AN AGENT-BASED MODEL PROPOSAL ON THE EMERGENCE OF DOMINANT WORD ORDER IN WORLD LANGUAGES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

BURCU ÇINAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF COGNITIVE SCIENCES

JANUARY 2023

AN AGENT-BASED MODEL PROPOSAL ON THE EMERGENCE OF DOMINANT WORD ORDER IN WORLD LANGUAGES

submitted by **BURCU ÇINAR** in partial fulfillment of the requirements for the degree of **Master of Science in Cognitive Sciences Department, Middle East Technical University** by,

Prof. Dr. Banu Günel Kılıç
Dean, **Graduate School of Informatics**

Dr. Ceyhan Temürcü
Head of Department, **Cognitive Sciences**

Prof. Dr. Hüseyin Cem Bozşahin
Supervisor, **Cognitive Sciences Deparment, Graduate School of Informatics**

Dr. Rabia Ergin
Co-supervisor,

Examining Committee Members:

Prof. Dr. Committee Member 1
Department, School

Assoc. Prof. Dr. Committee Member 2
Department, School

Assoc. Prof. Dr. Committee Member 3
Department, School

Assist. Prof. Dr. Committee Member 4
Department, School

Assist. Prof. Dr. Committee Member 5
Department, School

Date: 28.08.2019

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Burcu Çınar

Signature :

ABSTRACT

AN AGENT-BASED MODEL PROPOSAL ON THE EMERGENCE OF DOMINANT WORD ORDER IN WORLD LANGUAGES

Çınar, Burcu

M.S., Department of Cognitive Sciences

Supervisor: Prof. Dr. Hüseyin Cem Bozşahin

Co-Supervisor: Dr. Rabia Ergin

January 2023, 42 pages

In this study, we propose a computational model to explain the word order distributions used by today's languages. This model combines the preferences for word orders results of small deaf communities that had to produce new languages without traditional language inputs and other up-to-date experimental studies with the concept of iterated learning without real data. With iterated learning the model reflects the process through of a person picks up new behaviors after being exposed to those others. The model tries to explain how we transfer our word order preferences to our environment. It simulates the biological reproduction and development of communities while also incorporating our biases. It also examines the effects of the natural pressures of education and language on generations. Furthermore, it presents and tries to explain the results of possible scenarios with different parameters like community size, distribution of biases, communication network types and pressure effects.

Keywords: dominant word order, emergence of word order, communicational networks, agent-based model, iterated learning, frequency learning

ÖZ

DÜNYA DİLLERİNDE BASKIN KELİME DİZİLİMİNİN ORTAYA ÇIKIŞI ÜZERİNE AJAN TABANLI BİR MODEL ÖNERİSİ

Çınar, Burcu

Yüksek Lisans, Bilişsel Bilimler Bölümü

Tez Yöneticisi: Prof. Dr. Hüseyin Cem Bozşahin

Ortak Tez Yöneticisi: Dr. Rabia Ergin

Ocak 2023, 42 sayfa

Bu çalışmada, günümüz dillerinde kullanılan kelime sırası dağılımlarını açıklamak için hesaplamalı bir model öneriyoruz. Bu model, geleneksel dil girdileri olmadan yeni diller üretmek zorunda kalan küçük sağır toplulukların kullandıkları kelime sıralamalarının sonuçlarını ve diğer güncel deneysel çalışmaları, gerçek veriler olmadan yinelemeli öğrenme yöntemiyle birleştirir. Yinelemeli öğrenme ile model, bir kişinin diğer kişilerin davranışlarına maruz kaldıktan sonra yeni davranışlar edinmesi sürecini yansıtır. Model, cümle üretirken kullandığımız kelime sıralarını çevremize ve gelecek kuşaklara nasıl aktardığımızı açıklamaya çalışır. Önyargılarımızı da dahil ederken, toplulukların biyolojik üremesini ve gelişimini simüle eder. Ayrıca eğitim ve dilin doğal baskılarının nesiller üzerindeki etkilerini de inceler. Ek olarak olası senaryoların sonuçlarını topluluk büyüklüğü, önyargıların dağılımı, iletişim ağı türleri ve baskı etkileri gibi farklı parametrelerle sunar ve açıklamaya çalışır.

Anahtar Kelimeler: baskın kelime dizilimi, kelime diziliminin ortaya çıkışı, iletişim yapıları, ajan-tabanlı model, tekrarlamalı model, frekans öğrenme

Dedication here

ACKNOWLEDGMENTS

Acknowledgments here

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

S	Subject
O	Object
V	Verb
ABSL	Al-Sayyid Bedouin Sign Language
CTSL	Central Taurus Sign Language
ISL	Israeli Sign Language
NSL	Nicaraguan Sign Language
L1	First Language
L2	Second Language
EM	Expectation-Maximization

CHAPTER 1

INTRODUCTION

Most of today's sentences are consciously or unconsciously formed in a certain word order. Many languages are sensitive to word orders, and a canonical word order is important as it includes most of the meaning of the sentence spoken as well as different orderings of it can lead to other meanings. For example, the knowledge of who did what to whom can be a knowledge learned by ordering the words appropriately. There is a huge difference in 'man kisses woman" vs. "woman kisses man" if you are using English. The canonical word order of a language is defined by the order of the three main components of basic transitive sentences: subject (S), verb (V), and object (O). Subject is the one who does the action, verb is the action to be done, and the object is the one who is affected by the action. Logically, we can order these three items in 6 different ways: SOV, SVO, VSO, VOS, OVS, OSV. For example, while the dominant word order for Turkish is SOV [2] [3], for English it is SVO. In addition to these, there are also languages that do not conform to a certain basic element sequence. We can call these free word order languages. But these six possible word orderings are not evenly distributed amongst the world languages. According to [1] 41% of languages use SOV as their canonical word order, like Turkish, Korean, Persian. 35% of the languages use SVO as their canonical word order, like English, Chinese, French. 7% of the languages use VSO, 2% of the languages use VOS, 0.8% of the languages use OVS, again, 0.3% of the languages use OSV as their canonical word order. Also, 13.7% of the languages use free word order to communicate as given in world map in Figure 1. In these languages there is also a dominant word order, but other orderings are also possible to convey a true meaning while communicating. True meaning is the meaning that the speaker intends to convey to the listener. Both the speaker and the listener must derive the same meaning from the sentence.

So the overall distribution looks like this: (SOV, SVO) > free order > VSO > (VOS, OVS) > OSV. However, where does this canonical orders and the agreement in a certain order came from is still a matter of curiosity. To explore this, we need to go back a very long time, to our ancestors which is impossible with today's facility. Although experimental studies on the subject give various ideas, a holistic approach is not satisfactory due to the comprehensiveness of the research.

As mentioned before, it is not possible to observe the exact stages of the languages spoken in the world from their emergence to the present day. Therefore, in the researches on the subject, mostly the findings obtained from the young emerged sign languages (village sign languages like ABSL, CTSL and NSL) or the silent gestures (or pantomimes) of the speakers were used. SOV and SVO are also predominant word orders in sign languages according to [4]. This comparative study of 42 sign languages suggests that cognitive/communicative biases are involved in determining the dominant order in a language. These studies mostly done on basic sentences, complex sentences need more deep researches and experiments.

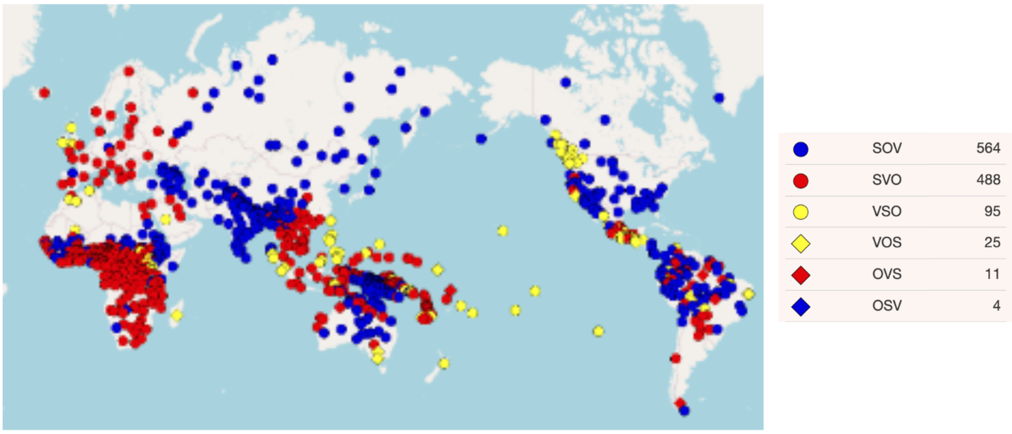


Figure 1: Dominant word order distribution in the world, taken from [1]

In this study, we also cover basic transitive sentences with three components (subject, object and verb). These transitive sentences also split into two kinds of sentences: reversible and irreversible transitive sentences. In reversible sentences it is semantically logical and possible to change the places of object and subject in the sentence, i.e. "the girl pushes the boy" could be "the boy pushes the girl" depends on the situation. On the other hand, irreversible transitive sentences does not allow us to change the places of subject and object semantically, i.e. "the girl cuts the paper" cannot turn into "the paper cuts the girl".

One of the explanations from [5] is that the observed word order frequencies may be the result of genetically encoded biases towards certain orders, as part of the universal grammar hypothesis. This could be a possibility, some word orders may be easier for us from birth. [6] suggests that the prevalence of SOV order across the world's languages may arise in part because SOV order is most compatible with how we conceptually represent transitive events. Also, [7] proposes that all living languages today descend from a single common ancestor, a proto-language uses the SOV word order. Thereof, it is possible that SOV and SVO have a special cognitive place.

[8] tries to explain the current frequencies of languages with three principles: "theme-first-principle", "verb-object-bonding", and the "animate-first principle". The frequencies are proportional to the number of principles they realised. According to [9] these three principles are realised in SOV and SVO, two are realised in VSO, one in VOS and OVS and none in OSV. It is also seen that the animacy level of subject and object also affected how we structure the sentences in many studies [4][10][6][11] [12][13], which will be explained in the literature review section in detail.

Similarly, this study is shaped using iterated learning, which tries to explain the process of a behaviour learning of an individual from other individuals' behaviour [14]. This process is also applicable for explaining the language evolution and how and why language is structured the way it is now.

1.1 Research Questions

The current studies propose some explanations for different questions, like why some orders more preferred over others, which are mentioned before, how network communication frequency and net-

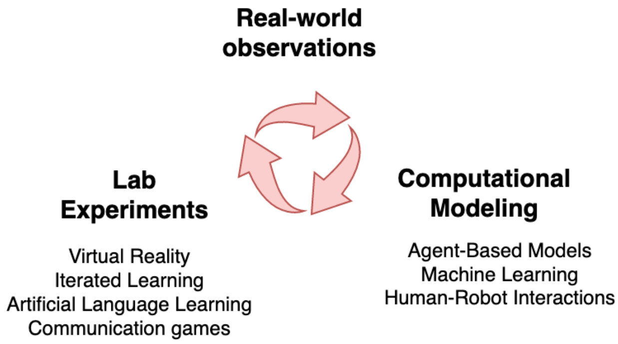


Figure 2: A cycle that contribute the language studies

work type help to make a language more systematic [15][16], how different generations behave while transferring the language by iterated learning [14], and so on.

There are several questions this study seeks to answer. These questions also serve as a source for the creation and testing of the model. First, previous studies suggest that SOV is cognitively the most basic word order to start a language. But what if this is not the case? If we start our model with another word order bias, will the word order converge to one of the most frequent orders today, SOV or SVO? Second, how do the community's size and different network structures affect the dominant word order emergence and evolution speed? We know that not all people are the same. We don't learn the same way or live the same way. Therefore, there should be an effect of if. We want to observe where the margin of error is different and how these differences will affect the convergence of the word orders. So, what will it change if many/few people change their language? And what does it tell us about real life? Will word orders remain fixed without language changers?

1.2 Contributions of the Study

This study was carried out to fill a few gaps in the literature. Although there are experimental studies and explanations on the subject, human life and opportunities do not allow a holistic study. For this reason, computers and models can enable us to combine different existing views. This study mainly focused on a computational model that connects the current explanations to strengthen our valid understandings and/or give another perspective, and of course expects to benefit the literature. We have adopted the cycle given in 2, real-word observations and lab experiments give us an insight to create a computational model.

The model produced in this study allows us to see how different and many effects affect the today's dominant word order distributions. It tries to be a playground for many people to try, observe and understand. And the good thing is, they don't need to go back or forth a century or so. This model that has not been created in the literature before and has a deficiency. Researchers can try different network models to communicate, different frequencies of communication, different sizes of population, different language types, different starting bias and different numbers of generations to predict. This model will give us a holistic perspective of what has been told in the literature before.

1.3 Study Limitations

First, this is a study based on inferences made upon the behavior of users of languages that have emerged in the near future. It is difficult to generalize to the whole world and to all times. This study captures the current understandings. Also, the study has built upon experimental data and the way of making sense of this data that there is no holistic information due to the limitations of experiments and setups. There is no actual data to feed the model. Hence, not all sentence types (like complex ones) could not be included we only cover reversible and irreversible transitive sentences.

1.4 Organisation of the Thesis

This master's thesis shows a model that can help us to understand how the word orders of today's languages distribute, while connecting different arguments from different studies. The organisation of the thesis is as follows: This first chapter explains the overall topic and general views about the topic, while presenting the research questions, contributions of this study to the literature and giving the limitations of the study.

Chapter 2 represents literature review on the topic. It will give more in-depth information on how the other previous studies contribute and suggests ideas to our study.

Chapter 3 will explain the overall model. There are different steps and stages in the model, all will be covered in this section.

Chapter 4 provides the different experiments setups on the model. And the results of these experiments will be given. Also, it will include explanations to understand how the model works and what the results tell us.

Chapter 5 is where the discussion will be done and represented, and the links and differences with existing studies also will be presented.

Chapter 6, the last chapter, gives the conclusion of this master's thesis and possible future work will be presented.

Last but not the least, references follows up all the chapters.

CHAPTER 2

BACKGROUND INFORMATION

In this chapter, related studies are given in detail. Since our study tries to make a model that connects different parameters of emergence in word orders in languages and transmission them to the generations in different circumstances, we need to understand the actual human behaviour. To do that, current literature needs to be examined. There are different studies done on the topic. They will be given to see what has been changed and how we can connect them with our model. This section is divided into three sub-sections in terms of their contribution to the model. The first sub-section tries to give the previous studies on the relationship and implications of nonverbal communication (gestures/pantomimes), newly emerged sign languages, and artificial language experiments were done in labs with word ordering behaviour of humans. The second sub-section examines the studies on iterative learning and reflects the effect of language on the transfer of generations. In the third sub-section, studies on the effect of the size of the community and the network structures on the systematicity of the language will be discussed.

2.1 Gesture Experiments, Emerged Sign Languages, and Dominant Word Orders

We can first understand how gestures help us to convey how humans learn languages. [17] supports that we all learn languages that we are exposed to after we are born. But if we are not able to acquire language as we supposed to be, what if we cannot hear? Human beings need to communicate, so, they will try alternative ways, like gestures. Furthermore, what if deaf people are the first ones in the population? They need to be the bridge between hearing ones and future deaf generations. Examining the word-order patterns in communication systems developed in laboratories is a popular method. Like in the studies, [18], [10] which wanted to explore cognitive biases for word order preferences, silent gesture, a practice in which adult hearing individuals transmit information exclusively using their hands and no voice, has been utilized.

Language learners are initially attracted to a language's canonical order (for L1 [19]; for L2 [20], [21], [22], [23]), because they are one of the sources to understand who (subject) did something to whom (object). Throughout the years, to comprehend the appeal of word order, there has been an increase in interest in sign languages while understanding language emergence in general. Emerging sign languages allow us to discover language emergence in real-time, which goes well beyond the possibilities of spoken language research. Equally valuable, the gestures of spoken people also help us to understand the cognitive effects of communication.

Most of the works consist of picture/video setups and experiments. Like [24], they carried out four experiments, two of them were with English (a subject-initial language) speaking children and the other two were with Fijian (a subject-final language speaking children. Children tried to learn a miniature artificial language based on names for two horses and two carts either ordering them in horse-cart or cart-horse. Then, they examined the influence of animacy with toy graders and toy boulders.

They showed that grader+boulder sentences are easier to learn than the reverse. So it is said that the agency of an item makes a representation easy to acquire and they distinct agent/patient property then the animate/inanimate property. Also, the results were the same in both English speaking group and the Fijian-speaking group. Hence, there is no effect of first acquired language's word order, this may be a cognitive bias for putting agents (subjects) before patients (objects).

[6] conveyed a cross-linguistic study. They wanted to see whether the language we speak is influential on how we communicate non-verbally. They asked English, Turkish, Spanish, and Chinese participants, where all their native languages have different dominant word orders. Participants carried out two non-verbal tasks: the first is describing an event without speech (using gestures only) and the second is reconstructing an event using pictures. Their results show that speakers' word order for their native languages had no effect on their non-verbal communication. They also found out that all of the speakers of four languages adopted the same ordering behaviour, agent-patient-action (or SOV). This is also the most used word order in many languages, as in newly developing gestural languages, and experimental artificial languages, too. The results showed that humans impose a natural order on occurrences when they are described languages non-verbally and when a new language is created.

[25] made two gesture-production experiments and one gesture comprehension experiment on native and normally hearing Italian and Turkish participants, where the languages have different word order preferences (Italian - SVO, Turkish - SOV). In the first experiment, they tested whether participants used their gestures by following their native language's structural regularities. In the second experiment, they used stimuli to get improvised gestures. They wanted to see if there is a shred of evidence for phrase structure in their improvised gestures. In the third experiment, they investigated whether the preferences found in gesture production also appear in gesture comprehension. And finally, they studied the preferences of phrase structure by assessing participants' order preferences for flat prosodic sequences of words in their native language. Results showed that, in the direct connection between the sensory-motor and the conceptual systems, SOV is the preferred order; the SVO order is favored by the computational system of grammar.

Observing different word order preferences in reversible (where both agent and patient are animate) and irreversible sentences (where the agent is animate and inanimate patient), [26] claimed that this state depends on the communicative pressures of semantic factors. SOV is used for irreversible events since it is clear who is doing what to whom (i.e. "MAN-BREAD-CUT"). On the other hand, in reversible events (i.e. "WOMAN-MAN-KISS"), gesturers preferred SVO over SOV by putting the action (verb) in the middle of two animated participants in the event, maybe to reduce any kind of misunderstanding between the agent and patient roles.

SOV is the most used word order in world languages. But where did SVO come from and why? [27] treated the subject differently and wanted to test whether SVO emerged for a reason, i.e. maybe it is not well suited for describing reversible events (i.e. GIRL KISS BOY). If we prefer to use SOV in reversible events, that sentence will be "GIRL BOY KISS", but this will only be understood if we know SOV is the agreed word order to be used. If someone uses OSV as their canonical word order, that sentence (now it is BOY GIRL KISS) will be understood as it was the boy who is kissing the girl by the ones who prefer to use SOV as their canonical word order (see Figure 3). Another reason is could be the pressures of efficiency (it is logical to mention agents before patients as a principle, which will rule out many other alternative word orders). They tested speakers of both English (SVO) and Turkish (SOV), by asking them to use pantomime to describe some reversible and irreversible events. Additionally, they gave some participants the task of teaching the experimenter the form of gestures while being consistent about them. These restrictions caused SVO to appear in both Turkish (SOV) and English (SVO) speaking individuals. Their results showed that being efficient, putting subjects before objects in the sentences, and avoiding SOV order for reversible events are the three requirements that SVO permits language users to achieve, and this is said to be at least part of the reason why SVO arises in the world's languages.

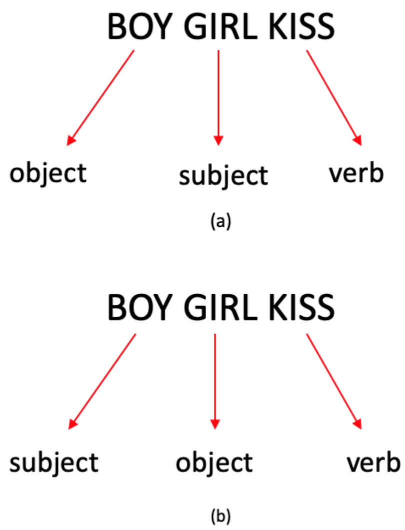


Figure 3: Same order, different meaning

Even if the words are in the same order, the meanings are different: (a) if OSV is the canonical word order and it is the girl who kisses the boy. (b) if SOV is the canonical word order and it is the boy who kisses the girl.

Another work [11] also used gesture production to see what causes preferences of SOV than SVO, or vice versa. To do that, they focused on the role of the verb in the sentences. They also mentioned extensional (i.e. throw) and intensional (i.e. think) verbs. Their results for events with extentional events showed that SOV is mostly the preferred order, which is consistent with the previous works ([6], [25], [26]). But, events with intentional verbs led to the SVO order instead. They concluded that the meaning of the verb is critical while ordering words in emerging language systems. These results also supported the language evolutionary point, which implies that semantics underlies the early formation of syntactic rules of the language.

[28] argued that the emergence of SVO is possible with exposure to a shared lexicon, which makes it possible to liberate adequate cognitive resources to use syntax. Eventually, they argued that SVO is a more efficient word order to express syntactic relationships. They gave Italian (SVO) and Persian (SOV) speakers a set of gestures to learn, and then they asked them to narrate basic events to verify

their hypothesis. It was revealed that after enough time, when both groups developed a consistent gesture repertoire, there emerged a coherent usage of SVO. This study is particularly important for the Persian-speaking participants since they switched their native language word order and started to use SVO once they were exposed to a new lexicon and become confident about it.

One piece of information to support the existence of a cognitive bias for SOV might be that the SOV is still dominant today. For example, a cross-linguistic study [29] argued that SOV is the oldest word order used, and different needs and conflicts made other orders emerge. The geographical distribution of word orders around the world also seems to support this opinion. On every continent but Africa, SOV is the dominant word order. SVO, on the other hand, is mostly limited to Africa and Eurasia, with little to no presence in the Americas and Austronesia.

On the side of sign languages, there are different orientations. Naturally emerged sign languages are said to have little or no influence on existing languages. Previous studies on these systems show that there is a great amount of variation in word order preferences in the initial stages of young sign languages. An emerged sign language in Colombia lacks a regular word order preference [30]. Likewise, in NSL, [31], it is reported that there are different word order preferences in different age groups with no clear tendency towards a certain word order.

Like [6], which said that the native language has little/no effect on gestural word order preferences, ABSL (used in a bigger community in Israel) is reported to have a different tendency of word order (SOV) in spite of their regional languages (Hebrew and Arabic - both SVO) [32], that is pointing out the word order preferences does not always influence by the local speaking language. The preference of word orders of ABSL users shows a systematic variance (OSV-SOV in different event structures) according to more recent research by [13].

According to [33], ISL has shown an SOV tendency as consistent word order in younger generations. In another study [34], it was reported that in an American and Chinese home sign system, the 'subject' in an event is frequently ignored and users preferred the OV order. This again shows that these emerging languages are independent of the languages spoken in the surrounding area.

For the young emerged languages, most of them lack clear word order tendencies in either reversible or irreversible events. There are studies that show clear preferences generally after the first generation. [35] and [36] show that CTSL (used in a small community in Turkey) has no plain word order preference in the first generation, but with the second generation SOV seems to be the dominant order for irreversible events and OSV is for the reversible events, and in the third generation, the previous preferences become more common even OSV is the least prevalent word order. It has been reported that SVO, the most common after SOV, has not yet emerged in this population. Correspondingly, similar results were obtained in ABSL, according to [13]. Instead of SVO, OSV has emerged in this community, too. This is an intriguing result because SVO is preferred over OSV in today's languages. These results make us ask similar questions again. Is SOV the simplest sort or do we have a bias for it? What caused the emergence of SVO?

The issue of who or what comes first when forming a sentence has also been the subject of many studies. We know that in world languages, subject-predicate before the verb is fire (S>V). To explain this, [37] refers to the "agent-first" principle. This principle says we have a bias that causes us to put the subject before the verb while forming a sentence. This principle is supported by late second language learners [38] and pidgin languages [39], who they prone to use subject before a verb to distinguish between roles in a sentence with an agent and a patient. It seems to be defending a cognitive tendency that causes us to use the subject of the sentence before the object naturally. On the other hand, there are also findings that contradict these views. These tendencies may not be observed in newly emerged sign language systems, like dominant OSV order in CTSL and ABSL. This may undermine the validity of the "agent-first" principle. The principle alone seems to be not enough, other pressures should have also taken place in these systems. In this regard, [13] tried to explain this situation with the animacy property of the characters in the event. It argued that regardless of whether it is a subject or an object, the human or the character most similar to a human being is put forward. But these principles, again, due to limited test cases, cannot be fully defended as true or false, as many events were left out. For example, sentences with inanimate subjects (with animate/inanimate objects) in new languages have not yet been studied. Likewise, situations where animal subjects and human objects are present, or where human subjects act on animal objects, were not studied yet.

The approaches to studying word order mentioned so far can be summarized as follows:

- a consistent word order may be absent in the very initial stages of a young natural system,
- there are cognitive biases (independent of surrounding/native languages) and communicative pressures shaping word order preferences, and
- considering the prevalence of SOV and SVO, there is so far robust evidence for the S>O pattern, displaying a bias for the agent preceding the patient.

On the other hand, [40] claimed in experimental studies (like gestures, and pantomimes) that people prefer an unnatural way. In such cases, although people prefer or have a chance to choose a certain word order to describe certain kinds of events, it is stated that people actually prefer to use the same order during actual natural speech. So, regardless of the types of events, we prefer to use the same order while talking. Taking all these into account, they argued that repeated use will eventually lead to regularity rather than naturalness.

2.2 Learning Models and Iterated Learning

There are different proposals on how learning occurs as a model. Models that involve learning the language from various aspects, including word learning, and frequency learning, have been proposed. However, a model of word orders does not exist so far. This study is essentially a model proposal to fill this gap. Yet, for our model, we are influenced by a few of the studies. Since there is no consensus on how language is learned, only the insights and studies we used for our model will be given here.

[41] proposed a framework based on the principles of Bayesian inference to capture human concept learning from examples. In the study, it was supported that a given collection of instances can provide significantly more information about the concept to the Bayesian learner, who can then use this knowledge to rationally estimate the likelihood that any new item will also be an instance of the concept. The proposed system consists of 3 components: the first is a prior probability distribution across a set of potential concepts in the hypothesis space; the second is the likelihood function to compute the probability of each hypothesis with the help of the provided set of utterances, finally, the principle for which has the learner calculate the likelihood of applying a concept to new objects by averaging all

hypotheses' predictions and weighting them according to their posterior probabilities. This study is important because it shows how it is possible for people to learn and generalize concepts from just one or a few examples which also support actual human behaviour. Although it may not be compatible fully, we will use the main components of this work. We will map concept learning to word order preferences.

One model that we considered is the iterated learning model where language is treated as cultural knowledge by [42]. Iterated learning is a way of transmitting information from one individual to the other. An individual's learning is provided by the other individual's output of their learning. At each utterance, each learner sees data, forms a hypothesis, then, produces the data for the next learner (see Figure 4). The aim of this process is simply to investigate the cultural evolution of linguistic structure.

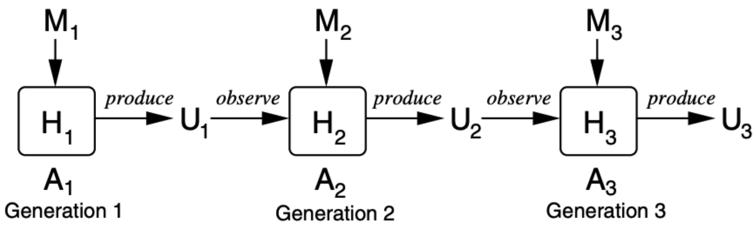


Figure 4: Iterated learning model

This model is also related to Bayesian inference. During each learning process, each learner updates their beliefs based on the rational procedure. For this study, for example, the preferred order to describe a reversible/irreversible event will be updated for events in related classes.

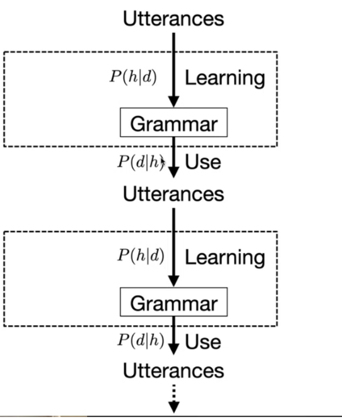


Figure 5: Modeling iterated learning based on the hypothesis

Another study [43] combined the ideas of iterated learning and Bayesian inference in an agent-based model. They created a model where learners compute a posterior distribution by combining prior knowledge of a language with the provided data. They showed that the prior and the amount of data they provided affected the whole model when they applied iterated learning. They argued that the problem is choosing the maximum probable posterior. Then, they suggested a variant of the EM algorithm.

They showed how iterated learning fits into language acquisition and create a connection between spoken and written languages; this suggests that information transferred through iterated learning will eventually come to represent the learners' ideas.

For learning the meanings of words, [44] proposed a Bayesian framework. Their theory simply described how learners may draw valid conclusions about a word's referents from a few examples by using inductive reasoning that combines prior knowledge of the word's meaning with the statistical structure of the observed cases using Bayesian inference. They included how adults and children differ when learning. They have done three experiments on adults and children. In three experiments, both adults and kids acquired referents for item categories at different levels, testing the predictions of the Bayesian approach. The Bayesian explanation outperformed competing theories in terms of both quantitative model fits and the ability to explain major qualitative events. Then they showed how Bayesian word learning models were more applicable. Of course, we must not forget that this model also only deals with simplified versions of real challenges as most of the models presented.

[45] wanted to test the iterated learning paradigm with an artificial language learning experiment with humans in the laboratory. This study mainly supported how the previous mathematical and computational models of iterated learning can really be applied in the actual world. Iterated learning is said to be the explanation of the cultural transmission of the language with particular and natural constraints of language. Their experiments showed that iterated learning is a vigorous model for understanding the behavior of humans in transmitting language to generations. Another study [14] also showed how iterated learning helps us to understand the origins of a language. They reviewed previous studies of iterated learning. In this study, they used computational simulations of agents; mathematical modeling; and laboratory experiments to support and show the power of iterated learning.

Meanwhile, some linguists emphasize the importance of frequency of occurrence in language learning. Our model, mainly, adopts the idea of frequency learning. So, studies should be included here. Usage-based theories hold that the acquisition of language is exemplar-based. It is said that the frequency with which we encounter information provides learning with what we associate with it.

The article [46] tried to demonstrate how the frequency of input has a profound impact on language acquisition. The way of humans form sentences and syntax are affected by the frequency provided. It is the result of the learning of many constructs and the frequency-biased abstraction of their internal regularities. The categories and patterns that make up linguistic regularities are said to be formed through given inputs and experience. They advocated that language change is greatly explained by frequency. It is finally concluded that the sensitivity of learners to the frequency in many linguistic areas has effects on how implicit and explicit learning theories interact.

Another study [47] wanted to examine the connection between children's language acquisition and the frequency of morphological, lexical, and syntactic forms as their input with some brief counter-arguments towards the effects of frequency in language learning. They have reported that they explored several variables that interact with the frequency effect since children do not have a simple way of mapping the input string directly. Their study was done in English. They tested this frequency-based explanation with this empirical work they have done. In the end, they discussed a relationship between the relative frequency of forms in the input of infants and errors, including morphological errors, optional infinitive errors, and accusative-for-nominative errors.

A study [48] wanted to prove the role that the frequency effect plays in language learning. They wanted to prove that the frequency effect must be considered in children's language learning. They studied simple syntactic constructions along with other respects of children's language acquisition and then presented theories. In the study, it was reported that the high frequency provides early acquisition, brings systematization while causing errors against competing forms, and also, interacts with the pattern learning mechanisms. Henceforward, this study documented the importance of frequency, regardless of any other language acquisition explanation.

Finally, a few models that we have influenced and adapted to our own model can be found on GitHub. We were altered by the materials of Labs 2, 3, and 4, used in Simulating Language class, one of which was taught at Edinburgh University [49]. Lab 2 is the simplest way that we could start. It is a simplified version of the model in [41], which is a simpler version of the model in [44] word learning. The model allows us to create meanings for words as referent lists. Then, a hypothesis space is created and Bayesian rules take care of the rest. Lab 3 model is a simple Bayesian model of frequency learning. This model allows us to explore the effects of the prior and the data on frequency learning as in the model [50]. Finally, Lab 4 presents us the iterated Bayesian learning model. This model combines the replication of the iterated learning model of the evolution of frequency distributions [51] and is built around the Bayesian model of frequency learning/regularisation from the Lab 3 model. This model also allows us to explore the effects of learning biases.

Also, we have gotten ideas from the tutorial of agent-based models tutorial [52]. In this repository, Part 3 has a complex model to represent a population of agents that interact with each other. Agents have different personality types (such as stubborn or flexible, which is a parameter affects learning flexibility). Then multiple simulations can be done with different random populations in this model.

2.3 Network Structures

Another issue that this study deals with is the structures of communities. Of course, today there are communities that communicate in different sizes and in different ways. A previous study reveals that the size and social structure of the community may play an important influence in the evolution of language ([53]; [54]; [15]). It was theorized that emerging sign languages that emerge in tiny groups (i.e. village sign languages) had less conventionalized structure. According to [55], languages forming in larger groups and/or communities with less common heritage (deaf community sign languages, for example) tend to be more uniform.

[56] used a laboratory setting to investigate how word order preferences vary with different parameters: learning biases, size of the community, and the amount of data that participants are exposed to. Their results showed that the size of the population and the amount of data they are exposed to play a significant role in language convergence. If the participants were exposed to data from one and single speakers, results show that listeners learn with frequency, and variability was seen to persist over three generations. But with the same amount of distributed data from multiple speakers, frequency learning was reported to be failed. They showed that there is a more prominent bias for SOV than OSV and VSO. In addition, it is reported that the effect of the amount of data exposed from multiple speakers contributed greatly to frequency learning.

[57] investigated the word order variability to test the hypothesis. They searched how word order becomes a standard in new communication systems that differ in their social structure and community

size by using real data. The results showed that there is significantly more variance in word order preferences in CTSL as opposed to those in ABSL, both within and across signers: CTSL signers show less convergence as a community and are less consistent in their own productions. These results support the hypothesis that the size of a language community has an effect on conventionalizing in the early stages of language emergence: the language of bigger communities is more uniform in structure than that of smaller communities.

CHAPTER 3

THE MODEL

In this chapter, we will explain the dynamics of our model and how we adopted the information mentioned in Chapter 2.

The primary linguistic process under examination in our study is the word order preferences in reversible vs. irreversible occurrences. In the previous section, we have given that events involving two characters from the same semantic category (animate-animate, or reversible events) present a communicative pressure on people. The roles of the agent and the patient (i.e. male or female) suffer from ambiguity, especially since there is no systematic preference for word order in young emerging communication systems. On the other hand, it is unlikely to experience ambiguity in non-reversible situations. As in an earlier example, it will be the man who cuts the bread, since bread will not be the subject here to cut the man, it will be understood correctly by the listener no matter how the sentence is formed if all other conditions are normal.

Our model consists of three main components: an individual learning agent class with appropriate attributes and functions; a model of the iterated learning functions of the population of these agents; and finally a series of functions to run the iterated learning of the population of these agents with different network structures and circumstances.

Basically, at the agents level, agents choose random values from existing lists according to their current weights correspondingly and start the learning process about it, then, update their existing knowledge given in Figure 6.

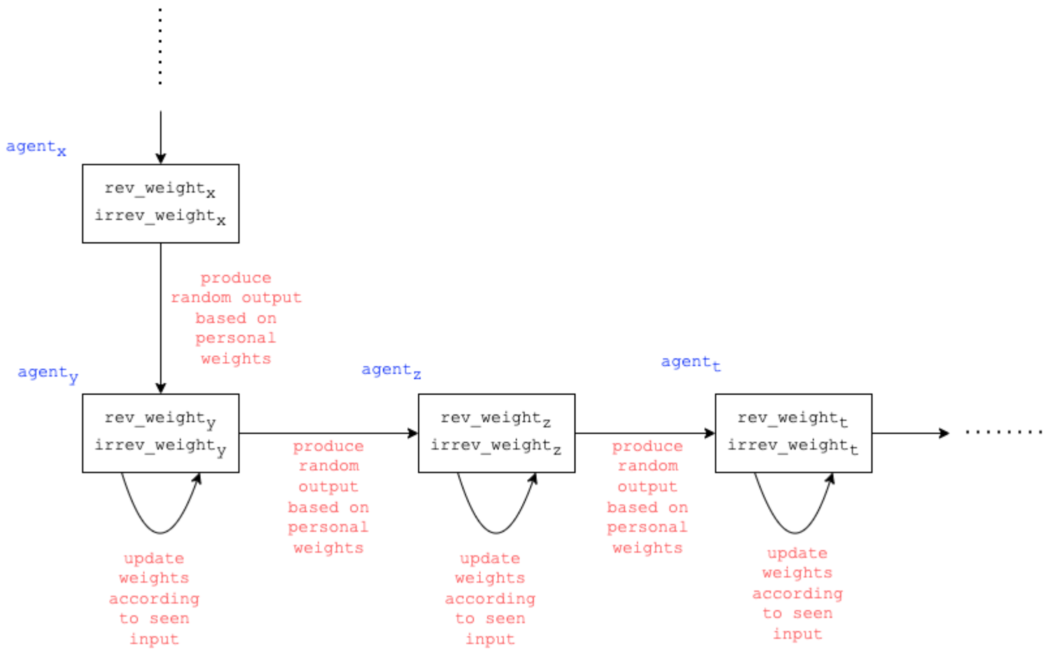


Figure 6: Learning process of agents

At the iterating learning level, populations are created and start communicating, then, each generation (that lives together) updates their knowledge according to their communication preferences, then, feeds the next generations given in Figure 7.

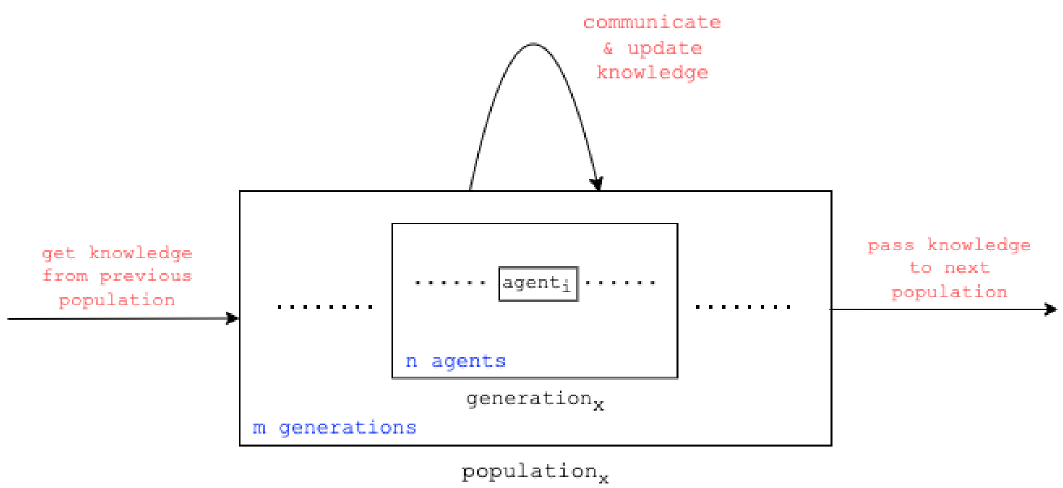


Figure 7: Iterated learning process of populations

3.1 Global Variables

There are some fixed variables in the model. Some will be modified throughout the simulating process (i.e. the number of children, starting biases, error/pressure rate, communication tendency for a word order); some will remain still (i.e. personalities, sentence types list, sentence weight list, and basic word orders list).

3.2 Agents

Agents are the units that represent people in this model. They have various attributes and functions. These will be explained in this section.

3.2.1 Attributes

Attributes are the properties hold by the class. All agents have generation numbers, personality types, reversible weights, and irreversible weights. If an agent is from the first generation, it receives starting bias weights; if it is from another generation, it receives the average weights of its mother's and father's weights, as in **??**.

The generation number increases when the current generation has children, and these children get that generation number. This generation number is important for the generations living together to form a population. A population is usually a community of 3-4 generations living together. In addition to these, there is also personality information. There are currently 2 types of personalities for this version of the model: flexible and stubborn. These will affect when agents update their weights for sentences to which they are exposed.

Each agent will produce word orders to communicate with other agent/s. Here, it is assumed that agents can distinguish reversible and irreversible sentences. For these two sentence types, agents separately record the word orders as weights that they encounter/hear. These weights represent the corresponding order in the fixed word order list defined at the beginning. An example is given in Figure 8.

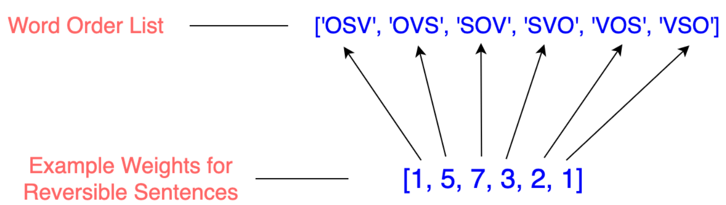


Figure 8: Example weights of corresponding word orders

3.2.2 Learning Process: Updating Weight Functions

Our model is based on matching the exposed word order with the related sentence type and then updating the weight of the corresponding word order. The basic process is explained in Figure 9. However, with different personalities and different pressure from the environment, agents update their weights differently. Pressure can be caused by education, network, or language. Education is repressive because it aims to make people learn the same way by reducing flexibility in the language. Education makes each agent obey the rules and learn in the same way. Besides, networks we build with our environment can also help with regularity in language. In the studies, [15][16], we have mentioned in the previous section, we said that larger groups create more regular languages. Therefore, larger and regular groups can provide regular languages.

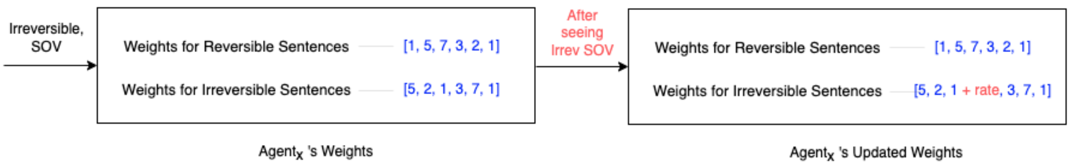


Figure 9: Example basic updating weights for an exposed sentence

Weights update functions are performed by adding a certain value to the corresponding word order in the sentence type encountered while handling personality traits, and pressure effects if applicable. For example, if learning takes place without any influence, the agent will only be affected by its personality. That is, it will update the weight of the corresponding word order in the sentence type it encounters. If the agent has a stubborn (S) personality, it will accept the situation hard and learns slowly, that is, it gives a low weight value; if the agent has a flexible (F) personality, it will give higher weight to the word order and want to make updates (add error value) in other word orders, too. These functions return a list of values to add the agent's current weight list of the corresponding sentence type.

Although the weight update function works in a similar way in the pressure environment, the agent degrades the value equally from the other word orders while updating the word order it encounters as in Figure 10.

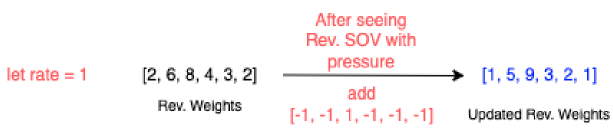


Figure 10: Example basic updating weights in an environment with pressure

3.3 Generations & Populations

Generations indicate a group of people (in our case, agents) born and living at about the same time. In this study, we will use the term "generation" as the birth period order. For example, while the generation number of the first agents of the community is 0, the children of this first generation will be considered generation 1, the children of generation 1 will be considered generation 2, and so forth. Here, it is necessary to distinguish between the first generation and the following generations. Communities will be built on their first generation.

3.3.1 First Generation

Unlike other generations, the first generation does not have a traditional input for language. So, we have to assume that there is no other generation to learn the language from. In this case, this generation should give direction to the language. The agents of this generation are given the starting bias. According to their bias, they create word orders to communicate.

3.3.2 Children

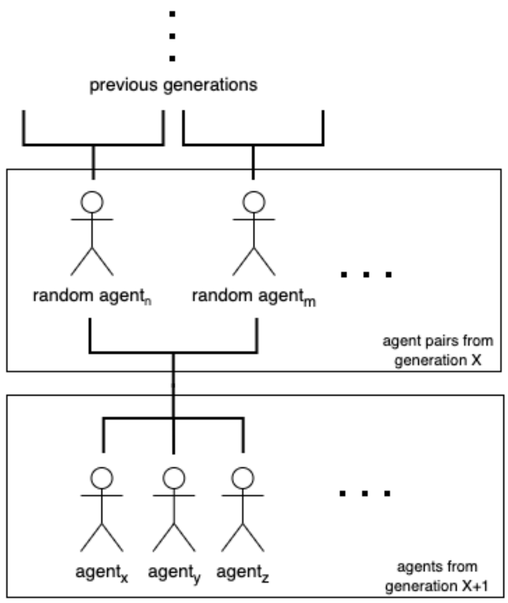


Figure 11: Example for the process of creating a next generation.

In the starting code, each pair of agents can have 1 to 3 children. For example, in Figure 11, a pair of random agents, in the same generation, have three child agents. It has been observed that a healthier and more realistic population distribution is created in the trials created with these numbers.

3.4 Iterated Learning

The basic elements of the model is given so far. But the core idea of the dynamics of the model is iterated learning. Iterated learning is seen as key mechanism of language evolution, as [14] have used iterated learning to explain the origins of structure in language, in our study we adopted this concept, too. All agents are updating their current knowledge with the other agents' knowledge. In Figure 12, it is shown that learning process is based on mapping an exposed word order to the sentence type. Word order is transmitted from one individual to another. This happens by the emergence of compositionality: parts of a sentence become systematically linked to parts of its meaning.

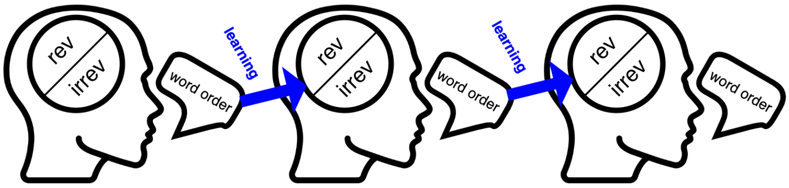


Figure 12: World order transmission

Over the time and transmission of language through generations, languages are changed and transformed. By applying iterated learning in the model, it is aimed to see a convergence to the today's languages' word order distributions. In the model, iterated learning has two levels: through individuals, through populations. In individual level of iterated learning, agents communicate with each other and update their knowledge. In population level, populations communicate each other and update their knowledge, accordingly.

3.5 Communication Process & Network Types

Word order transfer via communication will be given and explained in this section. The proposed model's communication process closely relates to how agents form a community. In some cases, forming a community can be one-on-one (point-to-point). Another way to form a community is where one speaks and others listen (star). where there are random conversations in a particular group and everyone listens (multiple stars in a group), or where everyone can listen to everyone else speak (totally connected/mesh). These were created from network topology types in computer science [58] as in expressed in Figure 13 and were inspired by the work [15][16] and all will be explained seperately.

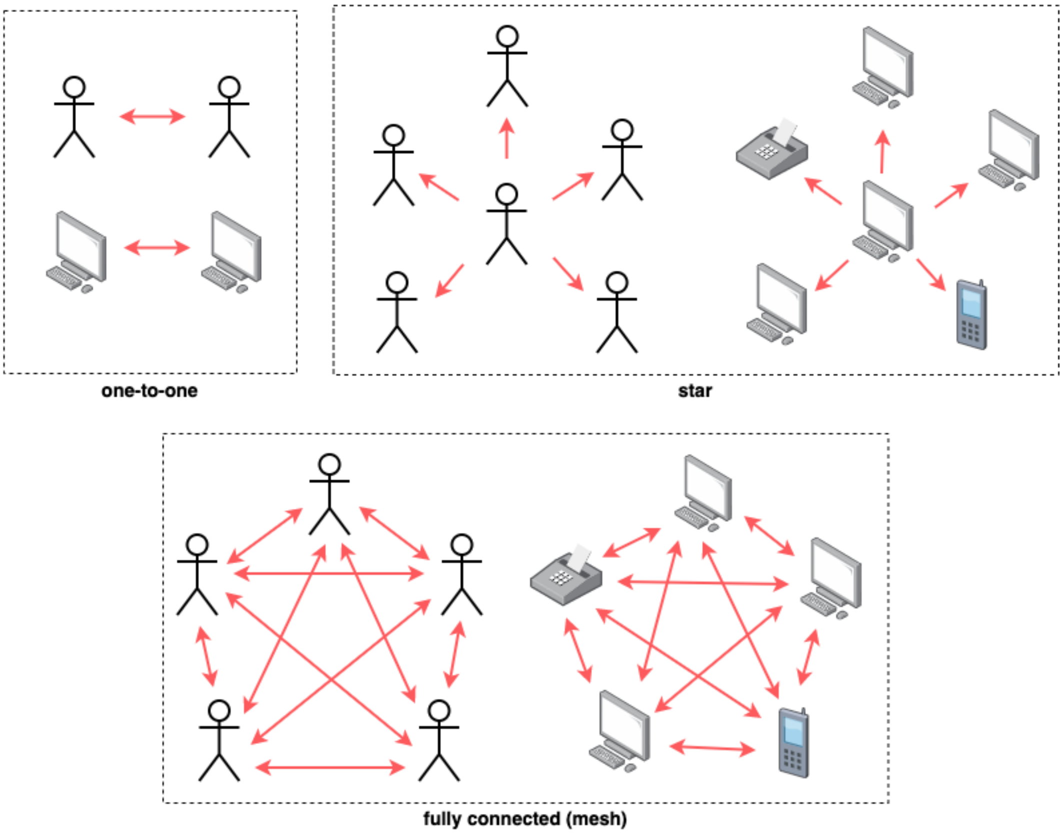


Figure 13: Network topologies and how agents form communication groups according to them.

The process basically consists of a speaker agent and its listener agent(s) and the function(s) they are using for communication. First, a list of sentence types is created, and then the speaker agent picks a word order according to the sentence type pops from the sentence types list, according to current weights for the picked sentence type produces a word order. Finally, listener agent(s) update the order they are exposed by adding them to their existing weighting system.

3.5.1 Sentence Types Producing

In each communication process, a list of sentence types that are 70% reversible and 30% non-reversible is produced for the given number of sentences. The model uses Python's random.choices function in several places. This method is used to select elements from a population randomly. This method also allows us to determine the weights assigned to the possibility of picking every value at random. If no weights are specified, the choices are selected with equal probability.

3.5.2 Word Orders Producing

Word order producing process can be visualised as a roulette wheel as in Figure 14. The probability for each word order type to be chosen is proportional to their areas in the roulette wheel as used in [49] models. Again, random.choices function does a simulation for picking up values from a roulette wheel. The speaker agent pops a sentence type from the sentence types list produced at the beginning of the communication.

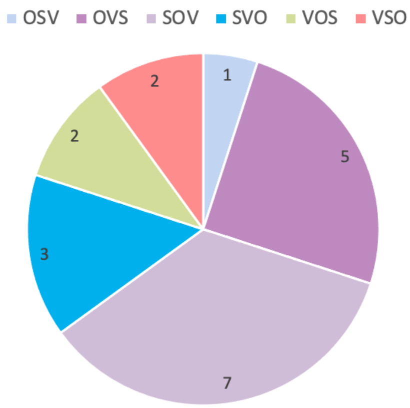


Figure 14: A roulette wheel areas according to weights of [1, 5, 7, 3, 2, 2]

3.5.3 One-to-One Type Communication (Two Agents)

According to community type, agents interact with each other differently. In one-to-one type communication only two agents communicate with each other according to Algorithm 1.

Algorithm 1 One-to-One: Two agents communication algorithm

$n \leftarrow$ number of sentences ▷ taken as parameter
$population \leftarrow$ the current population ▷ taken as parameter
$s_L \leftarrow$ createSentenceList(n)
$selectedAgents \leftarrow$ selectTwoRandomAgents(population)
for $i \leftarrow 1$ to n **do**
  $speaker \leftarrow$ randomIndex from selectedAgents
  $listener \leftarrow$ not randomIndex
  if $s_L[i]$ = 'irreversible' **then**
    $spokenWordOrder \leftarrow$ produceWordOrder(speaker.IrrevWeight)
    $listenerIrrevWeight \leftarrow$ updateIrrevWeight($spokenWordOrder$)
  else if $s_L[i]$ = 'reversible' **then**
    $spokenWordOrder \leftarrow$ produceWordOrder(speaker.RevWeight)
    $listenerRevWeight \leftarrow$ updateRevWeight($spokenWordOrder$)

3.5.4 Star Type Communication (N Agents: 1 Speaker N-1 Listener)

Star type communication is very similar to one-to-one communication but with many agents. In one-to-one type communication only two agents communicate with each other but in start communication, many agents create multiple one-to-one communications with randomly chosen speakers according to Algorithm 2.

Algorithm 2 Star: N agents communication algorithm

$nAgents \leftarrow$ number of agents ▷ taken as parameter
$n \leftarrow$ number of sentences ▷ taken as parameter
$population \leftarrow$ the current population ▷ taken as parameter
$s_L \leftarrow$ createSentenceList(n)
$selectedAgents \leftarrow$ selectNRandomAgents(population, nAgents)
for $i \leftarrow 1$ to n **do**
    $speaker \leftarrow$ randomIndex from selectedAgents
    if $s_L[i] =$ 'irreversible' **then**
        $spokenWordOrder \leftarrow$ produceWordOrder(speaker.IrrevWeight)
        for $listener \leftarrow 1$ to nAgents **do**
            if $listener \neq speaker$ **then**
                $listenerIrrevWeight \leftarrow$ updateIrrevWeight($spokenWordOrder$)
    else if $s_L[i] =$ 'reversible' **then**
        $spokenWordOrder \leftarrow$ produceWordOrder(speaker.RevWeight)
        for $listener \leftarrow 1$ to nAgents **do**
            if $listener \neq speaker$ **then**
                $listenerRevWeight \leftarrow$ updateRevWeight($spokenWordOrder$)

3.5.5 Mesh Type Communication (All Involved)

In mesh-type communication, all community members are guaranteed to be involved in the communication. For all members in a group, each agent is selected as the speaker once, and all others become listeners. Then, speaker agents produce n-word orders for n sentences, and all listeners update their weights for the spoken word order of the related sentence type according to the Algorithm 3.

Algorithm 3 Mesh: All involved communication algorithm

```
groupMemberIndices ← indices of group members            ▷ taken as parameter
n ← number of sentences                                   ▷ taken as parameter
population ← the current population                       ▷ taken as parameter
for s in groupMemberIndices do                            ▷ speaker
    for l in groupMemberIndices do                        ▷ listener
        if s ≠ l then
            s_L ← createSentenceList(n)
            for sentence in s_L do
                if sentence = 'irreversible' then
                    spokenWordOrder ← produceWordOrder(s.IrrevWeight)
                    lIrrevWeight ← updateIrrevWeight(spokenWordOrder)
                else if sentence = 'reversible' then
                    spokenWordOrder ← produceWordOrder(s.RevWeight)
                    lRevWeight ← updateRevWeight(spokenWordOrder)
```

CHAPTER 4

SIMULATIONS AND THEIR RESULTS

In the previous chapter, it is given how the model is shaped generally. In this chapter, it will be given the results of simulations with different parameters. We will try to answer the research questions given in the introduction. The output of the model is the frequencies of the word orders of the last population (3-4 generations) in the community. To get that, it is provided that each agent of the population produces n number of sentences, then all summed upped and given as a frequency graph. We will try to seek answers by changing the 7 suggested parameters for each research question. These are bias type (uniform, biased, random); the number of generations (5, 10, 25); the existence of a tendency towards a specific word order (yes, no), first community size (25, 50, 100); network type (mesh, star, one-to-one); (spoken) data size (less-1000, more-5000); and personality distribution in the community (F>S, F=S, S>F). With all these parameters, 972 possible test cases have been run.

The seeking answers to research questions can be summarized as follows: First, it was claimed that SOV is the cognitively most basic word order humans use. We asked what if, SOV is not the most basic word order to start a language. To answer that, we can start our model with different starting word order biases, and see if the word orders converge to one of the most frequent orders today (i.e. SOV or SVO) without changing another parameter. On the other hand, the tendencies of languages/communities may change over time, that is not emerged at the beginning of a language. We can check this by keeping the starting bias constant and observing it by changing the tendency to word order.

Second, we wanted to answer how a community's size affects the dominant word order emergence and evolution speed. To observe that, we can change the size of the first generation of the community and compare the results. With different first community sizes, population sizes will be different at an exponential rate. We can also observe how the number of generations can affect the dominant word order emergence and evolution speed. The more generations, the more communication and the more regular and mature languages.

Third, some languages allow only one order (i.e. English) because of the ancestors' preferences throughout the language evolution; but some languages allow multiple orders and handle problems with different mechanisms like case marking, phonology, etc. To capture this effect, the existence of a tendency towards a specific word order is used as a parameter. If a language allows only an order, then, there is a tendency. A tendency distribution is adapted from the current distribution in the world today given in [1].

In studies [15] and [16], it was said that the way communities are shaped and the size of groups create more regular languages faster as they feel more pressure to understand each other. Multiple simulations with different network types and community sizes will be executed to observe these effects.

Moreover, we want to observe how personal differences (stubborn & flexible, in our case) will affect the convergence of the word orders. So, what will it change if many/few people change their language? And what does it tell us about real life? Will word orders remain fixed without language changes? Simulations with more or less counter-personality agents will be run.

In this chapter, the values of the parameters other than the controlled parameter are as follows:

- **starting bias:** uniform
- **number of generations:** 25
- **tendency for a word order**: yes
- **first community size:** 50
- **network type:** mesh
- **data size:** 5000
- **personality distribution:** F=S

4.1 Different Simulation Parameters

To see the effects of how different parameters change language, language evolution and language evolution speed, the results will be given comparatively.

4.1.1 Starting Bias

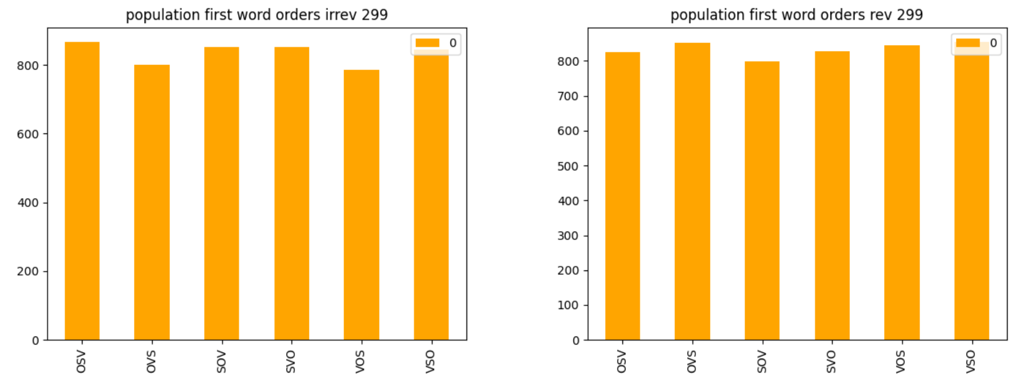
Three different starting bias types are used as a parameter in the model; these are uniform, biased and random biases. With this parameter, we want to observe the effect of our innate bias. Although studies say that we tend to certain word orders in newly emerged communities that produce new languages, and laboratory experiments; we want to see if it is possible to reach today's word order distribution without this effect.

Here, we gave two graphs of irreversible and reversible word order frequencies that represents bias distribution of the community before starting any communication.

4.1.1.1 Uniform

Uniform bias means the equal tendency to all word orders, that is these agents do not prioritize a word order as in Figure 15.

!

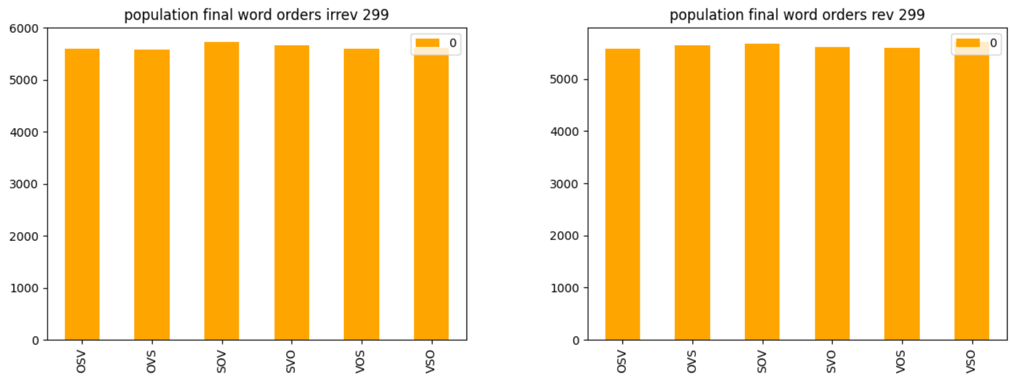


(a) irreversible sentences word orders

(b) reversible sentences word orders

Figure 15: First generation word order distributions with uniform bias of a population before any communication

!



(a) final irreversible sentences word orders

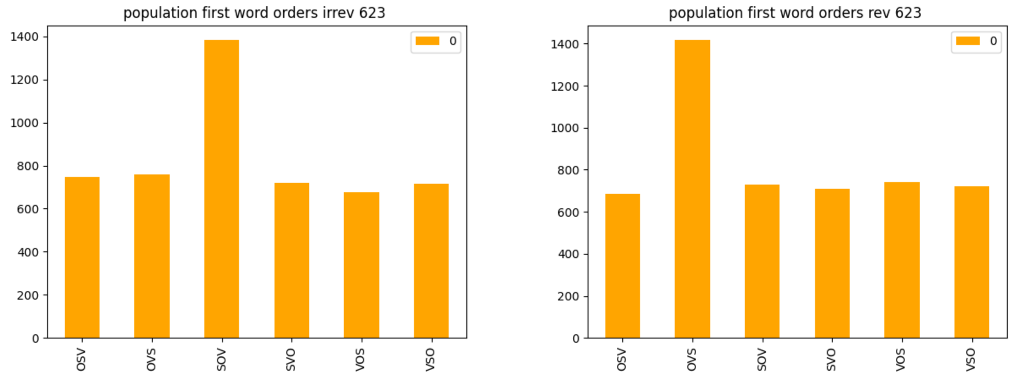
(b) final reversible sentences word orders

Figure 16: Final word order distributions with uniform bias of a population after all communications

4.1.1.2 Biased

Biased bias means the tendency towards a specific word order is present, that is people use a word order more frequently than others. Which word order this bias will be is randomly determined. Here one word order is determined to be twice as important as the others. So, for example as in Figure 17, OSV is prioritized for irreversible sentences, the weights for VSO, OVS, SOV, SVO, VOS, VSO are assigned as follows: 1x, 1x, 1x, 1x, 1x, 2x; and OSV is prioritized for reversible sentences, and the weights would be 2x, 1x, 1x, 1x, 1x, 1x, accordingly.

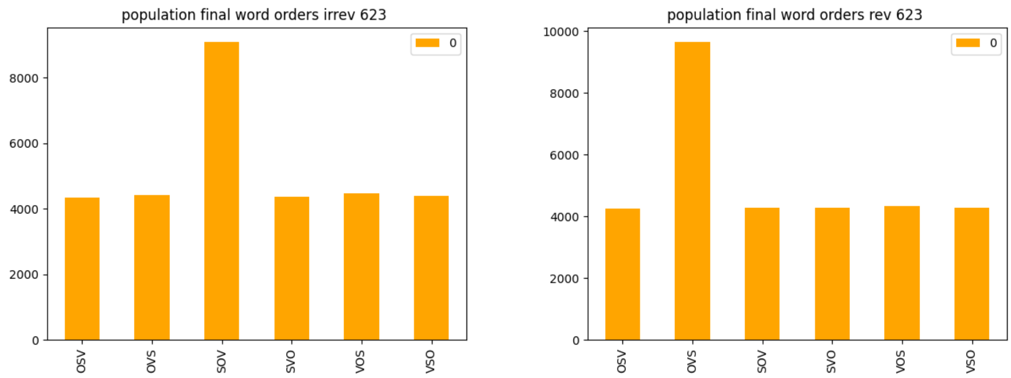
!



(a) irreversible sentences word orders (b) reversible sentences word orders

Figure 17: First-gen word order distributions with the biased bias of a population before any communication

!



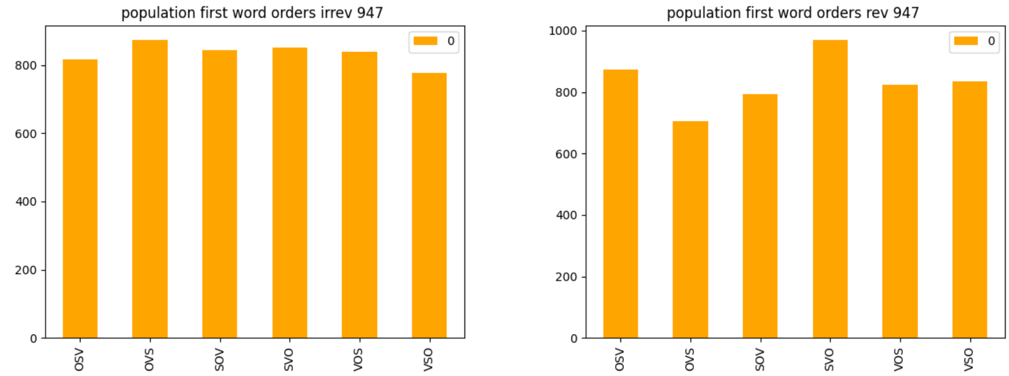
(a) final irreversible sentences word orders (b) final reversible sentences word orders

Figure 18: Final word order distributions with uniform bias of a population after all communications

4.1.1.3 Random (or Varied)

Random bias means that there is no specific tendency towards a word order and every member has their own preferences, that is people use different word orders with no logical selection as in Figure 19. This looks like the uniform bias but in uniform bias, everyone has the uniform bias, but in random bias, everyone has their own unique bias different from others. The reason that the result looks like the uniform bias is that everyone produces their own sentences and the whole graph converges to a uniform shape.

!

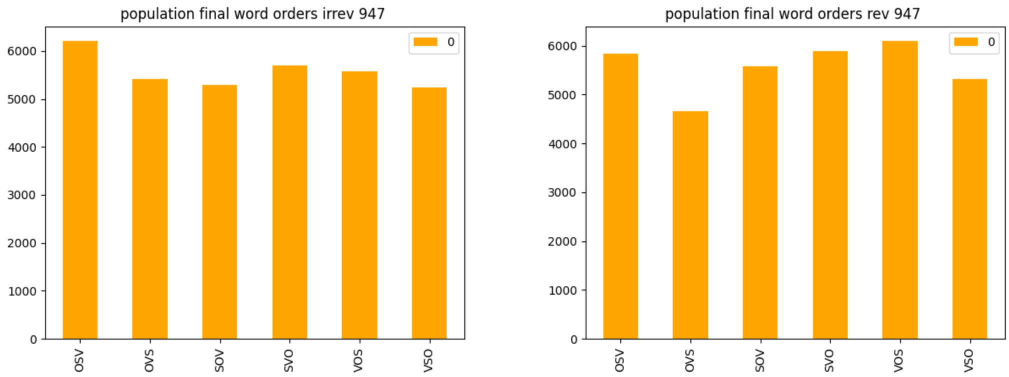


(a) irreversible sentences word orders

(b) reversible sentences word orders

Figure 19: First-gen word order distributions with the random bias of a population before any communication

!



(a) irreversible sentences word orders

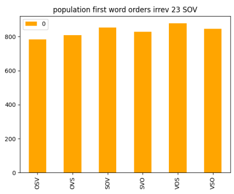
(b) reversible sentences word orders

Figure 20: Final word order distributions with the random bias of a population after communications

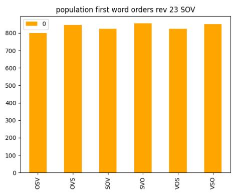
4.1.2 Number of Generations

Number of generations is one of the parameters used in the simulation. Here, we expect language to become mature and gain regularity as the number of generations increases. Because the ancestral preferences, rather than the uncertainties in a newly emerged language, will guide the next generations. Some word orders may disappear, or a new word order may be preferred than others.

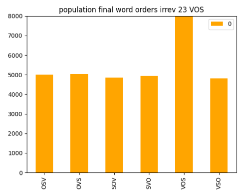
!



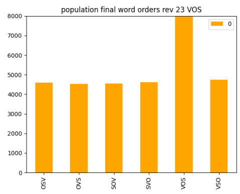
(a) 5 gen first irrev.



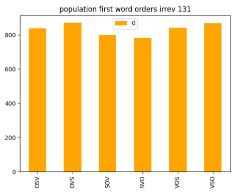
(b) 5 gen first rev.



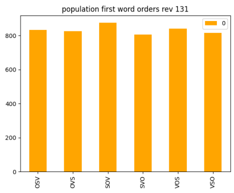
(c) 5 gen final irrev.



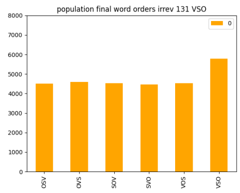
(d) 5 gen final rev.



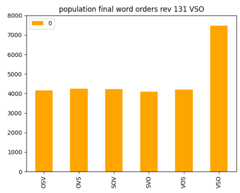
(e) 10 gen first irrev.



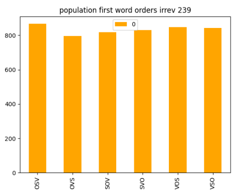
(f) 10 gen first rev.



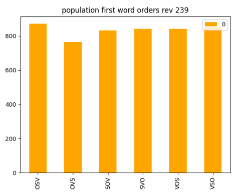
(g) 10 gen final irrev.



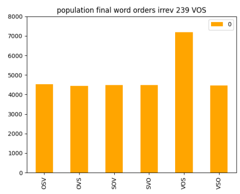
(h) 10 gen final rev.



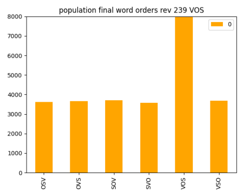
(i) 25 gen first irrev.



(j) 25 gen first rev.



(k) 25 gen final irrev.



(l) 25 gen final rev.

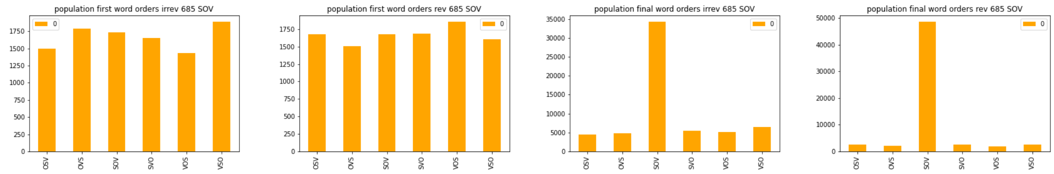
Figure 21: First and final word order distribution with different numbers of generations

The effect observed with the increase in the number of generations is given in Figure 21. Since the figures (a), (b), (e), (f), (i), (j) express the same starting bias frequencies, (c), (g), (k) and (d), (h), (l) should be compared. Although there does not seem to be a big difference in the proportions of frequencies due to the weights preferred in the model, as the number of generations increases, the dominant word order that emerges in the language is used more frequently than the others.

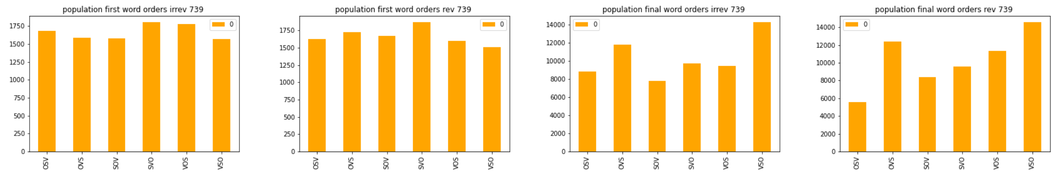
4.1.3 A Tendency for a Word Order

Another parameter in the model was the presence of a tendency to any word order. The tendency mentioned here may arise from the needs of the language itself, or it may be the case of returning to the more preferred word order over time, although it is not necessary. For example, logical languages like English have to follow the word order. Because the sentence gains meaning thanks to the order. However, in agglutinative languages such as Turkish where all ordering is possible, suffixes give meaning to sentences rather than word order. Despite this, the dominant word order in Turkish is said to be SOV.

!



(a) first irrev. (b) first rev. (c) final irrev. (d) final rev.



(e) first irrev. (f) first rev. (g) final irrev. (h) final rev.

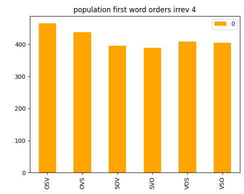
Figure 22: First and final word order distribution with (a, b, c, d) and without (e, f, g, h) the tendency/-pressure towards a word order

In Figure 22, (c) and (d) are the final word order frequencies of a community with a word order tendency. This can be thought of as English. The ancestors in this society may have set a word order as a rule and had to obey it; on the other hand, although there is no tendency in the case of (g) and (h), some orders may have tended to be used more frequently than others over time. Still, there is no obvious difference as in (c) and (d). This can be one of the languages that accepts all word orders, but some word orders are used more, like Turkish.

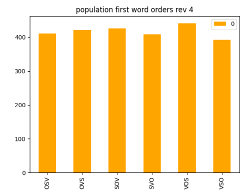
4.1.4 First Community Size

While running the simulations, the size of the first community is also given as a parameter. Here, the number of people in the community using an emerged language is meant. Again, it should be accepted that all of these people have connections with each other and contribute to the same language. In the chapter where we talked about background works, we said that large communities create more regular languages. The regularity here may be that in order for each member of the community to understand and communicate in the same way, they need to agree on a certain rule rather than variety. When the number of people in the community is small, it may be easier for people to understand each other and remember their preferences, even if everyone uses a different word order. But as the community grows, this will become more difficult.

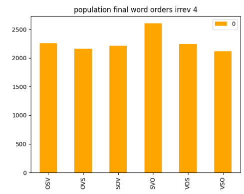
!



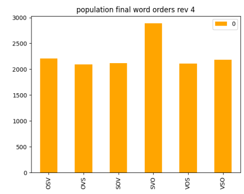
(a) first irrev.



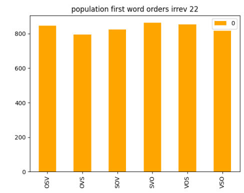
(b) first rev.



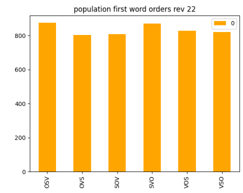
(c) final irrev.



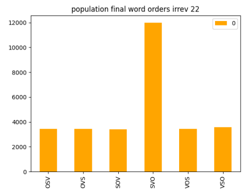
(d) final rev.



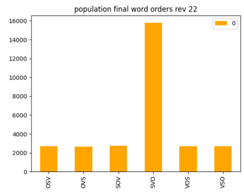
(e) first irrev.



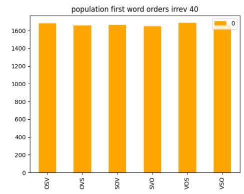
(f) first rev.



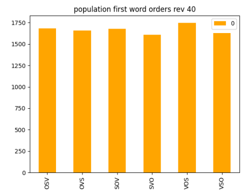
(g) final irrev.



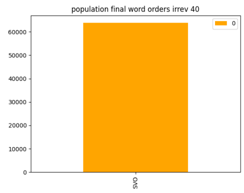
(h) final rev.



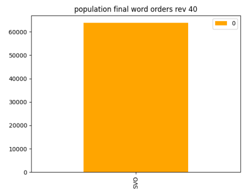
(i) first irrev.



(j) first rev.



(k) final irrev.



(l) final rev.

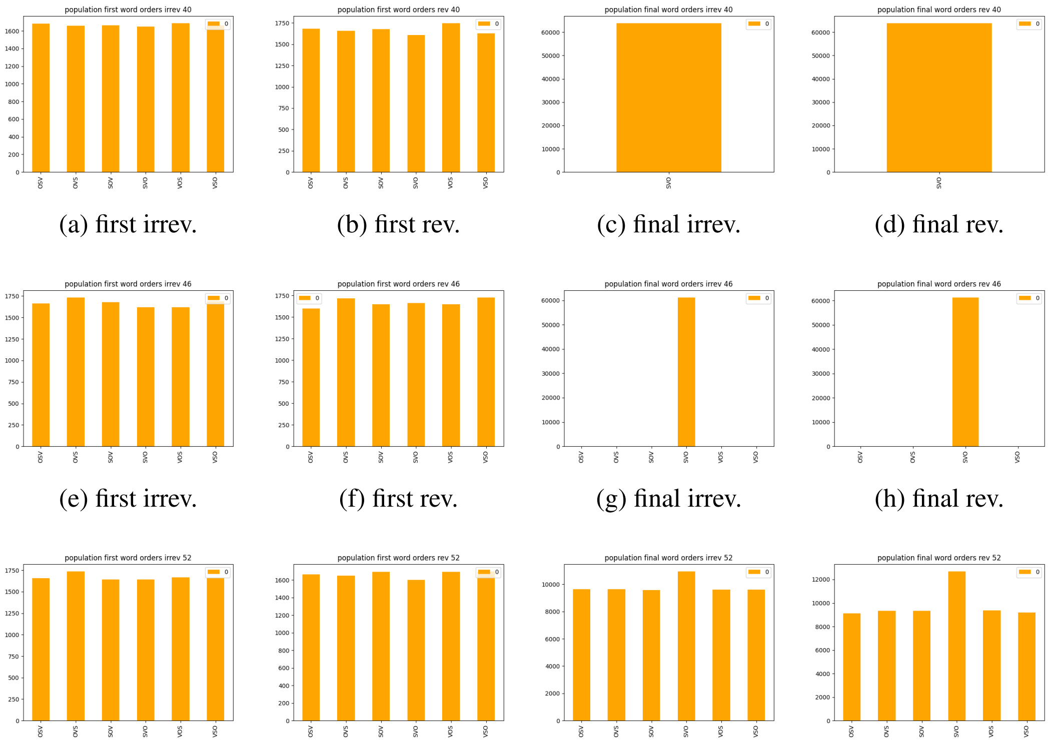
Figure 23: First and final word order distribution with 25 people in the first community (a, b, c, d); with 50 people in the first community (e, f, g, h); and with 100 people in the first community (i, j, k, l).

In our model, the frequencies of the word orders used, depending on the size of the initial community, are given in Figure 23. When (c)-(d), (g)-(h), (k)-(l) are examined comparatively, it can be seen that the preferred word order becomes more prominent (or even reduced to one as in (k)-(l)) as the number of first people in the community increases.

4.1.5 Network Type

As with the size of the community, the way the community connects with each other can also be an impact. We say that even if this is not result-oriented, it can speed up the regularization of the language. We mentioned earlier that we emulate network connection patterns in computer science. In society, the connection of people with other people may be different. While people in a particular area may determine their own effects on language when disconnected and unaware of others, there may be fully connected situations where everyone is exposed to the same form of language. Everyone does not need to talk to everyone, the important thing is to have access to the same form of language. For example, this can happen with education. If everyone is exposed to the language in the same way, they can use it in the same way.

!



(a) first irrev. (b) first rev. (c) final irrev. (d) final rev.

(e) first irrev. (f) first rev. (g) final irrev. (h) final rev.

(i) first irrev. (j) first rev. (k) final irrev. (l) final rev.

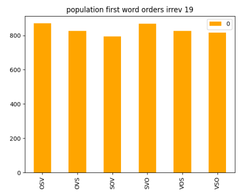
Figure 24: First and final word order distribution with mesh network type (a, b, c, d); with star network type (e, f, g, h); and with one-to-one network type (i, j, k, l).

In Figure 24, (c)-(d) show final word order frequencies when all agents are fully connected, only one word order is preferred at the end. (g)-(h) show final word order frequencies when all agents are not fully connected but partially connected with subgroups, and although one word order seems to be used more frequent than the others, there are other word orders in the community, but their lifespan seems short. And finally, (k)-(l) show final word order frequencies when all agents are neither fully nor partially connected, they are in the form of network where they communicate among themselves one-by-one in small groups, and as it is seen, there remains variety because no agreement can be reached in this disorganized and disconnected communication structure. It will be difficult to convince everyone to use one language in the same way.

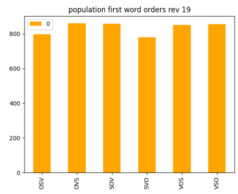
4.1.6 Data Size

In the model we proposed, we stated that a certain number of sentences (word orders) are produced in each communication cycle and agents update their current beliefs in each production. Here too, data size can be evaluated as the number of sentences produced (used) during communication. What we want to observe here is how more/less sentences will affect the regularity of the language.

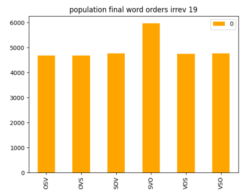
!!



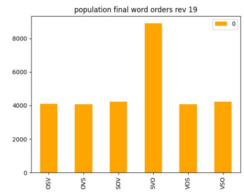
(a) first irrev.



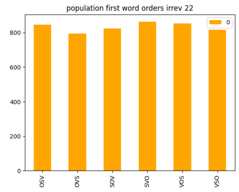
(b) first rev.



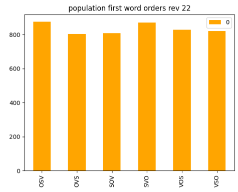
(c) final irrev.



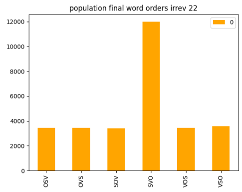
(d) final rev.



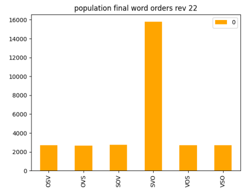
(e) first irrev.



(f) first rev.



(g) final irrev.



(h) final rev.

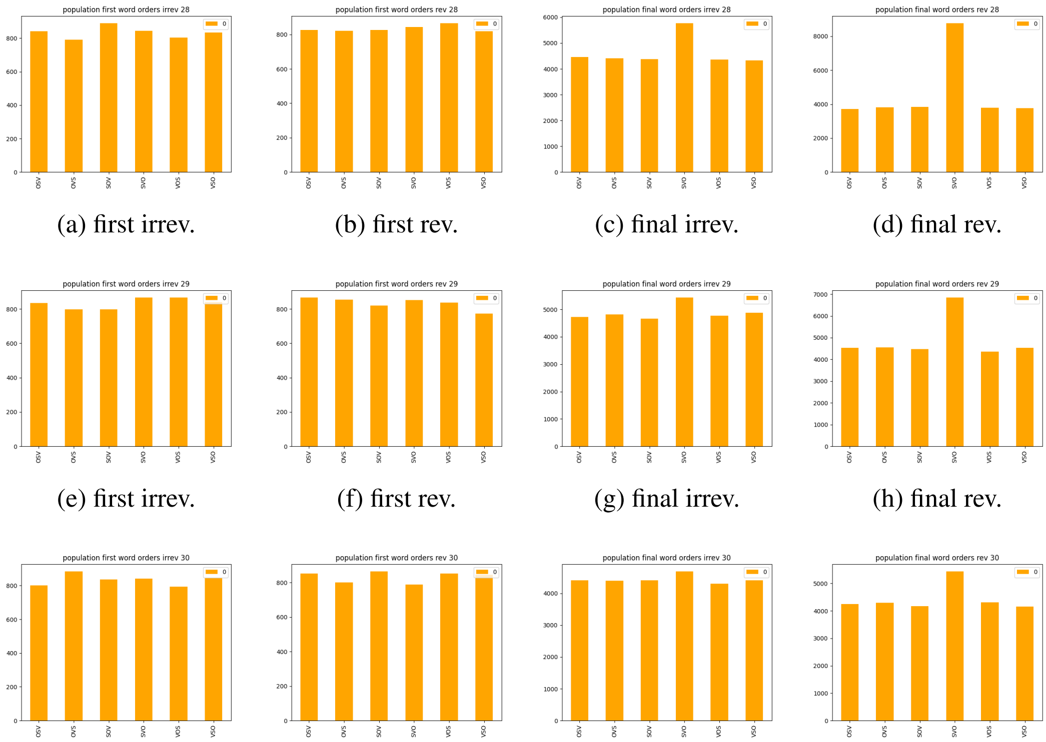
Figure 25: First and final word order distribution of community using 1000 sentences at each communication (a, b, c, d); community using 5000 sentences at each communication (e, f, g, h).

As seen in Figure 25, the dominant word order emerges faster if more sentences are used during communication (by comparing (c)-(d) and (g)-(h)).

4.1.7 Personality Distribution

As we have stated before, not everyone evaluates and uses the information they have acquired in the same way. If we think that the distribution of these people may be different as a social structure, we can look at how it can affect the emergence of the dominant word order. We mentioned 2 personalities in our model: stubborn and flexible. Stubborn people are more difficult to update their existing beliefs, and flexible people are more open to change and could update their knowledge more easily.

!



(a) first irrev. (b) first rev. (c) final irrev. (d) final rev.

(e) first irrev. (f) first rev. (g) final irrev. (h) final rev.

(i) first irrev. (j) first rev. (k) final irrev. (l) final rev.

Figure 26: First and final word order distribution of community where number of flexible agents is more than the number of stubborn agents (a, b, c, d); of community where F=S (e, f, g, h); and community where F<S (i, j, k, l).

In Figure 26, if flexible people is more than stubborn people (i.e. (c)-(d)) dominant word order may emerge more quickly. In the opposite case, it can be said that it may be difficult or late to mention a dominant word order.

In this section, we have summarized the simulations with one example for each to represent their behaviour in different parameters. Similar comments apply in all 972 test cases. The effects created by the comparisons appeared more or less compared to the situation of the other parameters.

In the next and last section, we will summarize what we have done and give a general conclusion. At the same time, we will discuss what changes can be made in the future, what the proposed model can be used for.

CHAPTER 5

CONCLUSION AND FUTURE WORK

In this last chapter, the results of the simulations set forth in the previous chapter are discussed and some future directions are given based on the findings of the study conducted in this thesis.

5.1 Overall Discussion of the Results

5.2 Future Directions

REFERENCES

[1] M. S. Dryer, "Order of subject, object and verb," 2013.

[2] G. L. Lewis, *Turkish Grammar*. Oxford: Clarendon Press, 1967.

[3] E. E. Erguvanli, "The function of word order in turkish grammar. berkeley," 1984.

[4] D. J. Napoli and R. Sutton-Spence, "Order of the major constituents in sign languages: Implications for all language," *Frontiers in psychology*, vol. 5, p. 376, 2014.

[5] T. Briscoe, "Grammatical acquisition: Inductive bias and coevolution of language and the language acquisition device," *Language*, vol. 76, no. 2, pp. 245–296, 2000.

[6] S. Goldin-Meadow, W. C. So, A. Özyürek, and C. Mylander, "The natural order of events: How speakers of different languages represent events nonverbally," *Proceedings of the National Academy of Sciences*, vol. 105, no. 27, pp. 9163–9168, 2008.

[7] T. Givón, *On understanding grammar*. Academic Press, 2014.

[8] B. B. Blake, "Russell s. tomlin, basic word order. functional principles. london: Croom helm, 1986. pp. 308.," *Journal of Linguistics*, vol. 24, no. 1, p. 213–217, 1988.

[9] L. Maurits, D. Navarro, and A. Perfors, "Why are some word orders more common than others? a uniform information density account," in *Advances in Neural Information Processing Systems* (J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, eds.), vol. 23, Curran Associates, Inc., 2010.

[10] M. L. Hall, R. I. Mayberry, and V. S. Ferreira, "Cognitive constraints on constituent order: Evidence from elicited pantomime," *Cognition*, vol. 129, no. 1, pp. 1–17, 2013.

[11] M. Schouwstra and H. de Swart, "The semantic origins of word order," *Cognition*, vol. 131, no. 3, pp. 431–436, 2014.

[12] A. Kocab, H. Lam, and J. Snedeker, "When cars hit trucks and girls hug boys: The effect of animacy on word order in gestural language creation," *Cognitive science*, vol. 42, no. 3, pp. 918–938, 2018.

[13] I. Meir, M. Aronoff, C. Börstell, S.-O. Hwang, D. Ilkbasaran, I. Kastner, R. Lepic, A. L. Ben-Basat, C. Padden, and W. Sandler, "The effect of being human and the basis of grammatical word order: Insights from novel communication systems and young sign languages," *Cognition*, vol. 158, pp. 189–207, 2017.

[14] S. Kirby, T. Griffiths, and K. Smith, "Iterated learning and the evolution of language," *Current opinion in neurobiology*, vol. 28, pp. 108–114, 2014.

[15] L. Raviv, A. Meyer, and S. Lev-Ari, "Larger communities create more systematic languages," *Proceedings of the Royal Society B: Biological Sciences*, vol. 286, no. 1907, p. 20191262, 2019.

[16] L. Raviv, A. Meyer, and S. Lev-Ari, "The role of social network structure in the emergence of linguistic structure," *Cognitive Science*, vol. 44, no. 8, p. e12876, 2020.

[17] S. Goldin-Meadow, *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. Psychology Press, 2005.

[18] M. L. Hall, Y. D. Ahn, R. I. Mayberry, and V. S. Ferreira, "Production and comprehension show divergent constituent order preferences: Evidence from elicited pantomime," *Journal of memory and language*, vol. 81, pp. 16–33, 2015.

[19] D. I. Slobin and T. G. Bever, "Children use canonical sentence schemas: A crosslinguistic study of word order and inflections," *Cognition*, vol. 12, no. 3, pp. 229–265, 1982.

[20] J. M. Meisel, H. Clahsen, and M. Pienemann, "On determining developmental stages in natural second language acquisition," *Studies in second language acquisition*, vol. 3, no. 2, pp. 109–135, 1981.

[21] M. Pienemann and M. Johnston, "A predictive framework of sla," *Manuscript: University of Sydney*, 1987.

[22] V. Cook and V. J. Cook, *Linguistics and second language acquisition*, vol. 12. Macmillan London, 1993.

[23] V. J. Cook, "Timed grammaticality judgements of the head parameter in l2 learning," *The Dynamics of Language Processes: Essays in Honor of Hans W. Dechert, Gunter Narr Verlag, Tübingen*, pp. 15–31, 1994.

[24] B. Byrne and E. Davidson, "On putting the horse before the cart: Exploring conceptual bases of word order via acquisition of a miniature artificial language," *Journal of Memory and Language*, vol. 24, no. 4, pp. 377–389, 1985.

[25] A. Langus and M. Nespor, "Cognitive systems struggling for word order," *Cognitive psychology*, vol. 60, no. 4, pp. 291–318, 2010.

[26] E. Gibson, S. T. Piantadosi, K. Brink, L. Bergen, E. Lim, and R. Saxe, "A noisy-channel account of crosslinguistic word-order variation," *Psychological science*, vol. 24, no. 7, pp. 1079–1088, 2013.

[27] M. L. Hall, V. S. Ferreira, and R. I. Mayberry, "Investigating constituent order change with elicited pantomime: A functional account of svo emergence," *Cognitive Science*, vol. 38, no. 5, pp. 943–972, 2014.

[28] H. Marno, A. Langus, M. Omidbeigi, S. Asaadi, S. Seyed-Allaei, and M. Nespor, "A new perspective on word order preferences: the availability of a lexicon triggers the use of svo word order," *Frontiers in psychology*, vol. 6, p. 1183, 2015.

[29] F. J. Newmeyer, *Language form and language function*. MIT press, 2000.

[30] W. Washabaugh, "On the sociality of creole languages," in *Dialect and Language Variation*, pp. 542–557, Elsevier, 1986.

[31] A. Senghas, E. L. Newport, and T. Supalla, "Argument structure in nicaraguan sign language: The emergence of grammatical devices," in *In E. Hughes & A. Greenhill (Eds.), Proceedings of the 21st annual Boston university Conference on Language Development*, Citeseer, 1997.

[32] W. Sandler, I. Meir, C. Padden, and M. Aronoff, "The emergence of grammar: Systematic structure in a new language," *Proceedings of the National Academy of Sciences*, vol. 102, no. 7, pp. 2661–2665, 2005.

[33] I. Meir, W. Sandler, C. Padden, M. Aronoff, *et al.*, "Emerging sign languages," *Oxford handbook of deaf studies, language, and education*, vol. 2, pp. 267–280, 2010.

[34] S. Goldin-Meadow and C. Mylander, "Spontaneous sign systems created by deaf children in two cultures," *Nature*, vol. 391, no. 6664, pp. 279–281, 1998.

[35] R. Ergin, *Central Taurus Sign Language: a unique vantage point into language emergence*. Tufts University, 2017.

[36] R. Ergin, I. Meir, D. Ilkbaşaran, C. Padden, and R. Jackendoff, "The development of argument structure in central taurus sign language," *Sign Language Studies*, vol. 18, no. 4, pp. 612–639, 2018.

[37] R. Jackendoff, "the autonomy of syntax," *Verb-particle explorations*, vol. 1, p. 67, 2002.

[38] W. Klein and C. Perdue, "The basic variety (or: Couldn't natural languages be much simpler?)," *Second language research*, vol. 13, no. 4, pp. 301–347, 1997.

[39] T. Givón, *Functionalism and grammar*. John Benjamins Publishing, 1995.

[40] M. Schouwstra, K. Smith, and S. Kirby, "The emergence of word order conventions: improvisation, interaction and transmission," 2020.

[41] J. B. Tenenbaum, *A Bayesian framework for concept learning*. PhD thesis, Massachusetts Institute of Technology, 1999.

[42] K. Smith, S. Kirby, and H. Brighton, "Iterated learning: A framework for the emergence of language," *Artificial life*, vol. 9, no. 4, pp. 371–386, 2003.

[43] T. L. Griffiths and M. L. Kalish, "Language evolution by iterated learning with bayesian agents," *Cognitive science*, vol. 31, no. 3, pp. 441–480, 2007.

[44] F. Xu and J. B. Tenenbaum, "Word learning as bayesian inference.," *Psychological review*, vol. 114, no. 2, p. 245, 2007.

[45] S. Kirby, H. Cornish, and K. Smith, "Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language," *Proceedings of the National Academy of Sciences*, vol. 105, no. 31, pp. 10681–10686, 2008.

[46] N. C. Ellis, "Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition," *Studies in second language acquisition*, vol. 24, no. 2, pp. 143–188, 2002.

[47] E. Lieven, "Input and first language acquisition: Evaluating the role of frequency," *Lingua*, vol. 120, no. 11, pp. 2546–2556, 2010.

[48] B. AMBRIDGE, E. KIDD, C. F. ROWLAND, and A. L. THEAKSTON, "The ubiquity of frequency effects in first language acquisition," *Journal of Child Language*, vol. 42, no. 2, p. 239–273, 2015.

[49] S. Kirby, K. Smith, and M. Spike, "Simulating language, academic year 2021-2022." `https://github.com/smkirby/simlang2022`, 2022.

[50] C. L. Hudson Kam and E. L. Newport, "Regularizing unpredictable variation: The roles of adult and child learners in language formation and change," *Language learning and development*, vol. 1, no. 2, pp. 151–195, 2005.

[51] F. Reali and T. L. Griffiths, "The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning," *Cognition*, vol. 111, no. 3, pp. 317–328, 2009.

[52] L. Raviv and B. Thompson, "A short tutorial on agent based modeling in python." `https://github.com/Limor-Raviv/Tutorial_Agent_Based_Models`, 2018.

[53] G. Lupyan and R. Dale, "Language structure is partly determined by social structure," *PLOS ONE*, vol. 5, pp. 1–10, 01 2010.

[54] I. Meir and W. Sandler, "Variation and conventionalization in language emergence," *Eds.), Language contact, continuity and change in the genesis of modern Hebrew*, pp. 337–363, 2019.

[55] I. Meir, A. Israel, W. Sandler, C. Padden, and M. Aronoff, "The influence of community on language structure: Evidence from two young sign languages," *Linguistic Variation*, vol. 12, 12 2012.

[56] A. Kocab, J. Ziegler, and J. Snedeker, "It takes a village: The role of community size in linguistic regularization," *Cognitive Psychology*, vol. 114, p. 101227, 2019.

[57] R. Ergin, L. Raviv, A. Senghas, C. Padden, and W. Sandler, "Community structure affects convergence on uniform word orders: evidence from emerging sign languages," The Evolution of Language Conferences, 2020.

[58] S. Contributor, "What is network topology? best guide to types amp; diagrams," Oct 2022.