood and accuracy in classification tasks.

7 Limitations

As in Zaslavsky et al. (2018)'s Information Bottleneck framework, we measure complexity as the mutual information between the object and word random variables. This quantity is upper-bounded by the entropy of the communicative need distribution. However, this measure does not fully capture the richness of natural languages: in principle, a language can achieve unbounded complexity by continually expanding its vocabulary.

We restrict our NN communication systems to *one-token* messages, whereas natural languages often employ compositional expressions to refer to kinship relations—i.e., components of the expression systematically correspond to semantic properties of the kin category. For example, in English, the prefix *grand-* consistently denotes *parent-of-*, while in Spanish, the suffixes *-a* and *-o* typically indicate female and male family members, respectively. Such compositional strategies cannot emerge in our current setup. Additionally, because the NN-Speaker generates names deterministically, the emergent language lacks synonymy, which is common in natural languages (e.g., in English mother may be referred to as either "mother" or "mum").

References

Diane Bouchacourt and Marco Baroni. 2018. How agents see things: On visual representations in an emergent language game. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 981–985, Brussels, Belgium. Association for Computational Linguistics.

Shaked Brody, Uri Alon, and Eran Yahav. 2022. How attentive are graph attention networks? In *International Conference on Learning Representations*.

Jon W. Carr, Kenny Smith, Jennifer Culbertson, and Simon Kirby. 2020. Simplicity and informativeness in semantic category systems. *Cognition*, 202.

Rahma Chaabouni, Eugene Kharitonov, Diane Bouchacourt, Emmanuel Dupoux, and Marco Baroni. 2020. Compositionality and Generalization In Emergent Languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4427–4442, Online. Association for Computational Linguistics.

Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. 2021. Communicating artificial neural networks develop efficient color-naming systems. *Proceedings of the National Academy of Sciences*, 118(12).

Rahma Chaabouni, Florian Strub, Florent Altché, Eugene Tarassov, Corentin Tallec, Elnaz Davoodi, Kory Wallace Mathewson, Olivier Tieleman, Angeliki Lazaridou, and Bilal Piot. 2022. Emergent Communication at Scale. In *International Conference on Learning Representations*.

Richard Cook, Paul Kay, and Terry Regier. 2005. *The World Color Survey Database*, pages 223–VII.

Mark Davies. 2002–2024. Corpus del español. Online linguistic resource providing historical and modern Spanish texts.

Mark Davies. 2010. The corpus of contemporary american english as the first reliable monitor corpus of english. *LLC*, 25:447–464.

Katrina Evtimova, Andrew Drozdov, Douwe Kiela, and Kyunghyun Cho. 2018. Emergent Communication in a Multi-Modal, Multi-Step Referential Game. In *International Conference on Learning Representations*.

Matthias Fey and Jan Eric Lenssen. 2019. Fast graph representation learning with pytorch geometric. *arXiv preprint arXiv:1903.02428*.

Edward Gibson, Richard Futrell, Steven P Piantadosi, Isabelle Dautriche, Kyle Mahowald, Leon Bergen, and Roger Levy. 2019. How efficiency shapes human language. *Trends in cognitive sciences*, 23(5):389–407.

Serhii Havrylov and Ivan Titov. 2017. Emergence of Language with Multi-agent Games: Learning to Communicate with Sequences of Symbols. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.

Charles Kemp and Terry Regier. 2012. Kinship categories across languages reflect general communicative principles. *Science*, 336:1049–1054.

Charles Kemp, Yang Xu, and Terry Regier. 2018. Semantic typology and efficient communication. *Annual Review of Linguistics*, 4:109–128.

Eugene Kharitonov, Rahma Chaabouni, Diane Bouchacourt, and Marco Baroni. 2019. EGG: a toolkit for research on Emergence of lanGuage in Games. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations*, pages 55–60, Hong Kong, China. Association for Computational Linguistics.

Satwik Kottur, José Moura, Stefan Lee, and Dhruv Batra. 2017. Natural Language Does Not Emerge 'Naturally' in Multi-Agent Dialog. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2962–2967, Copenhagen, Denmark. Association for Computational Linguistics.

Angeliki Lazaridou and Marco Baroni. 2020. Emergent multi-agent communication in the deep learning era. *CoRR*, abs/2006.02419.

Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. 2017. Multi-Agent Cooperation and the Emergence of (Natural) Language. In *International Conference on Learning Representations*.

George Peter Murdock. 1970. Kin term patterns and their distribution. *Ethnology*, 9(2):165–208.

Nelleke Oostdijk, Martin Reynaert, Véronique Hoste, and Ineke Schuurman. 2013. The construction of a 500-million-word reference corpus of contemporary written Dutch. In *Essential Speech and Language Technology for Dutch: Results by the STEVIN-programme*, chapter 13. Springer Verlag.

Nam Pham. 2024. vietvault (revision bee0136).

Terry Regier, Charles Kemp, and Paul Kay. 2015. Word meanings across languages support efficient communication. *The handbook of language emergence*, pages 237–263.

Dorea Ruggles, Hari Bharadwaj, and Barbara G. Shinn-Cunningham. 2011. Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proceedings of the National Academy of Sciences*, 108(37):15516–15521.

Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *15th Extended Semantic Web Conference (ESWC)*, pages 593–607. Springer, Cham.

Claude E Shannon. 1948. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423.

Naftali Tishby, Fernando C Pereira, and William Bialek. 2000. The information bottleneck method. *arXiv preprint physics/0004057*.

Hy Van Luong. 1989. Vietnamese kinship: structural principles and the socialist transformation in northern vietnam. *The Journal of Asian Studies*, 48(4):741–756.

Noga Zaslavsky, Charles Kemp, Terry Regier, and Naftali Tishby. 2018. Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences*, 115:7937–7942.

Noga Zaslavsky, Terry Regier, Naftali Tishby, and Charles Kemp. 2019. Semantic categories of artifacts and animals reflect efficient coding. In *Proceedings of the 41th Annual Meeting of the Cognitive Science Society, CogSci 2019: Creativity + Cognition + Computation, Montreal, Canada, July 24-27, 2019*, pages 1254–1260. cognitivesciencesociety.org.

A A Lower bound for Information Loss

In this appendix we derive a lower bound for information loss.

$$
\begin{aligned}
L =& -\mathbb{E}_{u\sim p}\mathbb{E}_{w\sim q_s(\cdot|u)} \log q_l(u|w) \\
=& -\sum_u p(u) \sum_w q_s(w|u) \log q_l(u|w) \\
=& -\sum_u p(u) \sum_w q_s(w|u) \\
& \log \left[p(u) \frac{q_s(w|u)}{p_s(w)} \cdot \frac{q_l(u|w) p_s(w)}{q_s(w|u) p(u)} \right] \\
=& -\sum_u p(u) \sum_w q_s(w|u) \log p(u) \\
& -\sum_u p(u) \sum_w q_s(w|u) \log \frac{q_s(w|u)}{p_s(w)} \\
& -\sum_u p(u) \sum_w q_s(w|u) \log \frac{q_l(u|w) p_s(w)}{q_s(w|u) p(u)} \\
=& -\sum_u p(u) \log p(u) \\
& -\sum_{u,w} p(u) q_s(w|u) \log \frac{q_s(w|u)}{p_s(w)} \\
& +\sum_w p_s(w) \sum_u \tilde{q}_s(u|w) \log \frac{\tilde{q}_s(u|w)}{q_l(u|w)} \\
=& H(U) - C + \mathbb{E}_{w\sim p_s} \left[D_{\mathrm{KL}}(\tilde{q}_s \| q_l) \right]
\end{aligned}
$$

where:

- $H(U) = -\sum_u p(u) \log p(u)$ is the entropy of the communicative need distribution;
- C is the complexity, as defined in Section 3.1;
- $\tilde{q}_s(u|w) = \frac{q_s(w|u)\, p(u)}{p_s(w)}$ is the Bayesian decoder of the Speaker.

Since the KL divergence is always non-negative, the information loss is lower-bounded by:

$$L \geq H(U) - C.$$

This bound is achieved when $q_l = \tilde{q}_s$, i.e., when the Listener's decoder is identical to the Bayesian decoder of the Speaker.

B Compatibility with the Information Bottleneck Framework (Zaslavsky et al., 2018)

In Zaslavsky et al. (2018)'s Information Bottleneck (IB) framework for color naming (Figure 6), the Speaker and Listener communicate about colors $u \in \mathcal{U}$, where $\mathcal{U}$ represents a continuous perceptual space. Upon perceiving a color u, the Speaker selects a meaning m, modeled as a distribution over $\mathcal{U}$, and then generates a name w using the encoder $q_s(w|m)$. The Listener decodes the message using:

$$\hat{m}_w(u) = \sum_m \tilde{q}_s(m|w)\, m(u),$$

which defines the *Bayesian-optimal listener*.

In this framework, assuming a Bayesian-optimal Listener, complexity is quantified as the mutual information between the Speaker's meaning variable M and the word variable W:

$$I_{q_s}(M;W) = \sum_{m,w} p_s(m)\, q_s(w|m) \log \frac{q_s(w|m)}{p_s(w)},$$

while informativeness is captured by the mutual information $I_{q_s}(W;U)$, measuring how much information the word conveys about the original object.

Following the IB principle, an optimal trade-off between complexity and informativeness is achieved by minimizing the following objective:

$$\mathcal{F}_\beta[q_s(w|m)] = I_{q_s}(M;W) - \beta I_{q_s}(W;U),$$

where $\beta \geq 1$ is a trade-off parameter that balances compression and informativeness.

When adapting the IB framework to a discrete domain such as kinship, the objects u are inherently discrete. In this setting, we can assume a one-to-one correspondence between the object set $\mathcal{U}$ and the agents' meaning space, allowing us to conflate u and m, as well as the corresponding random variables U and M. This assumption is consistent with Zaslavsky et al. (2018), who, in their color naming experiment, discretize the color space into a finite set of color chips, each of which is mapped to a distinct meaning. Under this assumption, the Bayesian Listener's decoder simplifies to:

$$\hat{m}_w(u) = \tilde{q}_s(u|w),$$

which is identical to the Bayesian decoder of the Speaker in our framework.

In this discrete setting, complexity and informativeness converge to the same quantity, and the IB objective reduces to:

$$
\begin{aligned}
\mathcal{F}_\beta[q_s(w|m)] &= I_{q_s}(U;W) - \beta I_{q_s}(W;U) \\
&= (1-\beta) I_{q_s}(W;U).
\end{aligned}
$$

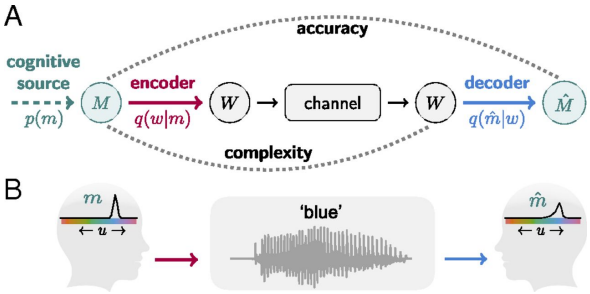

Figure 6: (A) Schematic of the communication model from Zaslavsky et al. (2018). (B) Example of color communication. (Figure adapted from Zaslavsky et al. (2018))

Two cases arise: (i) If $\beta = 1$, the objective equals zero, regardless of the system's informativeness or complexity; (ii) If $\beta > 1$, the objective is minimized when the system achieves maximal complexity, which corresponds to the entropy $H(U)$ of the object distribution. However, since natural languages tend to balance informativeness with efficiency rather than maximize complexity, the latter case is not of primary interest.

The first case ($\beta = 1$) clearly establishes that a system in which the Listener's decoder matches the Bayesian decoder of the Speaker achieves an optimal trade-off. This outcome is fully compatible with our theoretical framework. Nonetheless, the IB framework, by using $I_{q_s}(W;U)$ to measure informativeness, does not account for the impact of suboptimal listeners. Moreover, it assumes a fixed communicative need distribution $p(u)$, and thus does not capture cross-linguistic variability in communicative demands.

C Hyper-parameters

We show relevant hyper-parameters for all experiments in Table 1. Gumbel-softmax temperature controls the Gumbel-softmax sampling distribution: lower values tend towards a one-hot encoding, whereas higher values tend towards a uniform encoding.

D The evolution of NN Kinship System

We show in Figure 7 how NN communication evolved during training in the environments of the four languages: English, Dutch, Spanish, and Vietnamese.

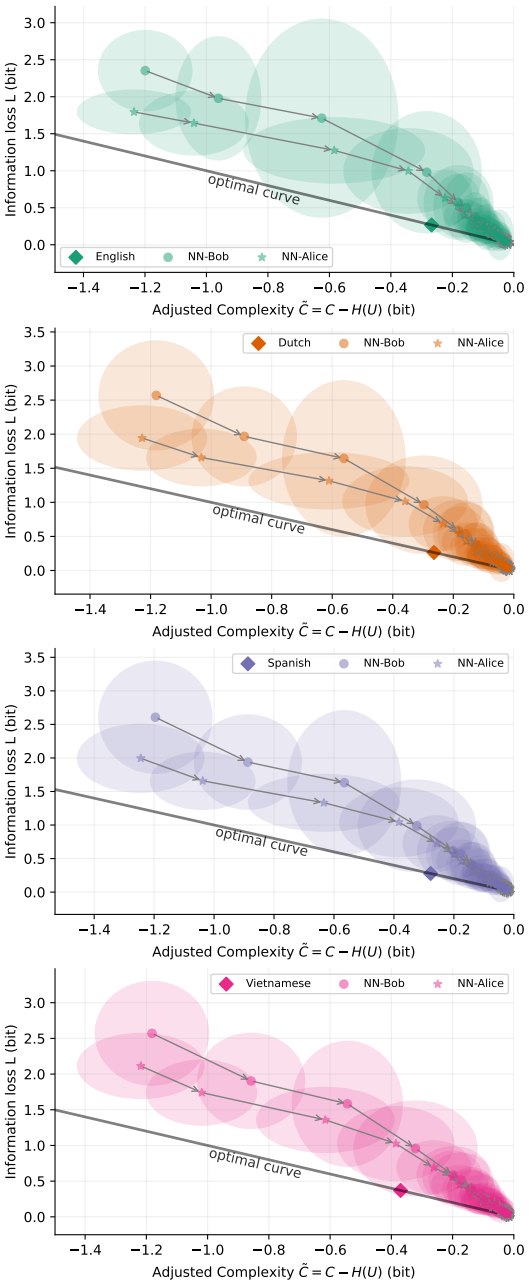

Figure 7: Evolution of complexity and information loss in ego-specific NN communication (average over 50 runs) under four communicative need distributions English, Dutch, Spanish, and Vietnamese.

Table 1: Hyperparameter settings used in our experiments. The third column reports the values selected for the main study, while the last column lists the values considered during architecture search.

	Hyperparameter	Value (main study)	Values (architecture search)
Model Architecture	embedding dimensions d	80	–
	hidden dimension d_h	20	–
	Graph neural net	RGCN	RGCN, GATv2Conv
	# graph net layers	3	–
	Vocabulary size $\|\mathcal{V}\|$	128	16, 32, 64, 128, 256
	Graph pruning	True	True, False
Training	Optimizer	Adam	–
	Learning rate	1×10^{-3}	–
	Batch size	50	–
	# distractors	5	–
	Gumbel-softmax temperature	1.5	–

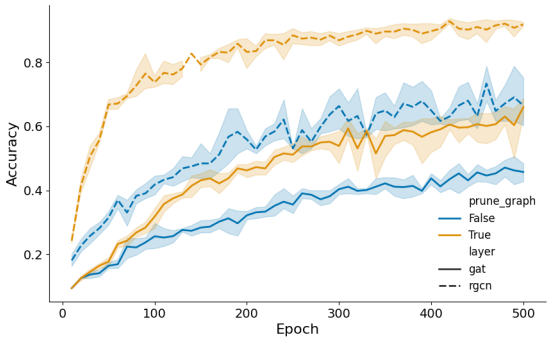

Figure 8: Evaluation accuracy, for simulations with and without pruning (n=40 runs, varying initialization).

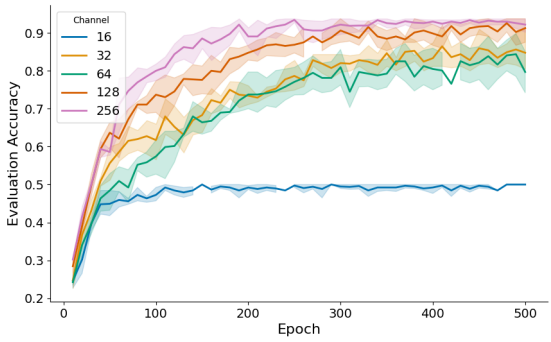

Figure 9: Evaluation accuracy, for simulations with pruning and RGCN layer (n=50 runs in total, with 10 different initializations for each vocabulary size).

E Impact of Model Architecture on Performance

We investigate the impact of architectural choices that led to the NN-agents used in the main study. Specifically, we examine three factors: *graph pruning*, *layer type*, and *channel capacity* (i.e., vocabulary size) in the environment of uniform communicative need distribution.

In the first study, based on the configuration described in Section 4 and summarized in Appendix C, we construct four variants by systematically varying the use of graph pruning and the choice of layer type (either RGCN (Schlichtkrull et al., 2018) or GATv2Conv (Brody et al., 2022)). As shown in Figure 8, both graph pruning and the use of RGCN layers are critical for achieving high communicative success, each contributing approximately 20 percentage points to the final communication accuracy of the NN-agents.

In the second study, we vary the vocabulary size (16, 32, 64, 128, 256) to examine the effect of channel capacity. Figure 9 demonstrates that a sufficiently large vocabulary relative to the size of the object set (32 kinship terms in our case) is crucial. For instance, a vocabulary size of 16 constrains communication to approximately 50% accuracy. In contrast, increasing the vocabulary size to 32 or greater substantially improves accuracy to 80% and above.

F Human Communication with Suboptimal HP-Listeners

We investigate the impact of suboptimal HP-Listeners on human communication across four language environments: English, Dutch, Spanish, and Vietnamese. Figure 10 reports both the distance to the optimal trade-off curve and the accuracy of the corresponding communication systems. The results reveal a linear relationship between the HP-Listener's error rate and both distance and accuracy. This finding supports our theoretical framework, confirming that communication with noisier HP-Listeners—i.e., those that deviate more from the optimal Listener—results in trade-offs

that lie further from the optimal curve.

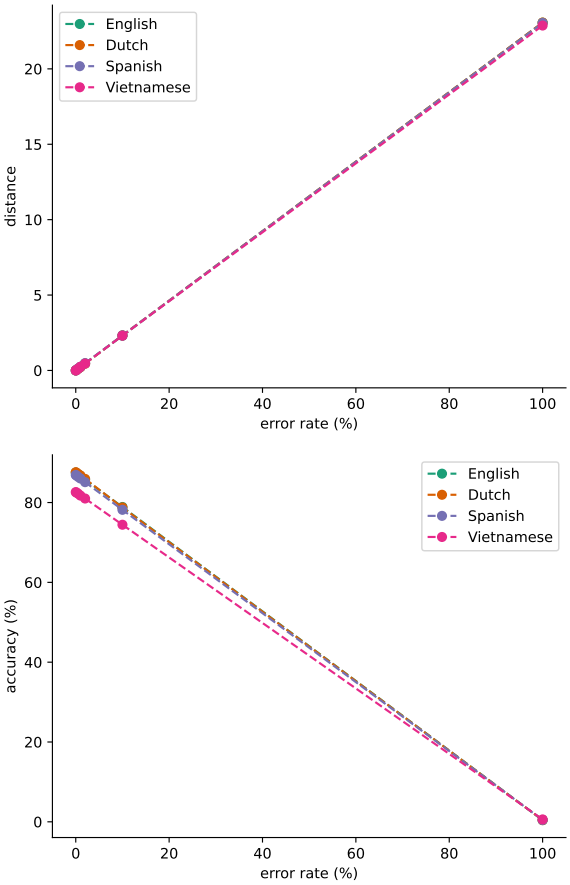

Figure 10: (Top) Distance to the optimal curve and (Bottom) accuracy of human communication with suboptimal Listeners at varying error rates $r_e \in \{0, 0.001, 0.005, 0.01, 0.02, 0.1, 1.0\}$. Dashed lines indicate the linear relationship between error rate and distance/accuracy.

G Kinship counts

We extract counts of family-member and kinship-term pairs (see Table 2) from text corpora in four languages: English, Dutch, Spanish, and Vietnamese.

- **English:** We use the Corpus of Contemporary American English (COCA) (Davies, 2010), a widely-used and balanced corpus of American English. It contains over one billion words from 1990–2019, covering eight genres such as spoken language, fiction, news, academic writing, and web content.

- **Dutch:** We use the SoNaR corpus (Oostdijk et al., 2013), a 500-million-word reference corpus of contemporary Dutch that includes both written and spoken data. SoNaR integrates material from various sources such as newspapers, newsletters, books, websites, and transcripts, offering broad coverage of modern Dutch across genres.

- **Spanish:** We use the NOW (News on the Web) corpus from the Corpus del Español (Davies, 2002–2024), which includes approximately 7.6 billion words from web-based newspapers and magazines across 21 Spanish-speaking countries, collected between 2012 and 2019. This corpus provides broad coverage of modern written Spanish as used in news media.

- **Vietnamese:** We use the VietVault corpus (Pham, 2024), a dataset filtered and curated from Common Crawl dumps prior to 2023. The full corpus contains 80GB of raw Vietnamese text spanning multiple domains. For our analysis, we sample a 5GB subset from the corpus to extract counts.

Table 2: Counts for (family-member, term) pairs from text corpora.

			English		Dutch		Spanish		Northern Vietnamese	
Gen.	**Label**	**Full Name**	**Term**	**Count**	**Term**	**Count**	**Term**	**Count**	**Term**	**Count**
1st	MM	mother's mother	Grandmother Grandma Gran	3949 882 25.5	Grootmoeder Oma	529 548.5	Abuela Yaya	8577.5 10	bà ngoại bà	210 411.5
	MF	mother's father	Grandfather Grandpa	3189 399.5	Grootvader Opa	720 361.5	Abuelo Yayo	8075.5 2	ông ngoại ông	125 1210
	FM	father's mother	Grandmother Grandma Gran	3949 882 25.5	Grootmoeder Oma	529 548.5	Abuela Yaya	8577.5 10	bà nội bà	266 411.5
	FF	father's father	Grandfather Grandpa	3189 399.5	Grootvader Opa	720 361.5	Abuelo Yayo	8075.5 2	ông nội ông	290 1210
2nd	M	mother	Mother Mom Mommy Momma	65458 29849 707 388	Moeder Mama Ma	18009 1511 1186	Madre Mama Mami Mamá	64464 1295 1071 52914	mẹ	18004
	F	father	Father Dad Daddy	63318 31100 3017	Vader Papa Pa	19939 896 1346	Padre Papa Papi Papá	77214 1250 512 47494	bố cha	5021 3390
	MZy	mother's younger sister	Aunt Auntie	1022 29	Tante	204.75	Tía Tita	1404.75 1.5	dì	243
	MBy	mother's younger brother	Uncle	1501.75	Oom	509	Tío Tito	1404.75 2.5	cậu	237
	MZe	mother's elder sister	Aunt Auntie	1022 29	Tante	204.75	Tía Tita	1404.75 1.5	bác gái bác	2.5 96.5
	MBe	mother's elder brother	Uncle	1501.75	Oom	509	Tío Tito	1404.75 2.5	bác trai bác	2 96.5
	FZy	father's younger sister	Aunt Auntie	1022 29	Tante	204.75	Tía Tita	1404.75 1.5	cô	435
	FBy	father's younger brother	Uncle	1501.75	Oom	509	Tío Tito	1404.75 2.5	chú	388
	FZe	father's elder sister	Aunt Auntie	1022 29	Tante	204.75	Tía Tita	1404.75 1.5	bác gái bác	2.5 96.5
	FBe	father's elder brother	Uncle	1501.75	Oom	509	Tío Tito	1404.75 2.5	bác trai bác	2 96.5
3rd	Zy	younger sister	Sister Sis	9587.5 43.5	Zus Zusje	2029.5 826	Hermana Tata	13610.5 82	em em gái	2425.5 849
	By	younger brother	Brother Bro	11687 63.5	Broer Broertje	2968 565	Hermano Tete	22723 4	em em trai	2425.5 459
	Ze	elder sister	Sister Sis	9587.5 43.5	Zus	2029.5	Hermana Tata	13610.5 82	chị	1752
	Be	elder brother	Brother Bro	11687 63.5	Broer	2968	Hermano Tete	22723 4	anh	3173
4th	D	daughter	Daugther	23571	Dochter Dochtertje	10378 1580	Hija	68744	con con gái	5296.5 2543
	S	son	Son	28815	Zoon Zoontje	8737 3753	Hijo	94575	con con trai	5296.5 2370
	ZyD	younger sister's daugter	Niece	326.25	Nicht Nichtje	71.75 121.75	Sobrina	746	cháu cháu họ	275.33 4
	ZyS	yougher sister's son	Nephew	339.25	Neef Neefje	177.75 76.5	Sobrino	1067	cháu cháu họ	275.33 4
	ByD	younger brother's daughter	Niece	326.25	Nicht Nichtje	71.75 121.75	Sobrina	746	cháu cháu họ	275.33 4
	ByS	younger brother's son	Nephew	339.25	Neef Neefje	177.75 76.5	Sobrino	1067	cháu cháu họ	275.33 4
	ZeD	elder sister's daughter	Niece	326.25	Nicht Nichtje	71.75 121.75	Sobrina	746	cháu cháu họ	275.33 4
	ZeS	elder sister's son	Nephew	339.25	Neef Neefje	177.75 76.5	Sobrino	1067	cháu cháu họ	275.33 4
	BeD	elder brother's daughter	Niece	326.25	Nicht Nichtje	71.75 121.75	Sobrina	746	cháu cháu họ	275.33 4
	BeS	elder brother's son	Nephew	339.25	Neef Neefje	177.75 76.5	Sobrino	1067	cháu cháu họ	275.33 4
5th	DD	daughter's daughter	Granddaughter	370	Kleindochter	87	Nieta	1304	cháu cháu ngoại	275.33 30
	DS	daughter's son	Grandson	478	Kleinzoon	117	Nieto	1685.5	cháu cháu ngoại	275.33 30
	SD	son's daughter	Granddaughter	370	Kleindochter	87	Nieta	1304	cháu cháu nội	275.33 25
	SS	son's son	Grandson	478	Kleinzoon	117	Nieto	1685.5	cháu cháu nội	275.33 25