

Emerging Languages Model

Connie de Vos, Seán Roberts & Bill Thompson

April 12, 2016

1 Acquiring & Using Speech and Signs

Individuals in the model acquire an inventory of signs and speech sounds by observing and using language in miniature/one-shot encounters with other individuals.

1.1 Exemplar Storage

We assume a simple exemplar model of signal storage: whenever a learner observes a signal x_i , it stores the signal in a memory $\mathcal{E} = \{x_i : i = 1, \dots, n\}$. Each signal x_i has a number of properties: the modality in which it was expressed, the form of the signal, and the meaning it denotes. We keep track of these properties to summarise an individual's linguistic competence. For example, the balance of signs and speech sounds in an individual's inventory is given by:

$$b = \frac{1}{n} \sum_{i=1}^n m_i \quad \text{for } i = 1, \dots, n, \quad (1)$$

where $\mathcal{M} = \{m_i : i = 1, \dots, n\}$ is a record of the modality for every sign encountered, such that $m_i = 1$ if signal x_i was manual and $m_i = 0$ if x_i was audible. Deaf individuals store only manual signals. Hearing individuals store both manual and audible signals.

1.2 Modalities in Action

Individuals' use of signals during interactions is governed by a set of simple heuristics. These heuristics express patterns of conditional usage of sign or speech as a function of the interlocutor's deaf status.

In each ‘stage’, every agent chooses another agent to interact with, in accordance with the social structure weights outlined below (they are more likely to interact with someone close to them in the social structure). The first agent is the speaker and the chosen partner is the listener. Therefore, every agent transmits a signal, but may receive many or none.

In an interaction, the speaker produces a signal which the listener stores (if possible). If the speaker is a deaf individual, she always produces a manual signal: if she already knows a sign, she selects randomly among these; otherwise, she innovates a new sign. If the speaker is a hearing individual, the modality of the signal produced depends upon the hearing status of the listener. If the listener is a hearing individual, the speaker samples randomly from all exemplars she knows, as such producing a manual sign with probability b . If the listener is deaf, she produces a manual sign, either sampled randomly from known signs or innovated if no signs are already known. Table 1 express the probability $P(m_{new} = 1)$ that the the produced signal m_{new} is in the manual modality.

Table 1: Probabilities that a produced signal is in the manual modality

	Deaf Listener	Hearing Listener
Deaf Speaker	$P(m_{new} = 1) = 1$	$P(m_{new} = 1) = 1$
Hearing Speaker	$P(m_{new} = 1) = 1$	$P(m_{new} = 1) = b$

A simulation consists of $Param_{nStages}$ stages.

1.3 Measuring communicative success

The communicative success of two given agents is measured as the similarity in the distribution of forms in the agents’ memories:

$$Comm(agents_i, agents_j) = |b_i - b_j| \quad (2)$$

2 Social structure

A world contains a fixed number of agents, controlled by $Param_{nAgents}$. Agents are assigned a random sex (stored in a vector *sex* and a set of genes (see section below). Each

agent has a hearing status (either hearing or deaf, stored in a vector *deaf*).

Agents belong to a compound, and a compound belongs to a clan. Compounds and clans are stored as vectors *clan* and *compound* of length $Param_{nAgents}$. The number of clans is fixed at $Param_{nClans}$. The maximum number of people in a compound is fixed at $Param_{MaxSizeCompounds}$.

The world is initialised by assigning an even number of agents to each clan (as possible). Agents are assigned to a compound within the clan. When the compound is full, a new compound is created.

The social structure S is a matrix of weights (dimensions $Param_{nAgents} \times Param_{nAgents}$) which determines how likely each agent is to interact with every other agent. S is initialised according to several factors, including whether the agents are in the same compound, the same clan, are of the same sex and are both deaf. Each factor is weighted by a parameter. Below, *same* is a function that returns 1 if the items are the same, and 0 if not. *both* is a function that returns 1 if both items are true.

$$\begin{aligned}
S_{i,j} = & same(clan_i, clan_j) Param_{clanWeight} \\
& + same(compound_i, compound_j) Param_{compoundWeight} \\
& + same(sex_i, sex_j) Param_{genderWeight} \\
& + both(deaf_i, deaf_j) Param_{deafCommunityWeight}
\end{aligned} \tag{3}$$

3 Marriage

A matrix M keeps track of which individuals are married. A matrix F stores the weights of likelihoods that each agent would marry every other agent.

There are some hard constraints on who can become married ($F_{i,j}$ is set to 0): Agents are not permitted to have more than $Param_{maxNumberOfMarriagePartners}$ marriage partners. Agents can only marry an agent with opposing sex. Beyond those constraints, F is set to reflect

$$\begin{aligned}
F_{i,j} = & same(clan_i, clan_j) Param_{clanWeightMarry} \\
& + same(compound_i, compound_j) Param_{compoundWeightMarry} \\
& + Comm(deaf_i, deaf_j) Param_{comSimilarityMarry}
\end{aligned} \tag{4}$$

Where $Comm(i, j)$ is a function that measures the communicative similarity of two agents.

A set of parameters fix the ideal proportion of married agents in the community. The proportion of married hearing agents is set in $Param_{proportionOfMarriedAgents}$ and the proportion of married deaf agents is set in $Param_{proportionOfMarriedAgentsDeaf}$.

Marriages occur periodically, with agent i and j being selected for marriage with a probability in proportion to $F_{i,j}$. Couples are selected, then married (updating M) until the ideal proportions are satisfied, or no more marriages are permitted. Marriage is either patrilocal (the female in the couple adopts the male's clan, compound and associated social structure) or matrilocal (the male in the couple adopts the females's clan, compound and associated social structure), depending on $Param_{MarriageLocality}$ (although only patrilocal is implemented, currently).

4 Death and Reproduction

Every 3 stages, a number of births and deaths may occur with probability $Param_{probOfBirthDeathEachStage}$.

A number of agents are added to the population, up to $Param_{maxNumberOfBirthDeathEachStage}$.

A random married couple is selected and a child agent is added to the population. The child inherits its genetics from its parents (see below). It inherits its clan, compound and social structure weights from either its mother or its father, depending on $Param_{CulturalInheritance}$ (currently only patrilineal is implemented).

After this, an equal number of agents are selected at random to be removed from the population (so the population size remains stable).

4.1 Genetics

Agents carry two genetic alleles, a 'deaf' allele and a 'hearing' allele. The deaf alleles are recessive, so an agent must have two deaf alleles to possess a deaf phenotype. Child

agents receive one randomly selected allele from each of their 2 parents.

The initial prevalence of genes are controlled by two parameters that control the proportion of recessive genotypes ($Param_{gDom}$) and the proportion of mixed genotypes ($Param_{gCarry}$).

5 Interpreting stages in years

Each agent keeps track of the number of stages they have experienced (their 'age'). The average age of agents at the end of the simulation is calculated, and equated with the average life expectancy of real people in the given situation. In the case of Kata Kolok, the average number of stages experienced is roughly 500, and the average life expectancy is roughly 70 years. That means that 1000 stages is roughly 140 years.

6 Estimating fluency

The real data on sign fluency was collected by having each person in the survey be interviewed by a deaf individual, and seeing whether they could do a series of basic communication tasks.

In the model, this is paralleled like this: - For each agent get them to have an 'interview' with a random deaf agent.

- Work out the difference in their lexicons, which returns a number representing the 'distance' between the two agents' lexicons. This is a kind of measure how hard it would be to communicate.

- Work out the mean distance and standard deviation for deaf agents being interviewed by deaf agents.

-For each hearing agent, work out how much bigger their distance is than the average deaf-deaf pair.

-Invert this difference so it's a measure of how 'close' each agent is to the mean deaf-deaf pair

-Divide the population like this:

- Non-signer: less than 1 standard deviation below the mean
- Non-Fluent: 0 to 1 standard deviation below the mean

- Fluent: 0 to 1 standard deviation above the mean
- Balanced: more than 1 standard deviation above the mean