An agent-based model of sign language persistence informed by real-world data

Katie Mudd
Vrije Universiteit Brussel, Brussels, Belgium
katie.mudd@ai.vub.ac.be

Connie de Vos | ORCID: 0000-0002-4800-4313
Tilburg University, Tilburg, the Netherlands;
Radboud University, Nijmegen, the Netherlands
C.L.G.deVos@tilburguniversity.edu

Bart de Boer
Vrije Universiteit Brussel, Brussels, Belgium
bart.de.boer@ai.vub.ac.be

Abstract

As evidence from sign languages is increasingly used to investigate the process of language emergence and evolution, it is important to understand the conditions that allow for sign languages to persist. We build on a mathematical model of sign language persistence (i.e. protection from loss) which takes into account the genetic transmission of deafness, the cultural transmission of sign language and marital patterns (Aoki & Feldman, 1991). We use agent-based modeling techniques and draw inspiration from the wealth of genetic and cultural data on the sign language Kata Kolok towards a less abstract model of sign language persistence. In a set of experiments we explore how sign language persistence is affected by language transmission types, the distribution of deaf alleles, population size and marital patterns. We highlight the value of using agent-based modeling for this type of research, which allows for the incorporation of real-world data into model development.
Keywords


1 Introduction

It has been a challenge to test theories of language emergence and evolution. Starting in the late 1980s, computer modeling and mathematical modeling have been used to test proposed theories (Smith, 2014). In turn, experimental and observational data have been used to verify the findings generated from these models. For instance, seminal modeling work by Kirby (2001) demonstrated that in the absence of biological evolution, the cultural transmission of language is sufficient in explaining the emergence of compositionality. Experimental work by Kirby, Cornish and Smith (2008) confirmed the results of the model. Despite researchers often providing experimental or naturalistic evidence to support model results, it has always been a challenge to incorporate real-world data into model development. Moving towards modeling techniques that allow us to incorporate levels of complexity needed to explain language evolution requires empirical data from a range of disciplines (Vogt & de Boer, 2010). Here, we present an agent-based model of sign language persistence, i.e. whether a sign language continues to exist, informed by real-world data from emerging sign languages.

The field of evolutionary linguistics has been increasingly interested in emerging sign languages (e.g. Senghas, Kita & Ozyurek, 2004). These are languages that emerge in a community of deaf individuals with no prior language knowledge (Meir et al., 2010). The only new instances of spoken languages are pidgins and creoles, which emerge due to contact between individuals of mutually unintelligible languages (Siegel, 2008). These spoken languages have been used as evidence of the essential features of a language (McWhorter, 1998), but they are not devoid of the influence of their root languages (Meir et al., 2010). Hence, emerging sign languages are the only example of languages which emerge in a communicative group without prior language knowledge that can be observed today. These languages have been used as a proxy for the origin of language because the first generation has no access to linguistic input at birth and the linguistic features found are largely independent of surrounding spoken languages (Meir et al., 2010). Thus, emerging sign languages can help reveal the linguistic features present in the earliest stages of cultural language evolution. It remains an open question what specific genetic and cultural con-
ditions allow for these languages to emerge and persist. The present study uses a computer simulation based on very detailed data from documented signing communities to study exactly what combination of factors contribute to sign language persistence.

The contributions of the present work to the study of language evolution include the following: first, emerging sign languages represent important case studies for language evolution. The present research explores under what genetic and cultural circumstances these languages persist. Secondly, the present work is more broadly relevant because it provides more detailed insight into gene-culture coevolution, and to what extent the simplifications made in previous, influential mathematical models (Aoki & Feldman, 1991; Feldman & Aoki, 1992) are valid. Thirdly, at the broadest level, the following work shows under which circumstances a communication system—and the gene(s) underlying it—can persist in a population when only a subset of the population learns the communication system, and how genes and culture coevolve in such a case. This may be relevant to early (biological) language evolution, in which it remains unanswered how language can persist—and genetic adaptations to language can spread—in a situation where only a small subset of the population members have the alleles necessary for language.

Depending on the social origin, emerging sign languages can be categorized as Deaf community sign languages (Woll & Ladd, 2003) or shared sign languages (Nyst, 2012). Deaf community sign languages are used by a large and dispersed group of deaf individuals, of which the vast majority are born to hearing parents (Mitchell & Karchmer, 2004). These languages emerge when deaf individuals congregate, for example in a city or a school. Here, the majority of language users are deaf, unrelated to each other and come from various backgrounds. Nicaraguan Sign Language is an example of a Deaf community sign language documented from its birth, which emerged due to contact between unrelated deaf individuals at a school in Managua (Kegl, Senghas & Coppola, 1999). In contrast, shared sign languages are shared by deaf and hearing members of communities and typically emerge in rural areas with a high incidence of hereditary deafness. These sign languages arise within a community, and the language transmission occurs within and between families (Nyst, 2012). Hence, there are a variety of genetic and cultural differences in the conditions affording sign language emergence and persistence. The precise conditions allowing for sign language emergence and persistence remain unclear and are a necessary subject of investigation for answering questions about language emergence and evolution.

Understanding the precise conditions giving rise to new sign languages is not straightforward. The number of documented emerging sign languages is

LANGUAGE DYNAMICS AND CHANGE 10 (2020) 158–187

---

Downloaded from Brill.com09/14/2023 12:25:47PM
via free access
small and not all features (e.g. genetic, linguistic, anthropological) of all emerging sign languages have been investigated. Hence, there is a small number of (often incomplete) data points which may be used to answer what conditions allow for sign languages to emerge and persist. Another methodology that has been used to understand how linguistic features of sign languages emerge and evolve is controlled experiments, for instance using the silent-gesture paradigm (e.g. Motamed et al., 2018; Schouwstra & de Swart, 2014). However, it would be impossible and unethical to conduct controlled experiments to fully grasp the contribution of the different genetic and cultural factors facilitating sign language emergence and evolution. Because of the complex relationship between genes and culture involved in sign language emergence and evolution, modeling is an appropriate methodology to tease apart the importance of the factors involved. In fact, sign languages have been the frequent subject of mathematical models of gene-culture coevolution (e.g. Aoki & Feldman, 1991; Nance & Kearsy, 2004). Here, we elaborate on one such model.

Aoki and Feldman's mathematical model (1991) addresses the conditions that allow for the persistence (i.e. protection from loss) of sign languages, as a result of the genetic transmission of deaf alleles, which are affected by marriage patterns, and the cultural transmission of a sign language. Though they explore a plethora of factors which may affect sign language persistence, their final model makes two assumptions: first, it assumes that deafness is due to a recessive phenotype, in which only individuals who acquire the allele causing deafness from both parents are deaf. Their model focuses on this type of genetic transmission of deafness, as it is the most common cause of deafness (Bitner-Glindzicz, 2002). Second, the model assumes vertical transmission (from parents to children) of the sign language from a signing parent to a deaf offspring. A key feature of models of sign language persistence is assortative mating for deafness, meaning non-random mating with regards deafness. For instance, in England and in the United States, deaf individuals have a tendency to marry other deaf individuals (Kyle & Woll, 1988); thus, in these cases there is a high proportion of assortative mating for deafness. Using the mathematical model, Aoki and Feldman (1991) conclude that the role of other types of sign language transmission, horizontal transmission (random transmission between individuals of the same generation) and oblique transmission (transmission between a child and a random individual in their parent's generation, other than the child's parents), do not significantly affect the persistence of a sign language.

A follow-up to their initial study investigated the role of grandparental transmission on sign language persistence (Feldman & Aoki, 1992). This transmission mode is particularly interesting because as deafness is modeled as a recessive trait, it may skip a generation, so sign language transmission from
a deaf grandparent to a deaf grandchild may be important for sign language persistence. To address this mathematically, Feldman and Aoki (1992) introduce a method based on triplets, consisting of two overlapping mother-father-daughter triplets, where the daughter in the first triplet is the mother in the second triplet. Hence, this method only accounts for grand-maternal transmission. From this, the authors conclude that the most important factor affecting sign language persistence is assortative mating for deafness, but that in some cases where vertical transmission is insufficient, grand-maternal transmission may contribute to sign language persistence (Feldman & Aoki, 1992). These models (Aoki & Feldman, 1991; Feldman & Aoki, 1992) outline the conditions for sign language persistence, but no empirical data from real sign languages was used to test their model.

When empirical data from a shared sign language was applied to Aoki and Feldman’s (1991) model, the sign language was incorrectly predicted not to persist. Gialluisi et al. (2013) tested Aoki and Feldman’s (1991) model with data from Al-Sayyid Bedouin Sign Language (ABSL), a shared sign language used by a small community in the Negev desert in Israel. As is the case for many shared sign languages, the ABSL community has a high incidence of hereditary deafness and all deaf individuals acquire the sign language from birth. Strikingly, in this community, marriages between deaf individuals did not occur in the early generations of this language (Kisch, 2012). As Aoki and Feldman’s (1991) model predicts that ABSL would not persist, it indicates that their model is incomplete. Gialluisi et al. (2013) emphasize that a realistic model of sign language persistence requires consideration of the social and genetic circumstances underpinning signing communities.

The approach described in this paper aims to incorporate real data from emerging sign languages into our model. In this way, we aim to create a model which accurately accounts for the persistence of shared sign languages. The approach moves away from mathematical modeling because there are certain features of sign language communities that are difficult to model mathematically. For instance, in differential equation modeling, there is no obvious way of representing certain social structures, like the family unit. Hence, in Aoki and Feldman’s (1991) model, all transmission besides vertical (from parents to children) was random, or unstructured. Our approach advocates modeling the persistence of sign languages using an agent-based model in which it is more straightforward to retain societal structure, necessary for implementing the features inspired by the real-world data from shared sign languages.

We use the outline of Aoki and Feldman’s (1991) mathematical model as a starting point for our model. Additionally, we draw inspiration from an agent-based model used to study the features that shape sign language fluency across
emerging sign language communities, informed by the shared sign language Kata Kolok (de Vos, Roberts & Thompson, 2016). The agent-based model presented here is created with the following aims:

1. to be more representative of the empirical data from shared sign languages
2. to model a finite population in order to observe the effect of community size on sign language persistence
3. to model the contribution of deaf individuals and hearing carriers of the deaf allele informed by real-world data
4. to model the family as the focal point of sign language transmission, following claims about the role of the families of deaf individuals as being key to the persistence of real-world sign languages (Marsaja, 2008).

The paper is structured as follows; in the following section (section 2), we delve into more detail on shared sign languages, with a focus on the remarkably well documented shared sign language Kata Kolok. The documentation of various genetic and cultural facets of this signing community is used as the starting point and as inspiration for the critical features of the agent-based model. In section 3, the agent-based model is described. Section 4 consists of the model results and a brief discussion, first of the baseline Kata Kolok model and then a set of four experiments exploring model parameters. This is followed by a general discussion (section 5) and a conclusion (section 6).

2 Shared sign language variation

The present research is focused on shared sign languages (as opposed to Deaf community sign languages) for the following reasons: first, because of an often small population size, shared sign language documentation has in some cases spanned entire communities (see the community sketches found in Zeshan & de Vos, 2012). Second, shared sign languages provide a unique window into the dynamics underlying gene-culture coevolution, as shared sign languages typically emerge due to a monogenic form (i.e. involving a single gene) of recessive hereditary deafness. Third, this work is part of a larger investigation into the emergence and evolution of the shared sign language Kata Kolok. The demographics of the Kata Kolok community are exceptionally well documented in comparison to other shared signing communities and this shared sign language is one of the ones that has persisted the longest (de Vos & Nyst, 2018).

Different shared sign language communities differ in many ways. Variation has been observed, for example, in the size of the signing community, the proportion of deafness, the age of the language and the context of language use.
Table 1 Examples of variation in shared sign languages

<table>
<thead>
<tr>
<th>Sign language</th>
<th>Incidence of deafness</th>
<th>Population size</th>
<th>Marriage norms</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high: 2500 (2012)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alipur Sign Language</td>
<td>low: 0.75%</td>
<td>high: 20000</td>
<td>no deaf-deaf marriages</td>
<td>Panda (2012)</td>
</tr>
<tr>
<td>Chican Sign Language</td>
<td>mid: 2.36%</td>
<td>low: 720</td>
<td>deaf-deaf and deaf-hearing marriages</td>
<td>Delgado (2012)</td>
</tr>
<tr>
<td>Kata Kolok</td>
<td>mid: 2.2%</td>
<td>mid: 2184</td>
<td>deaf-deaf and deaf-hearing marriages</td>
<td>Marsaja (2008)</td>
</tr>
</tbody>
</table>

(de Vos & Zeshan, 2012; Meir et al., 2010). Table 1 shows examples of how shared sign languages vary on several dimensions, including marriage patterns, population size and the proportion of deaf individuals in the population. Because the features of documented signing communities vary on so many levels, the present research investigates the precise combinations of conditions allowing for these languages to persist. It is worth noting that the data presented in Table 1 and those concerning the Kata Kolok community consist of a snapshot of these communities, but these communities have experienced drastic changes. For example, in the Adamorobe Sign Language community, the population has grown from 405 in 1961 to 2,500 in 2012 (Kusters, 2012).

The current project is inspired by the Kata Kolok community, which has been extremely well documented from many angles. Since the early 1990s, this community has been studied by geneticists (Winata et al., 1995), anthropologists (Hinnant, 2000), sociolinguists (Branson, Miller & Marsaja, 1999; Marsaja, 2008) and linguists (de Vos, 2012; Lutzenberger, 2017; Perniss & Zeshan, 2008). Because of the in-depth documentation of Kata Kolok from many perspectives around the early 2000s, it offers a unique window into the cultural and genetic dynamics allowing for the persistence of a shared sign language.

Kata Kolok is a shared sign language that emerged in a small village called Bengkala in Bali, Indonesia due to a high incidence of recessive hereditary deafness. Based on genealogical (Liang et al., 1998) and genetic (Winata et al., 1995) evidence, de Vos (2012) concludes that Kata Kolok is likely in its sixth generation. In the village there are 2184 residents of which 2.2% are congenitally deaf (individuals who carry two deaf alleles) and 17.6% of hearing individuals carry a deaf allele (Winata et al., 1995). Further, as reported for other shared signing
communities, this community has adapted to deafness; deaf people are highly integrated into the community and are not discriminated against, as evidenced by them being as likely to marry as hearing people, as well as deafness being present in the village's mythology (Branson et al., 1999).

A far-reaching sociolinguistic survey was conducted by Marsaja (2008) in 2000 about the use of Kata Kolok. In this survey, 83% of the population was interviewed, including 47 deaf individuals and 1770 hearing individuals. All families with deaf individuals took part. Interviews were guided by a list of 13 questions (see Appendix 4 in Marsaja, 2008), such as “where have you been this morning?”. According to this brief interview, participants were classified as fluent signers, less-fluent signers and non-signers. Based on this survey, 70% of the population has signing skills. More specifically, in Marsaja’s (2008) sample, all deaf individuals are fluent signers (47), all hearing individuals in families with a deaf member are fluent signers (78), and hearing individuals not part of a deaf family were divided into three sign fluency categories: fluent signers (449), less fluent signers (681) and non-signers (562). All deaf signers and hearing individuals in deaf families (i.e. families with a deaf member) are considered to be fluent signers of Kata Kolok. Marsaja (2008) considers families as consisting of extended family members who usually live in the same compound, such as grandparents, grandchildren, aunts, uncles, nieces and nephews. Marsaja (2008) notes the importance of deaf families in the persistence of Kata Kolok:

The use of Kata Kolok in Deaf families is the key to its overall use across the village. Both [deaf individuals] and [deaf] families use the sign language regularly at home, and as a result, they are all fluent signers. As regular and frequent users of the sign language, these people play crucial roles in the development and maintenance of Kata Kolok.

p. 101

Because of the breadth of rich empirical data collected on Kata Kolok, it is the ideal candidate to use as a starting point to help examine the conditions that may lead to the persistence of a shared sign language. The genetic (Winata et al., 1995) and sociolinguistic sources (Marsaja, 2008) from around the year 2000 provide detailed, community-wide data from the same time period, which is ideal for the current model. Using data from the documentation of the Kata Kolok community, the computational model presented in the following section attempts to build an ecologically valid representation of how shared sign languages persist.
3 Model

The following model description follows the ODD (Overview, Design concepts, Details) protocol for describing individual and agent-based models (Grimm et al., 2006, Grimm et al., 2010).1

Purpose The purpose of this model is to investigate which factors allow for the persistence of shared sign languages. We observe how several factors affect the number of signers and, in some cases, how they affect the number of deaf agents in the population. The input parameters are the number of deaf agents (d), the proportion of hearing agents carrying a deaf allele (c), the size of one generation (n) and the value for assortative marriage (m).2 In this model, persistence and transmission of deafness is determined by deaf alleles (d and c) and marriage patterns (m), while sign language transmission occurs within the family unit. It is worth noting that in the model an agent’s signing ability does not affect their fitness, as was the case in a previous model of Kata Kolok sign fluency (de Vos, Roberts & Thompson, 2016). In the current model we do not include this as a) our language component is rudimentary, with signing ability as a binary variable, and b) given what we know about the Kata Kolok community, there does not seem to be an effect of sign fluency on fitness for deaf or hearing individuals.

Entities, state variables and scales The model contains one type of entity: agents. Agents have the following properties: identity number, age, sex, genes, parents, partner and sign ability. Agents live for three generations (0, 1, 2) and get married and reproduce at age 1. Agents can transmit or acquire the sign language at any age.

Design concepts All agents belong to a family, the only collective in this model. Family units are comprised of grandparents, parents and children. Thus, agents can be a part of maximally three family units, as grandchildren, parents (if they have children) and grandparents (if they have children and grandchildren). Interaction within the model occurs based on family structure, which determines from whom an agent can learn from and transmit the sign language.

---
1 The model code and analysis documents are available at: https://doi.org/10.6084/m9.figshare.12179760.v2.
2 In the present work we refer to assortative marriage (as opposed to assortative mating, which was used by Aoki and Feldman, 1991). In assortative mating the offspring’s genotype is directly considered, while in assortative marriage we consider the probability of an individual pairing with an individual with the same phenotype (here, deaf or hearing) which in turn would have an effect on the offspring’s genotype. We wish to draw this distinction because, though in this case there is no difference between assortative mating and assortative marriage, in certain applications there could be differences.
Figure 1  Schema showing the order of processes and where parameters are used in the agent-based model. To create the initial population, the inputs are: the size of one generation ($n$), the proportion of deaf agents ($d$) and the proportion of hearing agents carrying a deaf allele ($c$). Subsequently, agents age, get married given the assortative marriage value ($m$) (determining if they have a tendency to marry a deaf or hearing agent, depending on if they are deaf or hearing), a new generation is created and agents have the chance to acquire the sign language from their family. This process continues for $g$ iterations.

to. For observing model output, the characteristics of agents are recorded (see Entities, state variables and scales).

Process overview and scheduling The simulation set-up is outlined in initialization. After the initialization phase, each time step comprises the processes in the order outlined in Table 2. For details, see Submodel sections. Figure 1 provides a schematic overview of the order of processes and parameter input. $n$ refers to the number of agents in one generation. The full population $n_{\text{total}}$ is comprised of three generations (aged 0, 1, 2) alive at once.

Initialization The model takes a list of input parameters: the proportion of deaf agents ($d$), the proportion of hearing agents carrying a deaf allele ($c$), the size of one generation ($n$), the value for assortative marriage ($m$) and how many generations the model runs for ($g$). The initial number of agents in the model is one generation ($n$), equalling a third of the total population size of a sign language community, calculated by $n_{\text{total}}/3$. The gene distribution for these agents is set according to $d$ and $c$ (see Submodel gene inheritance). In the first generation, all deaf agents are signers.

Submodel gene inheritance When an agent is born, it inherits one allele from each of its parents. From each parent, the agent inherits an allele causing deaf-
**TABLE 2** Processes, scheduling, pseudo-code and parameters

<table>
<thead>
<tr>
<th>Process</th>
<th>Pseudo-code</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>create initial generation</td>
<td>create $n$ agents</td>
<td>$n, d, c$</td>
</tr>
<tr>
<td></td>
<td>initialize according to gene inheritance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>if agent is deaf</td>
<td></td>
</tr>
<tr>
<td></td>
<td>it acquires the sign language</td>
<td></td>
</tr>
<tr>
<td>ageing</td>
<td>for each agent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>increase age by 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>if age is 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>remove agent from population</td>
<td></td>
</tr>
<tr>
<td>marriage</td>
<td>select agents with age 1</td>
<td>$m$</td>
</tr>
<tr>
<td></td>
<td>perform assortative marriage with value $m$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>perform marriage language acquisition</td>
<td></td>
</tr>
<tr>
<td>create new generation</td>
<td>repeat $n$ times</td>
<td>$n$</td>
</tr>
<tr>
<td></td>
<td>randomly choose couple</td>
<td></td>
</tr>
<tr>
<td></td>
<td>create agent with selected couple as parents</td>
<td></td>
</tr>
<tr>
<td>gene inheritance</td>
<td>for each new agent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>randomly acquire one allele from each parent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>if the agent acquires two deaf alleles</td>
<td></td>
</tr>
<tr>
<td></td>
<td>it is deaf</td>
<td></td>
</tr>
<tr>
<td></td>
<td>else</td>
<td></td>
</tr>
<tr>
<td></td>
<td>it is hearing</td>
<td></td>
</tr>
<tr>
<td>language acquisition</td>
<td>for each new agent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>find agents in same family unit</td>
<td></td>
</tr>
<tr>
<td></td>
<td>if any agent in family is deaf and if any agent knows the sign language</td>
<td></td>
</tr>
<tr>
<td></td>
<td>all agents in the family acquire the sign language</td>
<td></td>
</tr>
</tbody>
</table>
ness (a recessive allele (d)) or an allele that does not cause deafness (a dominant allele (D)). As the model pertains to recessive hereditary deafness, only agents who acquire two alleles causing deafness, homozygous recessive (dd), have a deaf phenotype. Agents who acquire only one allele causing deafness, heterozygous (dD) or (Dd), are carriers of the allele causing deafness, and, along with agents acquiring no alleles causing deafness, homozygous dominant (DD), have a hearing phenotype.

For the initialization of genes in the first generation, input parameters needed are the size of a generation \(n\), the number of deaf agents \(d\) and the proportion of hearing agents carrying a deaf allele \(c\). The number of deaf agents in the initial generation is calculated by \(n \cdot d\), and the number of hearing carriers of the deaf allele is calculated by \(n \cdot d\). The remaining agents in the population are hearing agents carrying no deaf alleles.

**Submodel language acquisition** Acquiring sign language is modeled as a binary variable; thus, agents either acquire the sign language or they do not acquire the sign language. Agents cannot unlearn the sign language once it is acquired.

Upon marriage, if any partner knows the sign language and any partner is deaf, the member of the couple who did not know the sign language acquires it.

**Submodel assortative marriage** All agents who are age 1 are eligible to get married. The assortative marriage value \(m\) specifies the number of agents seeking a partner with the same phenotype. Hence, we calculate the number of deaf agents in a deaf-deaf marriage and hearing agents in a hearing-hearing marriage. The number of deaf agents that marry a deaf agent is calculated by the number of deaf agents in a generation \(n_d = n \cdot d\), multiplied by the assortative marriage value \(m\), yielding \(n_d \cdot m\). The number of hearing agents that marry a hearing agent is calculated by finding the number of eligible hearing agents in a generation, \(n_h = n - n_d\), and multiplied by the assortative marriage value \(m\), yielding \(n_h \cdot m\). The remaining agents in a given generation, \(n - (n_d \cdot m + n_h \cdot m)\), are randomly married with respect to deafness. Marriages between siblings are prohibited.

4 **Results and discussion**

Model output consists of agent characteristics, recorded every time step. For the analysis, we use the agent characteristics deafness and signers. The number of deaf agents and signers is presented out of the total population size. All results presented here are of 1000 repetitions of the same parameter settings for
50 generations. On the figures presented, the dark line represents the median at each time step and the shaded area represents the first and third quartiles. If the error region (i.e. the shaded area representing the first and third quartiles) of one experiment overlaps with the median (i.e. the dark line) of another experiment, then there is no significant difference in result between the two.

4.1 Kata Kolok baseline model

In the Kata Kolok baseline model, the parameter values are set to the values recorded from the Kata Kolok community. The following parameters are constant throughout the simulation: the number of agents, the assortative marriage value and the sign language transmission method within the family (vertical, horizontal and grandparental transmission). First, the number of agents in the simulation is 2184, representing the population size at the time Marsaja’s (2008) study was conducted. Next, the assortative marriage value, determining if deaf people are more likely to marry a deaf person or a hearing person, is 58%, calculated based on a family tree of the Kata Kolok community in Winata et al. (1995). In addition, transmission of the sign language within the family is hypothesized to be crucial in ensuring sign language persistence (Marsaja, 2008). Other parameters are set at the start of the simulation but are not held constant throughout the simulation. These parameters are the fraction of deaf individuals which is initially 2.2% of the population and the fraction of hearing carriers of a deaf allele which is initially 17.2% of the population (Winata et al., 1995).

In our first experiment, we look at the role of sign language transmission type on sign language persistence. To do so, all other parameters influenced by the Kata Kolok community data are kept constant. In this experiment, sign language transmission is modeled as either purely vertical (VT), vertical and horizontal (VT+HT) or vertical, horizontal and grandparental (VT+HT+GT). Figure 2A shows that there is no fitness advantage to being a signer in the model, with the three models of sign language transmission yielding approximately the same number of deaf agents in the population. On the other hand, Fig. 2B shows that the persistence of a sign language is affected by sign language transmission type. Namely, adding horizontal transmission to vertical transmission appears insignificant while grandparental transmission appears to be important for sign language persistence. Strikingly, for the Kata Kolok baseline parameters, grandparental transmission has a great effect on the persistence of the sign language. As discussed by Feldman and Aoki (1992), grandparental transmission may aid in cases where recessive deafness has skipped a generation. Because assortative marriage occurs only slightly more than half of the time in this model, deaf individuals are likely to have hearing offspring.
Figure 2A shows the proportion of deaf agents over 50 generations and Figure 2B shows the proportion of signers over 50 generations. Both Figure 2A and Figure 2B show the results of three experiments of sign language transmission types: vertical transmission alone (VT), vertical and horizontal transmission (VT+HT), and vertical, horizontal and grandparental transmission (VT+HT+GT). The dark line represents the median at each time step and the shaded area represents the first and third quartiles.

Hence, due to the discontinuity between transmitting the sign language and genes causing deafness, grandparental transmission is shown to be important in sign language persistence. Marsaja (2008) defines families as more than just siblings, parents and grandparents, but considers the extended family which also comprises aunts, uncles, nieces and nephews who live together in a compound. Future work could also consider the importance of these additional family members involved in sign language transmission.

Aoki and Feldman’s (1991) mathematical model predicts the persistence of Kata Kolok by vertical transmission alone. However, as shown in Fig. 2B, the agent-based model does not. What is causing the difference in results using these two methodologies? We posit that this difference is largely a result of the less abstract representation of the Kata Kolok community in the agent-based model. With regards to marriage in the agent-based model, only female-male marriages are allowed between individuals of the same age (one third of the population) and marriages between siblings are prohibited. In small populations, this has a large effect because the pool of potential marriage partners is limited. Given the parameter settings used here, assortative marriage only presents a small fitness disadvantage for deaf agents, where 96.5% of deaf agents are married and 97% of hearing agents are married. In addition, there is a small number of deaf agents and if they are not selected to have offspring, this can have drastic effects on the persistence of the sign language. Crucially, here we show that to account for the persistence of Kata Kolok, vertical transmis-
sion alone is not sufficient, but the addition of horizontal and grandparental transmission helps to allow for the persistence of this sign language.

4.2 Parameter variation experiments
In the following sections, we present the results from the agent-based model and follow with a comparison of how the results are different from Aoki and Feldman’s (1991) mathematical model results. In this set of experiments, we investigate how sign language persistence and the number of signers in the population is affected by the spread of deaf alleles in the population, the population size and marriage patterns. For the remainder of the experiments, the family unit consists of the siblings, parents and grandparents.

4.2.1 Distribution of deaf alleles
In Aoki and Feldman’s (1991) mathematical model, the genetic basis of deafness is represented as the proportion of deaf alleles in the population. Because there is precise genetic information about the distribution of deaf alleles in the Kata Kolok community, the agent-based model considers the proportion of deaf agents (carriers of two deaf alleles) and the proportion of hearing agents carrying one deaf allele.

The proportion of deafness is varied in communities where shared sign languages have emerged (see Table 1). For Kata Kolok, the recorded incidence of deafness in the community is 0.022 (Winata et al., 1995). Keeping all other parameter inputs constant in the model, we vary the proportion of deafness to see the effect on the number of deaf individuals in the population over time, as well as on the number of signers in the population. The input parameter values investigated for the proportion of deafness \( d \) are 0.0075, 0.022, 0.025, 0.05, 0.075, 0.1, 0.11, 0.15. We investigate several values attested for the proportion of deafness in real-world signing communities: 0.0075 for Alipur Sign Language (Panda, 2012), 0.022 for Kata Kolok (Winata et al., 1995), 0.11 for Adamorobe Sign Language (Kusters, 2012), and other plausible values.

As expected, a higher proportion of recessive hereditary deafness in the population leads to a larger number of deaf agents (Fig. 3A) and a larger number of signers (Fig. 3C). As seen in Fig. 3B, regardless of the proportion of deafness in the population, the proportion of deaf agents in the population is on average above 0. When the proportion of deafness in the population is less than 0.0075, no deaf individuals are present in the initial generation, as the proportion is too low. Next, the proportion of deafness in the population influences the number of signers in the population. As seen in Fig. 3D, only when the proportion of deafness in the population is extremely low (0.0075) does the sign language not persist on average. This is different from what is observed in the
Figure 3A shows the proportion of deaf agents and Figure 3C shows the proportion of signers over 50 generations with varying levels of deafness in the population \((d)\) at generation 0. The dark line represents the median at each time step and the shaded area represents the first and third quartiles. Figure 3B shows the proportion of deaf agents and Figure 3D shows the proportion of signers at generation 50 for each experiment studying the proportion of deafness. The dots represent the median at generation 50 and the line represents the first and third quartiles.

Alipur Sign Language community, where the sign language has persisted with this attested value of deafness in the population. Hence, the agent-based model is not accounting for the persistence of this sign language.

Counting the number of deaf individuals in a community is straightforward, as deafness can be observed directly. But obtaining an estimate for the proportion of hearing carriers in a population is difficult, because of the need to collect genetic material from a large enough sample of the population. Kata Kolok is the only sign language for which we have this data: 17.2% of the hearing population carries a deaf allele (Winata et al., 1995).

With the following simulations, we ask, if the proportion of hearing carriers is altered, what is the effect on the persistence of the sign language? We
keep the Kata Kolok baseline parameters constant except for the proportion of hearing carriers of the deaf allele \( (c) \). We investigate the following values of hearing individuals who begin the simulation with one deaf allele: \( 0.0, 0.05, 0.1, 0.125, 0.15, 0.172, 0.2 \). Figure 4 shows the results of the simulations. The proportion of deaf agents in the population rises with the proportion of hearing carriers in the population (Fig. 4B). As shown in Fig. 4A and Fig. 4B, when there are no hearing carriers of the deaf allele in the population (when \( c \) is set to \( 0 \)), after only a few generations, there is a small number of deaf individuals remaining in the population. Consequently, this also affects the persistence of the sign language. Figures 4C and 4D show that when there are no hearing carriers of the deaf allele in the population, the sign language does not persist.
In this parameter exploration, we show that the proportion of hearing carriers of a deaf allele in the population is crucial to take into account when modeling sign language persistence. When the proportion of hearing carriers of the deaf allele is low, the sign language does not persist. Because the proportion of hearing carriers for Kata Kolok is a considerable proportion of the population (17.2%), this is shown to be critical for the persistence of the sign language. Interestingly, if the number of deaf individuals in the population is low, then the sign language still persists. This experiment highlights the genetic role that hearing carriers of the deaf allele play in sign language persistence. Because of the high proportion of deaf alleles in the population, even without a high proportion of deaf individuals, deaf children are likely to be born as a result of having hearing parents carrying deaf alleles. Previous models of sign language persistence assumed an equilibrium distribution (based on the Hardy-Weinberg principle) with respect to allele frequencies in the population (Aoki & Feldman, 1991; Feldman & Aoki, 1992). Because we have uniquely detailed information about the genotype distribution in the Kata Kolok community, which is not in equilibrium, we do not assume an equilibrium distribution in the population. Hence, this agent-based model moves away from an assumption of an allele and genotype frequency equilibrium distribution as it does not align with the community being modeled.

4.2.2 Population size
The population size is 2184 for the Kata Kolok baseline model, as reported in Marsaja (2008). Here, we vary the population size to see the effect on the proportion of deaf agents and on the proportion of signers in the population. Keeping the other parameter values constant, the following population sizes are considered: 500, 720, 1000, 2000, 2184, 3000, 4000, 5000, 7500, 10000 and 20000. As shown in Table 1, the smallest population size reported for a shared signing community is 720 for the Chican Sign Language community (Delgado, 2012) and the largest population size observed is 20000 for the Alipur Sign Language community (Panda, 2012).

By varying the population size, and keeping all other parameter values constant, we assess the role of the population size on the number of deaf agents and signers in the population. Figure 6 shows the results from the simulations of differing population sizes. Figure 6A, showing the proportion of deaf agents in the population, demonstrates that there are always deaf agents in the population irrespective of population size. However, Fig. 6C shows that in smaller populations, the sign language does not persist. In addition, as expected from the relationship between population size and variation, we find more variation in the number of signers when there is a small population (see Fig. 6D).
4.2.3 Assortative marriage

Based on a family tree of the Kata Kolok community in Winata et al. (1995), the assortative marriage value is 58%. In the baseline Kata Kolok model, the assortative marriage value is set to 0.58, meaning that 58% of deaf agents marry another deaf agent, and 58% of hearing agents marry another hearing agent. The remaining 42% are married randomly with respect to deafness.

Keeping the other baseline Kata Kolok model parameters constant (population size, incidence of deafness and hearing carriers), we vary the value of assortative marriage in a set of experiments with the following values: 0.0, 0.25, 0.3, 0.4, 0.5, 0.58, 0.75, 0.9, 0.95 and 1.0. As shown in Table 1, different marriage patterns have been observed across shared signing communities, including some where deaf individuals always marry hearing individuals. In this case, the assor-
The experiments demonstrate that when the assortative marriage value is low, the sign language is not predicted to persist. When the assortative marriage value is set to 0.0, deaf-deaf marriages are extremely unlikely and thus the probability that a deaf agent has a deaf offspring is less likely when marrying a hearing agent. Further, the sign language is transmitted only when there is a family with a deaf individual and an individual who knows the sign language within that family. Thus, the likelihood of having a deaf offspring decreases due to deaf-hearing marriages, which in turn affects sign language persistence (as shown in Figs. 6B and 6D). The results of the assortative marriage experiments thus reveal a similar result to that of Gialluisi et al.’s (2013) application.
of ABSL to Aoki and Feldman’s (1991) mathematical model, namely that this agent-based model does not predict the persistence of sign languages with low assortative marriage values.

The experiments on assortative marriage also reveal a counterintuitive result, namely that when the assortative marriage value is high, the proportion of deaf agents is low (Fig. 6B) and the sign language disappears from the population (Fig. 6D). When the assortative marriage value is 1.0, all deaf agents are married to another deaf agent. This creates a dynamic where the genetic transmission of deafness and the cultural transmission of the sign language occur in parallel. Because the starting signers in the model are the deaf agents of the initial generation, if only deaf-deaf marriages occur, no hearing individuals can take part in the signing community. But what is causing the number of deaf agents to decrease over time and the sign language not to persist when only deaf-deaf marriages occur? We believe that this is a result of a difference in the possibility for deaf agents to marry and hearing agents to marry. As previously discussed, this model allows only female-male marriages between agents of the same age. Additionally, sibling-sibling marriages are prohibited. Because there are few deaf agents (in comparison to the number of hearing agents) it is often the case that a deaf agent will not marry for the reasons mentioned. When assortative marriage in the population is set to 1.0, on average 79% of deaf agents are married and on average 97% of hearing agents are married. In contrast, when assortative marriage is lower, for instance for the Kata Kolok community baseline model, on average 96.5% of deaf agents are married and on average 97% of hearing agents are married. Thus, deaf agents in the model with assortative marriage at 1.0 experience a fitness disadvantage due to the features modeled which are absent in the mathematical models. We consider the influences on the choice of a marriage partner (e.g. no sibling-sibling marriages) to be a less abstract representation of emerging sign language communities.

It has been observed that marriage norms vary in different signing communities. There are cases where deaf-deaf marriages did or do not occur for various reasons. For example, in the Adamorobe signing community, deaf-deaf marriages have been prohibited by law since 1975, as deaf-deaf marriages were observed to yield deaf offspring (Kusters, 2012). Meanwhile, Kisch (2012) reports that in the ABSL community, initially no deaf-hearing marriages occurred. Due to a change in the norms of deaf education, since 2004 deaf individuals from this community have begun to marry deaf people outside the Al-Sayyid community (Kisch, 2012). While both Aoki and Feldman’s (1991) model and the present agent-based model do not predict the persistence of sign languages with no deaf-deaf marriages, the languages have persisted in reality. We are currently investigating if consanguineous marriage patterns can
account for the persistence of shared sign languages, even when deaf-deaf marriages do not occur. Considering more cultural and genetic factors from real-world communities is needed to strengthen our understanding of how sign languages persist. In mathematical modeling, retaining information about community structure is difficult. Modeling consanguineous marriage patterns requires keeping track of relatedness, which would be hard, if not impossible, using mathematical modeling.

5 General discussion

By incorporating features from real-world sign languages into the agent-based model, we further Aoki and Feldman’s (1991) investigation of the factors that allow for sign language persistence. Using a mathematical model, Aoki and Feldman (1991) found that assortative mating and vertical transmission of the sign language are the crucial components allowing for sign language persistence. In a follow-up study, Feldman and Aoki (1992) found that grandparental transmission may help in cases where vertical transmission is insufficient. In order to further explore what factors allow for sign languages to persist and to incorporate findings from naturalistic data into model development, we used an agent-based model, as many of the features of interest are difficult to investigate using mathematical modeling. In a set of experiments using the agent-based model we observed how the persistence of the sign language and the number of signers in the population varies due to changes in how the sign language is transmitted, the number of deaf individuals in the population, the number of hearing carriers of the deaf allele, the population size and the value of assortative marriage.

Using Aoki and Feldman’s (1991) mathematical model of vertical transmission, the shared sign language Kata Kolok is predicted to persist. However, when implemented in the agent-based model, with vertical transmission alone, we observe that the sign language does not persist. We believe that because the agent-based model presents more realistic constraints, the threshold for the sign language to persist is higher using the agent-based model. Thus, to account for the persistence of this sign language, vertical, horizontal and grandparental transmission of the sign language are necessary, as opposed to vertical transmission alone. In certain cases, the agent-based model does not account for sign language persistence when modeled with features of sign languages that we know exist in reality. The present model does not account for sign language persistence when there is a small population size, a small proportion of deafness and a low value for assortative marriage.
When the population size is small (under 2000 agents), with the other Kata Kolok baseline features, the sign language does not persist in the model. However, the community that uses Chican Sign Language has a population of 720 (Delgado, 2012). Hence, a small population size should not be a limiting factor in the persistence of sign languages. In the present model, we posit that the sign language not persisting is due to the small number of deaf individuals in these simulations, given the Kata Kolok default values and marriage constraints imposed by the age and gender of agents. Hence, the genetic transmission of deaf alleles and the cultural transmission of the sign language is interrupted, and due to the small number of deaf individuals, does not recover. Cultural and genetic features of these communities are obviously critical to the persistence of shared sign languages, and therefore should be studied more in depth.

When the proportion of deafness is low, the sign language was not observed to persist. For example, in the Alipur Sign Language community, the proportion of deafness was once 0.0075 (Panda, 2012), translating in the present model to five deaf agents in the initial generation. Because of marital preferences and the small number of deaf agents in the population, when the proportion of deaf agents is low the sign language does not persist. As stated, in order to understand how sign languages with such a low proportion of deafness can persist, these communities must be studied and their features must be incorporated into model development.

When the value for assortative marriage is low, the sign language does not persist in the model, parallel to the mathematical model results applied to a sign language with a low assortative marriage value (Gialluisi et al., 2013). Because the sign language is not predicted to persist in these cases, future work will aim to consider how sign language persistence is possible in communities where deaf-deaf marriages are not observed. Future work will consider the effect of consanguineous marriage patterns in the agent-based model, a marital pattern which has been observed in some shared signing communities. Though the agent-based model presented here is able to account for the persistence in many scenarios with observed features from real sign languages, it is still lacking features to account for the persistence of all shared sign languages. In order to further understand how sign languages persist in communities where deaf-deaf marriages do not occur, more genetic and cultural data is required to accurately model these communities.

Besides collecting genetic and cultural data specific to communities in order to explain their persistence, other features which have yet to be incorporated into the model could be considered, and may be necessary to explain the persistence of sign languages. For instance, the model does not take into account the high levels of endogamy reported for some shared signing communities.
or the geographical and family spread of deaf individuals in the community, as reported by Kusters (2009). It does not consider the role of a growing population size, as documented in many emerging sign language communities, such as in the Adamorobe Sign Language community (Kusters, 2012). The model could be extended to take into account language acquisition, and the differences in acquisition between native and non-native signers. In the current model, second language acquisition can occur through marriage, but the properties of the acquired sign language are not different from first language acquisition. Age-of-exposure has been shown to affect the linguistic features acquired, as evidenced by, for example, differences between native and non-native learners in phonological features of speech (Flege & Fletcher, 1992; Abrahamsson & Hyltenstam, 2009) and grammatical judgements (Johnson & Newport, 1989). Further, Mayberry, Lock and Kazmi (2002) show that early language exposure, regardless of the modality, bolsters future language acquisition. To consider the effect of these language learning factors on sign language structure would require modeling the sign language as more than a binary variable, which will be the subject of future work. Though the presented model represents a strong simplification of reality, by using agent-based modeling techniques it is possible to incorporate more realistic features in the future to consider their effect on sign language persistence.

Briefly, we would like to outline potential extensions of the current model. First, this model could be extended to address when a sign language is facing endangerment. All shared sign languages are considered endangered, often due to contact with Deaf community sign languages (de Vos & Zeshan, 2012). Lutzenberger (in preparation) outlines the threats to the vitality of shared sign languages, and more specifically developments in the Kata Kolok community affecting the persistence of Kata Kolok positively (e.g. language documentation and positive attitudes towards Kata Kolok) and negatively (e.g. tourism and increasing contact with Indonesian sign language). Further, as discussed by de Vos (2012), deaf individuals in the Kata Kolok community have begun to attend deaf boarding schools in Bali, hence they learn Indonesian sign language and show new marital preferences; now, deaf individuals often marry deaf individuals from outside of the village, who are unlikely to carry the same recessive gene causing deafness (de Vos, 2012). A similar trend has been observed in the case of ABSL (Kisch, 2012). Considering this in the model would require modifications to the genetic component because, in contrast to the majority of shared signing communities, in Deaf community sign language communities deafness is not always due to the same genetic cause. Because of the cultural and genetic differences underpinning the persistence of Deaf community sign languages and shared sign languages (for an overview see Meir et al., 2010), we do not foresee
this model immediately being extended to account for the persistence of Deaf community sign languages.

Second, the model may be extended to observe the role of genetic and cultural factors on linguistic properties of sign languages. For the case of Kata Kolok, recent linguistic studies reveal typologically unique patterns in negation (Lutzenberger, 2017) and sign-spatial mapping (de Vos, 2012), and additionally how sociolinguistic factors govern lexical variation (Mudd et al., to appear). Across emerging sign languages, linguistic properties have been documented in great detail (Meir et al., 2010; Zeshan & de Vos, 2012; de Vos & Pfau, 2015), but how factors like population size or marriage patterns influence linguistic properties is yet to be investigated computationally for these languages. Recent exceptions include an agent-based model by Thompson, Raviv & Kirby (2019) investigating lexical variation and population size, and an investigation of the emergence of the lexicon in homesign compared to Nicaraguan Sign Language (Richie, Yang & Coppola, 2014). In addition, there exists a much larger body of work investigating how social structure or transmission affects language structure from lab experiments (e.g. Raviv, Meyer & Lev-Ari, 2019) and from computational modeling (e.g. Gong, 2010; Gong & Shuai, 2016; Cuskley et al., 2017). Of particular relevance to the current model is a study by Gong and Shuai (2016) demonstrating the importance of grandparental transmission on maintaining the mutual understanding of linguistic forms across generations.

The emergence and persistence of sign languages provides a unique window into how the interaction between genes (alleles causing deafness) and culture (assortative marriage and language transmission) can give rise to new languages. As discussed by Levinson and Dediu (2013), the story of shared sign languages highlights that a genetic mutation or bias may propel language evolution in a new direction, even if it only affects a small proportion of the population, ultimately giving rise to a feedback loop between genes and culture. More work is needed to tease apart how various genetic and cultural factors affect language. This type of research highlights the value of creating models inspired by empirical data, and presents a fruitful direction in language evolution research. By using empirical data to inform computational models, it is possible to ask complex questions that are ecologically valid.

6 Conclusion

To begin to understand the effect of various factors relating to sign language persistence, we built an agent-based model simulating the genetic transmis-
ision of deafness and the cultural transmission of sign languages. To investigate this genetic and cultural transmission, we informed our model with features from Kata Kolok, an exceptionally well-documented shared sign language, and in a set of experiments, cultural and genetic factors were varied to understand their role on sign language persistence.

In contrast with a previous model investigating these factors, we made several additions and changes to further this exploration. Aoki and Feldman's (1991) model stresses the importance of vertical transmission and assortative mating on deafness state, in an infinite population with a given proportion of deaf alleles. Following this, Feldman and Aoki (1992) show that grandparental transmission may aid in cases where vertical transmission is insufficient for sign language persistence. The agent-based model presented here is of a finite population in which sign language is transmitted in the family, with deaf alleles determining deafness and the proportion of hearing carriers of a deaf allele. In a set of experiments, we found that sign language persistence is affected by the proportion of deafness in the population, the proportion of hearing carriers of a deaf allele, the population size, assortative marriage for deafness and by the method of sign language transmission (i.e. vertical and horizontal). More work is required to account for the persistence of all shared sign languages.

All in all, this parameter exploration of sign language persistence highlights the value of using real data to inform the choices made in modeling. The better informed our models, the more valuable they will be in explaining real-world phenomena.

Acknowledgments

We would like express our gratitude to three reviewers for constructive comments on a version of this manuscript. Special thanks to Bill Thompson, Sean Roberts and Yannick Jadoul for providing information about a related project modeling Kata Kolok sign fluency across the community (de Vos et al., 2016) and to Yannick Jadoul and Marnix van Soom for help with coding and fruitful comments throughout the project. A big thanks to Hannah Lutzenberger for sharing her insights into the Kata Kolok community which helped in making modeling decisions. This research was supported by the FWO-NWO grant “The emergence of phonology within six generations” awarded to Bart de Boer, Paula Fikkert and Connie de Vos.
References


