

qL

ATLAN

Introduction to a proposed philosophical
international auxiliary language

道可道，非常道。
名可名，非常名。
無名天地之始；
有名萬物之母

ㄣ △ ∇ | ˘ ∅ ㄣ |
♂ △ ∇ ♂ ˘ ∅ ㄣ |
ㄣ ∅ ∇ ♂ ˘ ㄣ ∅ 卜 \ ㄣ 十
ㄣ ∇ ♂ ˘ ㄣ ∅ ！ 卜 ㄣ ㄣ

*The Path that can be trodden,
is not the eternal Path.*

*The Name that can be named,
is not the eternal Name.*

*Nameless, it is the Originator
of heaven and earth
Named, it is the Mother
of ten thousand things.*

-Dao De Jing, chapter 1

Foreword - by Ana Bosnić

A distinct pleasure to write the foreword to this book by Stijn, Niek, Jarno, Jep, Jonathan and Max, students of the Humanities Honors Program at Utrecht University. As this book shows, they have attempted to create a new language (one with the potential for universal communication), thus embarking on a linguistic adventure with challenges, curiosity, and a dash of audacity. Creating a new language is a complex, interdisciplinary endeavor requiring skills, time, effort, dedication, motivation, creativity, and above all – (linguistic) knowledge. It also needs to be a deeper understanding of linguistics, notions and theories, complexities that underpin communication, the cultural diversity that influences our lexicon, and the paradox of simplicity and expressiveness.

The authors have embarked on this journey to create a language trying to develop a universal, neutral, and simple language; a language that should be able connect people across all boundaries, transcend boundaries, and be embraced by diverse communities. Ambitiously, they have attempted to craft a linguistic framework that would not favor any specific group or group, but rather provide a neutral platform

human communication. Needless to say, that this is a titanic task, and the mere act of attempting it is seworthy.

As a constructed language, Atlan benefits from the aid of tools that an organically generated language cannot. Thus, for example, the creators of Atlan have tapped vast reservoirs of linguistic data to inform their decisions, allowing them to create Atlan's phonetic inventory, script and vocabulary. This fusion of human creativity and computational insights laid the foundation for their linguistic invention, offering a boost in their pursuit of universality, unambiguity, and parsimony. It is without saying that this innovative approach to language creation serves as a testament to the boundless possibilities that arise when human creativity converges with computational tools.

In conclusion, this book encapsulates the result of arduous work carried out by Stijn, Niek, Jarno, Jep, Nathan, and Max in their quest to create a universal neutral auxiliary language. Throughout the pages of this book, readers will bear witness to their tireless pursuit to create a language that transcends borders and fosters effective global communication. It is my hope that

ourney inspires further research into the intricacies of language, serves as a reminder that the power of linguistic ingenuity knows no bounds, and shows that learning can be fun!

Can Atlán become a new lingua franca? Well, we will tell.

Ana Bosnić

Contents

Contents	iii
Introduction	1
1.1 Opening Remarks – Niek Elsinga	1
1.1.1 About the project	5
1.2 The story of King Atlas – Stijn Janssens . .	9
1.3 Need for an IAL – Jonathan Roose	13
1.4 Eco’s words – Jonathan Roose	16
1.5 Linguistic relativity – Max Geraedts	18
1.5.1 Language and Culture	32
1.5.2 Culture	34
1.5.3 Language without culture	35
1.5.4 Language with many cultures	36
Phonology – Niek Elsinga	40

TS

The Neurologic Basis of Language	46
Consonants in Atlan	49
Vowels in Atlan	53
Decisions regarding the phonetics: tone and prosody	56
Morphology in Atlan	57
Concluding remarks	60

Writing Atlan Stijn Janssens 62

Writing system – Jarno Smets	62
Direction of reading	76
Punctuation	77
Transliteration	77
Numerals	78
Mathematics	82
Typography	82
Names, loanwords, and cartouches	83
On Dyslexia	87
T _E X and Atlan – Jarno Smets	93

Phonology and syntax 95

Unambiguous Syntax – Jarno Smets	95
Atlan’s grammar	113
Greenberg’s universals	117

CONTENTS

Our Ontology – Stijn Janssens	127
5.1 Parsimony in semantics	128
5.1.1 <i>Oligosynthesis</i>	128
5.1.2 Taxonomy	135
5.2 The anthropocentricity of language	147
5.3 Phenomenological ontology and qualia . .	151
5.3.1 Vision	155
5.3.2 Sound	160
5.3.3 <i>Taste and Scent</i>	165
5.3.4 Physical sensation	167
5.3.5 Emotions	169
5.3.6 Consciousness states	171
5.4 Universal Semantics	172
5.5 Concluding remarks	175
5.6 Appendix	176
Lexicon – Jep Antonisse	181
6.1 Comparison methods and language families	183
6.1.1 Cosine Similarity	183
6.1.2 Language families	189
6.2 Using cognates to generate words	193
6.2.1 Language selection	195
6.2.2 Weights	205

TS

6.2.3 Lexi and her workings	206
Semantic atom vocabulary	211
Protocol for translation	228
Jargon translation	231

Pragmatics – Jonathan Roose 233

Implicatures	238
7.1.1 Prosody	239
7.1.2 Intentions in speech	241
Culture and Language	246
Is language grammar?	250
Conclusion	253

Other suggestions 256

Sign Language – Stijn Janssens	256
Language Variation – Niek Elsinga	257
8.2.1 Language variation and change: inevitable?	261
8.2.2 Mechanics of variation and change	264
8.2.3 Implications for Atlan in language development	266
8.2.4 Conclusion	268

Sample texts 269

CONTENTS

9.1	The Story of Babel – Jonathan Roose	269
9.2	The Epic of Gilgamesh – Niek Elsinga . . .	285
	Exercises – Max Geraedts	291
10.1	Solutions	295
	ilogue	297
	liography	298
	ex	323
	ex	324

Chapter 1

Introduction

Opening Remarks – Niek Elsinga

limits of my language mean the limits of my world.”

Ludwig Wittgenstein

Language, the remarkable construct that binds humanity together, possesses an unparalleled power to shape our thoughts, connect individuals, and cultivate understanding. It is through language that we express our deepest emotions and convey ideas, as well as preserve the vastness of human knowledge. This vast linguistic landscape barriers and bor-

1.1. OPENING REMARKS – NIEK ELSINGA

rise, resulting in imperfect communication and the impediment of the exchange of ideas across cultures and nations.

Philosophically, language can be perceived as more than a mere tool for communication. It shapes our understanding of the world, influences our perspectives, and defines our cultural identities. Language is not merely a means of conveying information, but a reflection of our collective history, aspirations, and values.

Natural languages that have emerged throughout human history presents both a marvel and a challenge. While they showcase the richness and diversity of human expression, they also lead to barriers and misunderstandings amongst speakers and cultures. In an attempt to solve these problems, several languages have been constructed; so-called constructed languages, or “conlangs” as we will see, the language we develop here is an international auxiliary conlang. *International* meaning being as inclusive and as accessible to as much languages in the world as possible. *Auxiliary* meaning that we *don’t* intend our language to replace natural language. Our language should be seen as a tool for communicating clearly internationally.

1. INTRODUCTION

The creation of artificial language is a practice that is as old as humanity, and to this present day, still very much so. Many attempts at a language for international communication have been made in the past. Of course, not all constructed languages are made with the purpose of being international auxiliary languages. Some other constructed languages are made for film, such as Klingon in Star Trek. Others are more personal. The group of people called *Conlangers* is a flourishing community of people who create constructed languages. You might know some of these languages, such as Esperanto. The quest for a constructed international auxiliary language, however, is much older. It has its roots in the early 20th century, when linguists, philosophers, and idealists alike envisioned a language that could serve as a bridge between nations and foster understanding among diverse cultures. Their motivation was grounded in the belief that a language constructed with careful consideration of phonology, grammar, and vocabulary could provide a common ground for international discourse, transcending the boundaries of individual native languages.

This book takes you on a captivating journey through the intricacies of constructing an international auxiliary

1.1. OPENING REMARKS – NIEK ELSINGA

language. It explores the fundamental principles underlying language construction, delves into the complexity of phonological categories, and examines the neurobiological basis of language acquisition and comprehension. Additionally, it investigates the challenges and opportunities presented by the creation of a culturally neutral and inclusive language.

As we embark on this exploration of language and its creation, we invite you to contemplate the immense potential that lies within a constructed language - a language that aspires to be a unifying force, bringing together individuals from diverse backgrounds, fostering global communication, and ultimately transcending the limitations imposed by our native tongues.

Join us on this intellectual odyssey as we delve into the realm of linguistic possibilities, guided by the belief that language, at its core, reflects our shared humanity. Through the creation of a constructed international auxiliary language, we may pave the way for a more inclusive and interconnected world.

1. INTRODUCTION

About the project

This book is authored by Jep Antonisse (artificial intelligence), Niek Elsinga (language and culture studies – logic), Max Geraedts (artificial intelligence), Stijn van der Vlist (philosophy), Jonathan Roose (literature studies), and Arno Smets (language and culture studies – logic). It was written for the Humanities Honours course ‘Reminiscence’ at Utrecht University, under supervision of Dr. Ana Bosnić (linguistics). Our project was to construct a language ‘Atlan’, and write a book about it. From February until June of 2023, we met every week to work on constructing the language, writing literature review essays on the different aspects of the language, programming different tools, and finally putting it all together in this book as the final project. We all enjoyed working on the project, and had many interesting discussions about language, philosophy, literature etc., as well as establishing informal friendships. The language is based on sketches made by Stijn, who had made an earlier attempt at constructing a language that would fit the proposed constraints, but was dissatisfied with the final results. He collected notes and reflections on different aspects that would have to be put

1.1. OPENING REMARKS – NIEK ELSINGA

the language (the writing system and phonology had already been assembled), but after realising the sheer time and ambition required to attempt completing it, he put the project on ice for a few years.

When the project for the Research Seminar was first introduced, Ana gave a short introduction of herself and her work, mentioning the practical application of linguistics seen in constructed languages. Stijn was reminded of the old project he was still intending on finishing some day and realised that with the help of the two AI students, the project would be a lot more achievable, having the power of computation on our side. As though through serendipity, the rest of the group members happened to be standing in close proximity when Stijn pitched the idea to them, generating much enthusiasm from everyone, and thus the project was decided upon the very same day.

Before we proceed, we must mention that the current language being presented in this book should be seen as a first draft of a fundamental proof of concept. It is far from flawless, and we invite the critical enthusiast-reader to contribute to help us add to the language and revise problems or inconsistencies. The main aimed

R 1. INTRODUCTION

of the language is to be a competitor of English as a global language. Therefore, it's main goal is to be better for such a role than English is. It might be finished as a final 'perfect language', partly because language is always alive, but if, we may hope, it is to be more suitable than English, our goal seems to be achieved.

The constructed language 'Atlan' is designed to be an international auxiliary language (IAL) and a philo-language (PhilIAL). It is built along three pre-constraints:

human universality / cultural-linguistic neutrality

ambiguity

elegance / form from function / parsimony

The first constraint covers the goal of the language to be an IAL: a truly unbiased auxiliary language does not give a disproportionate favour of one specific language over any other, as is now very much the case with English being the main IAL (the reason why this book is written in English). It cannot be a mix of a few European languages, like Esperanto for example. Nevertheless, absolute neutrality is impossible because there is

1.1. OPENING REMARKS – NIEK ELSINGA

...true ‘centre’ to different linguistic structures, and the number of different languages and their relative number of speakers will also shift the balance in the total world population (this will be accounted for with the aid of AI, chapter 6.2).

The second constraint overlaps in political and philosophical relevance: a language that is to be learned and commonly spoken by speakers of any language on Earth intended to unify and overcome language barriers, as to ‘undo the confusion of tongues’, and to construct a modern Adamic language’. Therefore, miscommunication and ambiguity should be avoided as much as is historically possible. Within the analytic tradition, philosophy is often regarded as the ‘perfecting of language’ through making statements logically consistent and definitions clearly defined (Stanford, 2022). These concerns whether require Atlan to have an orthography that is phonologically consistent, a lack of homonyms and synonyms that do not add any meaningful nuance and a syntax that does not (easily) allow for grammatically confusing or logically ambiguous statements.

The third constraint is the most ideal and philosophical in nature. ‘Elegance’ here is meant in a similar way

CHAPTER 1. INTRODUCTION

Mathematicians and physicists praise simple and forward formulas that describe and predict a vast phenomena and data. The goal is thus to have as necessary parts as possible; less is more. This is called *parsimony*. This means that any form of communication, be it orthographic, semantic or syntactic, should be an emergent property of the combination of its parts.

The story of King Atlas – Stijn Janssens

We have chosen to name our constructed language ‘Atlas’. It consists of the words ‘AT’, meaning all / everything, and ‘LAN’, meaning speak / talk / language. Therefore, the name can be understood literally to mean ‘Universal Language’. Although the majority of Atlas’s content was generated by an AI programmed on natural language data, the syllables ‘AT’ and ‘LAN’ were consciously assigned their meaning as a symbolic homage to the mythical titan Atlas.

In Greek mythology, Atlas was said to have been condemned by the Gods to uphold the firmament for eternity after having lost in the Titanomachy, an epic battle

1.2. THE STORY OF KING ATLAS – STIJN JANSSENS

between the Titans and the Gods. The Greek poet Pindar located him at the extreme West, at the edge of the known world (which back then mostly referred to the landmasses surrounding the Mediterranean¹ sea). This led him later to be identified with the Atlas Mountains in northern Morocco. This seems to coincide with a folk tale of the local Mauri people, also known as Berbers, present-day Morocco, who to this day still tell of the legendary King Atlas of Mauritania. Because of this, a suggested etymology for the name is the local Tamazight word ‘*ádrār*’, meaning mountain.

According to Greek mythology, he was encountered by the hero Perseus. Upon arrival in Atlas's Kingdom, Perseus asks for shelter, claiming to be the son of Zeus. Atlas refuses, because of a prophecy that once told him that a son of Zeus would come to steal golden apples from his orchard. Because of this, Perseus turned Atlas into a mountain range, with his head at the peak with his feet for hair, and his shoulders as the ridges. Perseus, however, was not the prophesised apple thief. The real thief was rather his grandson and half-brother (thanks to Zeus' incestuous practices), Heracles. When fulfilling

¹From Latin, meaning *middle earth*

R 1. INTRODUCTION

re labours, he was sent to steal some golden apples from Hera's orchard, which was tended by Atlas's daughters, the Hesperides. Atlas and Heracles tricked Hera into carrying the firmament, until Heracles was able to escape with the apples.

Atlas is said to have invented the celestial globe, and perhaps the first having established the science of astronomy. He was supposed to be expert in philosophy, mathematics and astronomy. Perhaps this led to his reputation of carrying the firmament. King Atlas inspired cartographers such as Gerardus Mercator,

and for the Mercator projection of earth, to name his Atlantic Ocean after him. The Atlantic Ocean was named after the titan, as well.

In his late dialogue Timaeus, the philosopher Plato described King Atlas as being the first ruler of Atlantis,

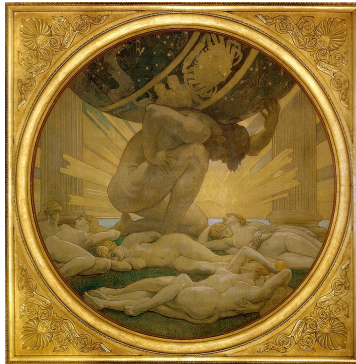


Figure 1: Titan Atlas and the Hesperides, by John Singer Sargent, ca. 1922-1925

1.2. THE STORY OF KING ATLAS – STIJN JANSSENS

city established by Poseidon. Perhaps this city might be referred to a place which is now known as the Richat structure, a geological formation of concentric circles in northern Mauritania just below the Atlas Mountains, matching the description given by Plato. During the purported existence of this city, 12.000 years ago during the African humid period, this area was a lush and fertile land, until the sudden catastrophic global warming event known as the Younger Dryas took place, turning the area into the Sahara Desert we know today. Neolithic artefacts from this era have been found around the Richat structure, as well as fluvial and torrential deposits from the time the Younger Dryas is believed to have taken place. Perhaps this was the origin of the myth of its sudden destruction, having been passed on through oral tradition among North African peoples, until it reached the Egyptian historian Herodotus, who Plato claims to be the source of the story.

We have chosen to name our language after Atlas because of his legendary reputation as being the ruler of a civilisation, a symbol of knowledge, as well as its connotation with philosophy and organised knowledge about the world. It seems appropriate to us to name

R 1. INTRODUCTION

Language Atlan, being somewhat of an encyclopaedic
philical language, after this ancient cross-cultural
presenting wisdom and the bridge between heaven
n.

eed for an IAL – Jonathan Roose

ally the diversity of languages has been both a
and a curse. On the one hand has the variety of
been a database of ways to understand the world
an expression, on the other it has also led to bar-
in- and outgroups. This is why five of my co-
and I have taken up the ambitious task of cre-
o-called International Auxiliary Language (IAL
(), a language that will allow its users to bridge
e barriers and lead to mutual understanding be-
eakers with different mother tongues, a neutral
on which all international communication can

lingua francas of today's world that are used in
onal relations, like French, English or Swahili
archival importance to the language of one par-
group and/or state, these languages are based on

1.3. NEED FOR AN IAL – JONATHAN ROOSE

tical power and historical conditions, they cannot be neutral, they have become international languages because of political interactions and thus are always a political matter. The aim of an IAL is to be a meeting ground for all people without it being dependent on power relations and historical animosities. This project has a lot in common with other IALs, Esperanto, for example, created by L.L. Zamenhof. He hoped that Esperanto would reduce the violence between nations. However, why Esperanto succeeded is also why it is limited. It was made for bringing speakers of European languages together and it is, for a small part, however, only speakers of European languages. Such languages are commonly referred to as 'protoclones'. Our goal with this project is to create a unifying international language for the whole world, which thus not be limited to only a small set of language groups.

The ambition we have with the language Atlan is to create a language that is based on nothing more than a human condition. Later in this book Stijn will explain how we intend to do this however, for now I would like to introduce a term that might help to better understand what we hope to achieve with Atlan. A *tertium*

R 1. INTRODUCTION

tertia is a wish of many translators to have some way to compare the meaning of the original text with the translation. Deriving from the Latin for 'a third person' this term describes the want for a 'perfect language' that could be used to completely translate a text. Of course, the translation is meant to have the same meaning however, some meaning will always be lost. The comparison is to see whether the new meaning has not lost the essence of the original. Whether the totality of the original meaning is captured in the translation. What translators want is a semiotic system that can show the essence of a message in a way that can be compared to all other languages. In Atlan we hope to create a linguistic code, the rules of grammar and syntax of a language, that can function as such a *tertia* by making it based on the essential human experiences. A language that can get to the essence of a message by basing it on the basic human ontological being. The language will function as an IAL that is neutral and universal, it is a language based on the human condition and not on any human experiences.

1.4. ECO'S WORDS – JONATHAN ROOSE

Eco's words – Jonathan Roose

This ambitious project we are indebted to the numerous projects that predate ours with the same or similar aims. Not only is there Zamenhof's Esperanto many other thinkers have dealt with the quest for an IAL. To name all would be too numerous however we can mention a book that has introduced many of the language projects to us. Umberto Eco's book *The Search for the Perfect Language* has been a great source of inspiration for this project. Like Esperanto the book is mostly concerned with Europe. Nonetheless to finish this introduction to Atlan we end with a passage from his book summarising the project:

“Is it possible to reconcile the need for a common language and the need to defend linguistic heritages? Both of these needs reflect the same theoretical contradictions as well as the same practical possibilities. The limits of any international common language are the same as those of the natural languages on which these languages are modelled: all presuppose a principle of translatability. If a universal common language claims for itself the capacity to re-express a text written in any other

R 1. INTRODUCTION

language, it necessarily presumes that, despite the individual genius of any language, and despite the fact that each language constitutes its own rigid and unique way of seeing, organizing and interpreting the world, it is still always possible to translate from one language to another. However, if this is a prerequisite inherent to any universal language, it is at the same time a prerequisite inherent to every natural language. It is possible to translate from a natural language into a universal and artificial one for the same reasons that justify and guarantee the translation from a natural language into another. The intuition that the problem of translation itself presupposes a perfect language is already present in Walter Benjamin: since it is impossible to reproduce all the linguistic meaning of the source language into a target language, one is forced to place one's faith in the convergence of all languages. In each language 'taken as a whole, there is a self-identical thing that is meant, a thing which, nevertheless, is accessible to none of these languages taken individually, but only to that totality of all of their intentions taken as reciprocal and complementary, a totality that we call Pure Language [reine Sprache]'" (Eco 1995:345)

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

Linguistic relativity – Max Geraedts

start I would like to explore linguistic relativity. It is an important term within the study of linguistics, and I would like to explore the possible consequences it has for a universal language. For those of you who are unfamiliar with this term, it refers to the hypothesis that language and thought are intertwined. Benjamin Lee Whorf and Edward Sapir – two linguists – developed about how the structure of a language can influence our thinking. Sapir and Whorf developed two hypotheses about this presumed phenomenon. A strong and weak hypothesis, the strong one argues that language determines thought and that linguistic categories limit and determine cognitive categories. Effectively stating that the language one speaks limits their cognitive abilities. This hypothesis is now disregarded by many modern linguists. The weak hypothesis, however, is still a main point of discussion among linguists. It argues that language influences thought but does not determine it. This weaker position is much easier to accept. A good example of this is the way in which different languages have different conceptions of colors, representations of time and other elements of cognition. So, while it is safe to say that the

R 1. INTRODUCTION

hypothesis is false it is difficult to deny that languages have an influence on our way of thinking. It is our way of representing the world. A difference in language can lead to a difference in our representation of the world.

ideas and views that would eventually go on to be-define linguistic relativity are first found in anthropology. However, it only began to enter mainstream research in the eighteenth and nineteenth century, with German romantic philosophers on the forefront. (German nationalism fuelled the discussion about language relationship with culture and unity at this time. Wilhelm von Humboldt – a Prussian philosopher, linguist and government functionary – stated in 1820:

The diversity of languages is not a diversity of signs and sounds but a diversity of views upon the world (Traband, 2000).

With this movement in Europe, American scientists were discussing this same subject in the early twentieth century. At this time the idea that some languages were superior to others was commonly accepted. It was believed that lesser languages maintained their speakers in a state of mental poverty (Migge, 2007). This caused some

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

American linguists to seek to eradicate Native American languages, they thought that its speakers were savages needed to speak English to become civilized.

The first linguist that began refusing these beliefs was the American Franz Boas, during his studies he became fascinated with the Inuit. After learning their language and culture he began stressing the equal worth of all cultures and languages. There were no such thing as inferior languages according to Boas. Boas' student Edward Sapir went back to the Humboldtian idea that language is vital to understand the unique perception everyone has of our world (Leavitt, 2010). Sapir argued that no two languages could never be perfectly translated to another. This dissonance in language continued in the world view of individuals according to Sapir:

No two languages are ever sufficiently similar to be considered as representing the same social reality. The worlds in which different societies live are distinct worlds, not merely the same world with different labels attached (Sapir, 1929).

This did not however mean that Sapir agreed with the Sapir-Whorf hypothesis, he did in fact disagree with it. Stating

:

R 1. INTRODUCTION

would be naïve to imagine that any analysis of experience is dependent on patterns expressed in language (Sapir, 1946).

It seems that in these middle stages of the development of linguistic relativism views on the subject changed dramatically over the years. As we continue the history, we arrive at the linguist Benjamin Lee Whorf who was one of Sapir's students and has been associated with linguistic relativity more than any other.

One of his best-known examples regards the words the Inuit have for snow compared to the words we have for it in English. This example showed that English could not perfectly translate even simple concepts such as snow between languages. This example was later contested as a misinterpretation by Pullum (1991). Another example of Whorf's linguistic relativity was the time in Hopi. Whorf argued that Hopi did not have countable units of time compared to the SEA – standard European languages – the latter instead regarded time as a single continuous concept. This notion was however also later contested by linguists. In the 1980's Ekkehart Malotki claimed that he had not found any evidence for the claims Whorf

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

made about the Hopi. This refute was then in its turn tested by relativist scholars who criticized Malotki's for forcing the Hopi language into a grammatical model that didn't fit the data (Lee, 1996). How Whorf approached the Hopi is an example of the structure-centered approach. This approach focuses on a structural difference between languages. It then examines the possible sequences and ramifications of this structural difference. The Hopi and the peculiar structure time has in its languages is a prime example of this approach (Lucy, 1997). Whorf died at 44 and left many unpublished papers, these were eventually published in a single volume called *Language, Thought and Reality*. Since neither Sapir nor Whorf had officially formulated a hypothesis Brown and Lenneberg – two influential linguists from the twentieth century – formulated their own in 1954:

"The world is differently experienced and conceived in different linguistic communities" and
(ii) "Language causes a particular cognitive structure" (Brown, 1954).

These were later reformulated by Brown into the *weak* and *strong* formulations:

R 1. INTRODUCTION

Structural differences between language systems will, in general, be paralleled by non-linguistic cognitive differences, of an unspecified sort, in the native speakers of the language. (Weak)

The structure of anyone's native language strongly influences or fully determines the worldview he will acquire as he learns the language. (Strong)

Now we have arrived at the creation of the Sapir-Whorf hypothesis. Which was not created by Sapir nor Whorf. What we have also seen is the difficulty of quantifying linguistic relativity. We came across many bold claims which have all in turn been contested by others. In this reflection we arrive at the last stretch of the development of linguistic relativity. In 1996 the anthology *Linking Linguistic Relativity* was published. It distinguishes linguistic relativity that focuses more on cognitive and social aspects of language. For example, men in the Guugu Yimithirr could give directions based on a compass-like system of north, south, west and east (Gentil, 1998). This shift of focus alongside the development of better means of conducting research ushered in much new research seeking to not only define but also challenge linguistic relativity.

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

Brown and Lenneberg thought that languages described the same objective reality. They decided to research if there was a difference in describing this reality could be proven to have influence on behaviour. For their experiments they decided to focus on the different descriptions of colour in different languages. For one of their first experiments, they tested whether it was easier for English speakers to remember colour shades for which there existed a specific word opposed to shades which were more difficult to describe with words. Later they also compared results between English and Zuni speakers – Zuni classifies green and blue as the same – and it was found that Zuni speakers did have more difficulty making distinctions between shades in the green/blue category (D'Arade, 1995). These studies by Brown and Lenneberg are part of a tradition of investigating linguistic relativity using colour terminology. Real differences could be seen between the perception of colour by an individual and the language they speak. These studies however also received criticism because colour perception is hardwired in the brain. This causes it to be universally restricted to the same factors for all humans (Lucy, 1997). I however have some nuance to add to this argument. While it is

R 1. INTRODUCTION

colour perception is hardwired into our neu-
m, I believe linguistic relativity to be a quale. A
y regarding our experiences, thoughts and inner
. While it is undoubtedly true that colour per-
are *biologically* the same for all of us, I believe
difference lies in our *mental* representation of
ogical phenomenon.

ur research was continued by Berlin and Kay, an
biologist and linguist respectively who are most
own for their research in colour. During their
they found clear universal conventions when it
o colour naming. For example, they found that
different languages have different colour ter-
y, there are universal trends among them. Lan-
who only have three colour terms all have the
ee colours, black, white and red (Berlin, 1969).
colour naming was originally thought to be ran-
s new information was seen as a powerful argu-
inst linguistic relativity (Grumperz, 1996). This
has since in turn been criticised by relativists
Lucy who argued that the conclusions from Berlin
were skewed because they insisted that colour
ly encoded colour. According to Lucy, this made

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

men blind to instances where colour terms contained and provided other information that might be considered as linguistic relativity (Lucy, 1992). As we see and discuss the aspects of linguistic relativity it should become clear that it is a very broad and contested hypothesis.

Advances in cognitive psychology and cognitive linguistics again brought a new wave of studies that focused on linguistic relativity. George Lakoff, for example argued that language is often used metaphorically and that metaphorical use can give us insight in the cognitive structure of language. He gave the example that in the English language time is often likened with money, a lot of metaphors including time talk about it like it can be invested, saved and spent. This cognitive relationship emerging through language can be a sign of linguistic relativity. Especially considering that other languages do not talk about time this way. Other metaphors like up and down that are based on human experience are languages where up is associated with good and down with bad. This association can be seen in many myths and folklores, such as heaven being high up in the skies and hell being down. Lakoff also argued that metaphors play an important role in political debates such as the “right to

R 1. INTRODUCTION

right to choose” (Lakoff, 1980). Lakoff revitalized linguistic relativity not only because of his newly published results, but also because he reappraised linguistic relativity thus rendering past criticisms moot. He did this by concluding that the debate regarding linguistic relativity had been confused. To clear up this confusion he described four parameters on which researchers differed in their opinions on what constitutes as linguistic relativity. These were his four parameters:

The degree and depth of linguistic relativity
Perhaps a few examples of superficial differences in language and associated behavior are enough to demonstrate the existence of linguistic relativity. Alternatively, perhaps very deep differences that permeate the linguistic and cultural system suffice.

Whether conceptual systems are absolute or whether they can evolve

Whether the similarity criterion is translatable or the use of linguistic expressions

Whether the focus of linguistic relativity is on language or in the brain (Lakoff, 1987)

He concluded based on these definitions that past criticisms of linguistic relativity had based their criticism

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

novel definitions of linguistic relativity. According to him this rendered their criticism superficial.

Up to this point we have mostly seen the broad general way linguistic relativity has developed through history. In this last part I want to focus more on some specific cases and thoughts I have about linguistic relativity. Beginning with its influence on constructed languages in literature. Because there are many instances where authors have used language – natural or constructed – in their stories. One of the best examples of this is how George Orwell showed how linguistic relativity might be exploited for political purposes. The authoritarian state in his novel 1984 created a language Newspeak which made it impossible for people to criticize them (just like in real life, Newspeak also has some Oligo-synthetic features: see chapter 5.1.1). Another example is Rand's *Anthem*, a story about a dystopian communist society who erased the word "I" from their language to erase individuality. These examples illustrate not only the possibility of language influencing us but also the fact that we can think about language in this way. The fact that we can imagine it having such an influence on ourselves speaks volumes.

Looking back in history we can see the influence lan-

R 1. INTRODUCTION

as had on us and our actions. Book burning illustrates this perfectly. The earliest occurrence dates to 600 BC. Maybe most famous example coming from the 1930's and 40's when the Nazi's burned countless Jewish books. The Nazi's sought to erase Jewish culture and saw burning their books which were written in their language about their culture to be necessary. Terrible, it does illustrate that language is inseparable from culture. Seeking to eradicate one demands creating the other. Which in turn means that creating one requires creating the other. Linguist and author J.R.R. Tolkien did exactly this when writing stories set in his world, the most famous of those being *The Lord Of The Rings* and *The Hobbit*.

Others have sought to create a language to enable a higher level of cognition. They believe that by speaking a new language humans can reach higher levels of understanding. One of these languages is Loglan and its evolution is called Conlang. This conlang is extremely logic based. They wanted it to be as logical as possible. The creators wanted to test whether linguistic relativity exists. Because the language is entirely based on logic, they thought that it would make its speakers think more logically. Speak-

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

of Lojban reported that they did feel like they thought more logically when speaking Lojban (Nicholas, 2003). This is another example of how language can influence our thoughts in a specifically directed way. Another linguist who sought to do this using her Conlang is Suzette Haden Elgin. She has invented the language Láadan which according to its creator makes it easier to express a female world view. Elgin argued that SEA languages have a male centered world view. Making use of linguistic relativity, she sought to counter this using language. The Toki Pona language was created with the same intent. Its creator – Sonja Lang – wanted to create a simple universal language which focused on happy thoughts. It quite literally aims to make its speakers happier (Lang, Sonja). Because of its simple nature (having only 123 words to begin with), however, it cannot be used to express more detailed or complex meaning: its word for ‘complicated’ is even the same as the word for ‘bad’, *ike*. We once again see how language can have a directed influence on our thoughts. It is not a stretch to pose that we are all confined by our language. It is our way of expressing our thoughts, desires and feelings. The following quote by Von Humboldt illustrates this beautifully:

1. INTRODUCTION

There resides in every language a characteristic worldview . . . By the same act whereby [man] spins language out of himself, he spins himself into it, and every language draws about the people that possesses it a circle whence it is possible to exit only by stepping over at once into the circle of another one (Von Humboldt, 1988).”

Throughout this chapter we have seen the evolution and development of linguistic relativity. We have seen that it is a topic to pin down and reach consensus on. We have also seen that it does have a remarkable impact on our thinking and understanding of the world. From the way we see colours to how we feel. We have seen how we can create languages to infuse its speakers with a specific world view. The power of language is evidently underestimated and we can only guess at the possibilities. Will there be one universal language one hundred years from now? Is one universal language desirable? In one way or another, language has and always will be an integral part of our being. For without it we are nothing.

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

1 Language and Culture

Language and culture have long been inseparable. They influence each other and evolve alongside each other. Culture needs language and language needs culture. Mastering a new language has made this painfully obvious to me. At a certain point you figure out that it is not sufficient to just look up the meaning of a word according to the dictionary. You then use grammar to construct sentences. Language is a lot more intricate; words can mean one thing in each context and only for another context to change its meaning to the complete opposite. Some words are not even in the dictionary. Some words have an entirely different meaning than the one stated in the dictionary. The meaning of some words changes dramatically over time. Some might even say that depending on which language we speak our view of the world can change. Not only this, but language grows over time, it is never in a stable state. Language is our world and culture.

We can see clearly that culture influences our language. When culture changes our language changes with it. A good example of this is Dutch 'straattaal' - literal translation; street language – an unofficial dialect spoken by the youth subculture in the Netherlands. 'Straat-

R 1. INTRODUCTION

sists of mostly normal Dutch words and sentence structure, it has however a few exceptions. It introduces new words and ignores grammar in certain situations. Hereby marking itself different from standard Dutch and dominant Dutch society. This diversion is not an accident. These youths don't want to be a part of the 'adult' or 'adult' society. They seek to define themselves. Creating their own language plays a big role in this. It creates a very strong in-group – people who speak the same language and can communicate with each other – and a very distinct out-group – people who do not speak the same language – this helps in creating subcultures. The existence of a lot of subcultures have their own variations of the language of the dominant culture marks the importance of language in society and the bilateral relation between language and culture.

It makes it difficult to imagine a language without a culture attached to it. This is, however, one of the challenges we have with Atlan. We want to create a language that is as universal as possible. We cannot have one dominant culture associated with Atlan as this would result in it not being useful for people from that culture. In this chapter we will explore by what we really mean when we say

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

universal language and what our vision is of the culture that could be attached to Atlan in the future. Because a language without any culture is impossible.

2 Culture

As you have read in the introduction our goal with Atlan is to create a universal auxiliary language. Not based on one country, culture or region but based on human experiences. I however believe – as implied in the introduction – that a language is impossible to exist without attached culture. I view language as I do the chicken and egg dilemma. It is impossible for one to exist without the other also existing. This conclusion seems like a problem for Atlan. Our goal is to create a universal language but at the same time it is impossible for a language to be without culture. And therefore, it is also impossible for a language to be without biases. I, however, believe to have found two possible solutions to this problem. The first option is to accept that Atlan has no culture and therefore is not a proper language. This might seem like a shocking conclusion, and I will elaborate on this later. The second option is to attempt to create a new culture attached to Atlan. A culture based on human ex-

R 1. INTRODUCTION

S.

language without culture

option I want to discuss is the language without culture option. Seeing how I have stated earlier that a language without culture is impossible you are confused by this option. Let me explain what I mean when I say language without culture. This option originates from the dilemma of making a universal language. For this to be true it cannot depend on a culture. If it did it would not be universal anymore. It is also true that without a culture Atlan cannot be a language. I will not go into detail on the precise conditions something has to satisfy to be a language, but I conclude that having an associated culture is one of the conditions. The result then of the decision to not have any culture associated with Atlan is that Atlan is not a language. Of course it will still satisfy a lot of the conditions of a language. It can be spoken and written, and its purpose is to communicate with other people. It will not be a language like we know them. It will not have a culture. It will not have a country where it is the official language of. It will not have a history. It

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

in some sense be more like a computer language. It not naturally evolve over time it will instead receive updates when deemed necessary.

This might feel like it makes Atlan a very cold and empty thing. Which it does. I think, however, that for the purpose we devised for this language this is a necessary sacrifice. Atlan will be a universal language, used for communication between people who speak different languages. Atlan does not need to be a language as we know it today because it will fulfill a different purpose. It's okay for Atlan to not have its own culture, history, country and people because we already have enough languages who have those things. Atlan will be used as a worldwide communication language; it is allowed to be cold and lifeless. For those languages with identity already exist and will continue to exist in the future. The purpose of Atlan will be to bridge the gap between these cultures. It will be cold and cultureless for everyone; this

4. Language with many cultures
make it an even playing field for all those who speak
This would be the main future I see for Atlan without you strongly disagree with the idea of language without culture.
If you are lucky because I have a second option, language with many cultures. It is impossible for Atlan to have one culture because it would be biased towards

R 1. INTRODUCTION

ure. It would create an 'in' and 'out' group, a fa-
for a supposed universal language. To avoid this
, we could have many cultures associated with
his would create many groups who all have their
ation on Atlan. They can understand each other
will also each have their own identity. This way
n be used to communicate internationally but it
have an identity, culture and history. In fact, it
many different ones. This would create the op-
different countries/cultures to develop their own
of Atlan with which they will build their own his-
a. Of course, these variations cannot be too big
e these different groups will not be able to un-
each other anymore. But apart from this re-
this solution offers a much more alive version
than the previous one.

obvious argument against this option would be
argues for one universal culture that is associ-
h Atlan. This seems like the perfect solution.
ll have an associated culture and it will be a uni-
e. Thus, not excluding anyone and maintaining
ose as a universal language. I don't think this
le unfortunately. Creating one universal cul-

1.5. LINGUISTIC RELATIVITY – MAX GERAEDTS

is a worthy ideal but I am afraid it is not yet possible. As I have said before a culture creates ‘in’ and ‘out’ groups. I believe that culture not only creates these groups it needs them. It originates from them. We can see throughout history that a common enemy brings people closer together. This is also the case for culture. The idea of trying to create a universal group with everyone is impossible. There needs to be some sort of ‘out’ group. The effect of this choice would be very similar to the language with many cultures’ choice. There would be many nations in Atlan, all of them associated with their own culture. It is best to have this view from the very beginning of Atlas’ journey, giving the speakers this freedom rather than having them take it.

I see these as the two possible solutions for the problem I stated in the introduction. It remains a fact that a universal language cannot have a culture. It would not be universal anymore. Having a language without a culture would mean it will not be a language anymore. This creates a problem. I offer two possible solutions for this problem. The first one is to create a language without a culture. The second solution is to create a language with many cultures. These solutions are extremes on the same

R 1. INTRODUCTION

n. I don't know which of these solutions is the
tion. I do think that they both solve the prob-
it in their own way. They would have massively
consequences for Atlan in the future. I look
to seeing how Atlan will develop in our society
ture.

Chapter 2

Phonology – Niek Elsinga

Language is a system through which an individual can communicate with others, which is structured by grammar and vocabulary. Languages are usually spoken, but can also be conveyed by signs as with sign language, or with script. The definition of language is quite a contested topic. Multiple theories about the purpose of language have been proposed. One of the first definitions of languages was put forward by Ferdinand De Saussure. De Saussure saw language as self-contained, self-regulating system, in which the elements are characterised by their relationship with other elements in the system. De Saussure named his vision on linguistics

R 2. PHONOLOGY – NIEK ELSINGA

biology’, but this was posthumously named structure by other linguists (Matthews, 2014).

Today, linguistic scholars deem the structuralist approach outdated, and favour more recent explanations. While some linguistic scholars such as Noam Chomsky and Steven Pinker see language as a biological, formal, mathematical system of signs that are dictated by grammatical rules to convey an utterance (Chomsky, Pinker, 1994), other scholars such as Nicholas Evans take a more ‘functional’ approach and see language as a system of communication that allows for the exchange of information (Evans & Levinson, 2009). One other view sees language primarily and purely as a ‘tool’ that can be used by humans to undertake linguistic behaviour, where language is solely a means of producing and understanding utterances that evolved over the course of human history (Fitch et al., 2005).

It is clear that these definitions more or less convey the same meaning: “a system through which an individual can communicate”. The difference between these views lies not so much in what language is for, but what it encompasses. They are not mutually exclusive to a certain extent. Nonetheless, contemporary scholars predomi-

tly adopt Chomsky's biological approach. However, this view has been contested, on the grounds that neuroscientific studies have found neither biological nor neurologic evidence for the existence of Chomsky's theory on the application of WH-questions, i.e., what, where, when, who(m/se), why, which, and how (Kluender & Kuhl, 1993).

English is still the most spoken language of academia worldwide, and the *lingua franca* of the western world (Lauranen, 2003). It has not, however, gained this position because it is easy to speak or learn. Pronunciation of English vowels, for example, is unlike its graphemic representation, due to phonological shifts of vowels after the standardisation of English spelling in the 15th and 16th centuries (Denham & Lobeck, 2007). English did not gain this position because of the purported absence of cultural influence of English, as stated by Knapp and Meierkord (2002). English fulfils the need of a global lingua franca, and it has spread to large areas of the world due to various factors. These include the adoption of the Latin alphabet worldwide, the invention of the internet and its widespread use in the United States of America.

The development of the American research univer-

R 2. PHONOLOGY – NIEK ELSINGA

subsequent adoption of English as *the* academic language have also been of tremendous importance its widespread use. However, there exist more sinister factors, such as widespread colonization brought about by the British, American cultural hegemony, and the spread of Christianity through western missionaries (Navarro, 2006). The use of English in academic contexts has long been postulated by some to be ‘neutral’ and free of cultural influences (House, 2003).

However, as of late this claim has been challenged. Scholars such as Seidlhofer (2006), and Knapp and Meierkord have claimed that English is imperialistic by default due to the use of English by colonists. These scholars have subsequently decreed that English would be the language to be spoken in countries which do not have English as its endogenous language, and as such the use of English is seen as a form of oppression (Macedo & Bartolomé,

and other scholars have presumed that English can be neutral to a certain degree, and that it is up to the speaker to give partiality to one’s words and actions (Gordon, 1997). If this view is mirrored against the notion of the impartiality of language and that language

culture are interwoven to their very core as famously articulated by Kramsch (2014), it is possible to surmise that any language that has evolved naturally in humans through use and repetition without conscious planning or remediation is intrinsically biased, due to the fact that culture and language are inherently linked (Lyons, 2011).

Atlan is designed to be an auxiliary constructed language, a language that is created with the purpose of facilitating communication between people who have different native languages. This decision has been made because we are of the opinion that a language that is used in academic context should be neutral. This does not imply that the language shall solely be used for academic purposes, nor does it mean that it should replace other languages.

With the creation of the language, multiple goals have been kept in mind. The primary purpose in the creation of the language is to be as culturally neutral as possible, so that no group of people will be especially favoured or disfavoured when learning the language with regards to its similarity to their own. Creating a language from scratch can procure this cornerstone.

R 2. PHONOLOGY – NIEK ELSINGA

her main goal is that the language should both to speak and understand. The notion of unambiguity is another tenet, with the goal of reducing confusion and misinterpretation within communication as much as possible. This means being as sparse as possible, with the fewest elements of the language, where simplicity is preferred and complexity should arise from the combination of basic elements. This is, of course, of utmost importance in phonology and morphology. If a different system is used, it would change the entirety of the language. The same applies to morphology, where the distinction needs to be made between who the actor and who the recipient is.

This paper will serve as an overview regarding the phonological and morphological considerations that have been made for the language. In the first section of this paper, I will elaborate on the neurology concerning speech production. The second section will cover the choices that have been made regarding the phonology for consonants, vowels, and prosody. Finally, I will close this section by summarising what has been stated, and giving concluding remarks.

2.1. THE NEUROLOGIC BASIS OF LANGUAGE

The Neurologic Basis of Language

Neurolinguistics is the study of how the brain produces, comprehends, and acquires language. It combines both the traditional framework of humanities, namely the language as a social phenomenon, with a neuroscience approach. The two traditional areas that are correlated with the production and comprehension of language and speech with respectively Broca's area in the frontal lobe, and Wernicke's area in the temporal lobe (Geschwind, 1972), which are connected through the *fasciculus arcuatus* (Bernal & Altman, 2010). These areas are not bilaterally localized, and solely exist in the left cerebral hemisphere.

The production of speech occurs according to three main principles: conceptualization, formulation, and articulation. In the first stage, conceptualization, an individual with the intention to create speech links the desired concept to the particular spoken words. This preverbal message contains the to-be conveyed thoughts to be expressed. The second stage is formulation, in which the linguistic form for the desired message is formulated. This stage involves knowledge of grammar, phonology, and phonetics applied to the preverbal message. The third stage

R 2. PHONOLOGY – NIEK ELSINGA

articulation of the message, in which motor functions are activated to produce the utterance.

Perception of language or speech begins at the hearing of the sound signal and the process of audition. Subsequently, speech sounds are further processed in order to extract information regarding acoustic cues and phonetic information. This information can then be used for processes that are considered to be 'higher-level' language processes, such as word recognition (Levelt, 1999). These processes are then further processed in the auditory cortex to reach the brain.

Research has indicated that the auditory cortex processes voiceless and voiced phonemes differently in ferrets, which have similar structures in the processing of speech information when compared to humans (Meszner et al., 2008). Phonemes are, put very simply, sounds, the smallest units of speech. Phonemes are usually divided into consonants and vowels (Yallop & Fletcher, 2005). Consonants are created by constricting the airway in the vocal tract when air is forced out of the lungs, which is mostly done by the tongue.

Consonants can also be created by, among other things, the nose and vocal tract. Voiced consonants are con-

2.1. THE NEUROLOGIC BASIS OF LANGUAGE

ants that incorporate the vibration of the vocal cords in the articulation of the letter occurs. Some examples of voiced consonants are the /b/, /d/, and /g/. Voiceless consonants on the other hand do not make use of vocal cords. Examples of voiceless include /p/, /t/, /k/. Some languages, such as Arabic, do not have the voiceless bilabial plosive /p/ in their phonological inventory (Al-Ani, 1970). When a speaker of Arabic wants to say the word 'pizza', they would pronounce it as 'bizza', where the voiced bilabial plosive /b/ is used instead of the voiceless /p/. If an Italian on holiday in an Arabic-speaking country would order a pizza, pronouncing the word with the voiceless bilabial plosive /p/, a monolingual speaker of Arabic would not have any hindrances whatsoever with the comprehension of the utterance (Versteegh, 2014).

This can be linked to another research by Liégeois-Duval et al. (1999) on the inquiry of the perception of voiced and voiceless phonemes. In this research, a speaker produced voiceless and voiced phonemes, with the following vowel being /a/ (/pa/, /ta/, /ka/ for voiceless, /ba/, /da/, /ga/ for voiced) in a random order. Neurologic tests were carried out using a tool called 'electroencephalography' (EEG). An EEG maps where in the

R 2. PHONOLOGY – NIEK ELSINGA

Electrical pulses occur, i.e., where and which areas of the brain are activated when an individual is exposed to a stimulus. The EEG has shown that the auditory cortex is able to process syllables with voiced consonants and syllables with voiceless consonants in the left hemisphere; however, the right hemisphere was not able to make this distinction and solely processed acoustic stimuli. Furthermore, the auditory cortex was not able to distinguish between syllables with voiced consonants and voiceless consonants. The results from the EEG showed no significant differences between syllables with voiced and voiceless consonants. However, a differential coding of voiced and voiceless syllables is preserved. This would mean that an individual is able to distinguish these differences (Liégeois-Chauvel et al., 1999).

Consonants in Atlan

As previously stated, the word 'pizza' would be pronounced differently by according to Arabic phonology (Al-Ani, 1970). The example also states that 'pizza' and 'bizza' would be understood as the same word. This is because in Arabic, the 'b' and 'p' are variants of the same phoneme, a phenomenon called allophony.

2.2. CONSONANTS IN ATLAN

Furthermore, certain languages (or language families) use scripts that do not implicate the voicing of a consonant, such as Tamil. Tamil uses both voiced and unvoiced consonants, however, it is decided by context (e.g., linguistic- register), and not by its script. A consonant being voiced or unvoiced does not imply that a word gets a whole new meaning, but gives meaning to the context of the word. Consonant voicing thus is not contrastive in Tamil (Keane, 2004; Schiffman & Arokianathan, 1986). Regarding the phonology of our language, the decision has been made that both voiceless and voiced consonants are allophones. For example, a speaker of our language would perceive both the voiced bilabial plosive /b/ and the voiceless bilabial plosive /p/ as the same phoneme.

The script is meant to reflect this, as is the case with Tamil. Furthermore, because not every language has the same set of phonemes nor the same number of phonemes, we have decided that nine distinctive categories should be made. The phonemes that belong to each respective category are allophones in our language. The categories were chosen according to mutual intelligibility, proximity according to the consonantal chart of the International Phonetic Alphabet, and manner of articulation.

R 2. PHONOLOGY – NIEK ELSINGA

efoged, 1999). Furthermore, consideration has been given to the frequency of each phoneme and its parent category. Every category contains a phoneme with a high rate of frequency in languages worldwide. To retrieve the information regarding salience of phonemes, the UCLA Phonological Segment Inventory Database (UPSID) and the Phonetics Information Lexicon (PHOIBLE) were used (Maddieson, 1984, Bran & McCloy, 2019). These databases document the frequency of every existing phoneme.

Categories are as follows: The first category contains (bi-)labial plosives [b, p]. The bilabial plosives are found in 98.89% of all languages worldwide according to UPSID. The second category consists of the coronal plosives, i.e., the dental, dento-alveolar, alveolar, and postalveolar plosives, [t, d, t̪, d̪]. The coronal plosives are found in almost every language according to Liberman (1967), however, no exact percentage is given regarding its frequency. The third category contains the velar plosives and dorsal fricatives [k, g, q, ɢ] and [x, ɣ, ʁ]. The velar plosives and fricatives are found in 99.30% of all languages worldwide according to PHOIBLE and UPSID. The fourth category consists solely of the bil-

2.2. CONSONANTS IN ATLAN

al and labiodental nasal [m, ɱ]. According to UPSID, PHOIBLE, and Maddieson(2013a), the bilabial nasal is a phoneme with the highest degree of frequency worldwide, with over 96% of all languages containing it. The category consists of the coronal and dorsal nasals [n, ɲ, ɳ]. No exact percentage is known of the frequency of the non-bilabial nasals, however PHOIBLE states that over 80% of all languages contain a phoneme of this category.

The sixth category is what is called the ‘liquid’ consonants. All trills, laterals, and lateral approximants, as well as the coronal and dorsal flaps and taps [r, ʀ, ɾ, ɽ, ɻ, ɭ, ɮ, ɬ, ɭ, λ, ʟ, ʡ, ɺ] belong to this category. This category is considerable in size, but carefully chosen. Many languages contain one of these consonants, and different phonemes are usually considered allophones if a different phoneme is used, as is the case with /r/ and /l/ in Japanese and Korean (Ladefoged & Maddieson, 1996; Maddieson, 2013b; Takagi & Mann, 1995). No exact percentage is given for the frequency of these liquids.

To the seventh category belong the labial fricatives and labial approximants [ɸ, β, f, v, θ, ð, ʋ, ɹ̥, w]. These phonemes are found in 84.49% of languages worldwide

R 2. PHONOLOGY – NIEK ELSINGA

g to PHOIBLE and UPSID. The eight category
coronal sibilant fricatives [s, z, ʃ, ʒ, ɬ, ɮ, ɕ, ʑ]. Ac-
to UPSID, these phonemes are found in 88.03%
ages worldwide. The ninth category consists of
al consonants [ç, j, j], which according to UP-
PHOIBLE are found in 90% of all languages. A
asi-category was made for glottal and pharyn-
sonants; however, we have decided to give these
es no meaning.

also makes use of the glottal stop [ʔ]. However
nd is not notated in its orthography. Rather, it
s to differentiate two of the same vowels when
ext to one another. For example, ‘KA.AK’ could
sed with ‘KAK’ if there is no pronounced dis-
between the two syllables, therefore the former
e pronounced as ‘KAʔAK’.

vowels in Atlan

sing the vowels was considerably more difficult,
ing that vowels cannot be placed on an axis of
articulation’ and ‘manner of articulation, as is
with consonants (Ladefoged, 1999). Vowels can

2.3. VOWELS IN ATLAN

placed on a spectrum, with one axis from 'close' to 'open', and another from 'front' to 'back'. The close-to-open axis refers to the position of the tongue placed against the roof of the mouth. 'Close' in this context means that the tongue is positioned as close as possible to the roof of the mouth as it can be without creating a constriction, whereas 'open' means that the tongue is positioned as far as possible from the roof of the mouth. The front-back axis refers to the position of the tongue in the mouth. 'Front' in this context means that the tongue is positioned as far forward as possible in the mouth, 'back' means that the tongue is positioned as far backwards as possible in the mouth (Yallop & Fletcher, 2007).

Vowels considered to be close-front include [i] as in the English word 'free' and the Dutch 'vieren', and close-back include [u] as in the Dutch 'voet' and the English 'foot'. Open-front vowels include [a] as in the British English 'hat', and open-back include [ɑ] as in the Dutch 'vader' (Gussenhoven, 1992; Roach, 2004).

Because the quality of vowels is a spectrum and not every vowel exists in every language, a certain degree of allophony exists in vowels as with consonants. In Indonesian, [ɪ] and [ʊ] are allophones of /i/ and /u/, while in

R 2. PHONOLOGY – NIEK ELSINGA

they are contrastive (Gussenhoven, 1992; Soderstrom, 2008).

In our language, five categories of vowels were made. For the consonants, these are based on the saliency of the vowels and its frequency in languages worldwide. For the vowels, this is based on the same tools as for the consonants; UPSID and PHOIBLE. Three of these categories were easily made because most languages contrast high and low vowels. These are, from high to low, [i, u, a], with respectively 92, 88, and 86% occurring in languages worldwide. For the two remaining categories, a substantial lower frequency is noted for [e, o], respectively 61 and 60% of the languages worldwide, according to UPSID and PHOIBLE.

The five categories were chosen because these five vowels are found in every language, and the frequency of the vowels [e, o] were found in roughly the same percentage in language families worldwide, with the exception of (some) Australian languages (Butcher, 2018; McCloy, 2019).

Another extra vowel is used in our constructed language, namely the schwa [ə], but this vowel is not no-

DECISIONS REGARDING THE PHONETICS: TONE AND PROSODY

d. Its function is to differentiate two of the same sonants that occur next to one another, similar to the use of the glottal stop. For example, in spoken At- the difference between 'AK.KA' would be barely distinguishable from 'A.KA', therefore the former would be pronounced 'AKəKA' to retain the distinction.

Decisions regarding the phonetics: tone and prosody

In some languages, tone (i.e., the use of pitch), is a meaning-distinguishing feature. For example, Mandarin is a tonal language and depending on the pitch or variation in pitch, the word 'ma' can have five different meanings, such as horse, mother, scold, or as a marker for a question (Lee et al., 1996). Pitch can be as important as vowels themselves for comprehension of words and grammatical functions. Following on from the previous section, we decided that phonetic properties, such as tone, have no semantic nor pragmatic value. Prosody has no intrinsic value either. Prosody consists of intonation and rhythm. Intonation are the changes in pitch used for, e.g., conveying the speaker's attitudes and emo-

R 2. PHONOLOGY – NIEK ELSINGA

to highlight or focus an expression. The concept of rhythm in language is dubious at best, and is better explained by the notion that the perception of rhythm is based on the language that an individual alludes to, and is thus irrelevant to precisely define rhythm (Bauer & Beneke, 2012). Likewise, lexico-semantic characteristics are conveyed through phonemes, not the intonation or pitch of said phonemes. This is due to the fact that only a select few languages make use of pitch differences for semantic meanings, such as Mandarin, Cantonese, Vietnamese, Yoruba, and Navajo (Bauer & Beneke, 2012; Yip, 2002).

Morphology in Atlan

What is the difference between ‘walking’ and ‘walked’? Both words convey the meaning that something or someone is moving at a regular pace by lifting and setting one foot in turn, never having both feet off the ground at once. However, while they convey the same meaning in movement, saying: “Stijn walking there yes” would be incorrect, as is “Jep is walked there”. The difference here is the suffix: a morpheme added at

2.5. MORPHOLOGY IN ATLAN

end of a word to form a derivative (e.g. -ation, -fy, -ness). A morpheme, and its subject of study called morphology, examine the smallest meaningful units of language, which can be individual words or parts of words (Matthews, 1991). The main goal of morphology is to understand how words are constructed and how they convey meaning. It analyzes the various types of morphological processes, such as affixation, where morphemes are added in, around, before, or after a word, or compounding, where two or more words are combined to form a new one, and inflection, in which the form of a word is altered to indicate grammatical information like tense, number, or gender (Booij, 2007).

Some languages make heavy use of these morphemes and these context-related factors, other languages do not. Languages exist on a continuum in regard to morphology but can be more or less categorised. On one end of the spectrum exist isolating or analytic languages, in which words are composed primarily of individual morphemes that are each distinct and carry a specific meaning, such as Vietnamese (Comrie, 1989). In isolating languages, each morpheme generally corresponds to a specific concept or grammatical function.

R 2. PHONOLOGY – NIEK ELSINGA

At the other end of the spectrum, there are polysynthetic languages, in which words are composed of multiple morphemes that are fused together to express complex ideas and convey a wealth of information within a single word. In polysynthetic languages, a single word can contain a combination of roots, affixes, and grammatical markers, allowing for the expression of entire sentences' worth of information (Baker, 1998). In polysynthetic languages, the process of word formation involves the morphological affixation, compounding, and inflection. An example of a polysynthetic language is Nahuatl, spoken in Mesoamerica before the colonization by Spanish conquistadores (Rolstad, 2001; Suarez, 2001). Somewhere in this continuum, agglutinative languages exist.

In agglutinative languages, morphemes are typically added to the root or stem of a word to express various grammatical features such as tense, aspect, mood, number, and person. Unlike other synthetic languages like Nahuatl, agglutinative languages maintain a one-to-one correspondence between morphemes and specific grammatical functions. Generally, agglutinative languages exhibit a great degree of transparency in their morpholog-

2.6. CONCLUDING REMARKS

systems. The term “transparent” means that the relationship between the morphemes and their meanings is relatively straightforward and predictable. The affixes are typically added in a consistent and regular manner, allowing for clear distinctions between different grammatical features (Durrant, 2013). A wonderful example of an agglutinative language is Turkish (Lewis, 2001).

Atlan, like Turkish, makes use of an agglutinative system for morphemes. One of the core elements of Atlan is that unambiguity is a prerequisite. In choosing an agglutinative system for morphemes, we are of the opinion that this keystone has been achieved.

Concluding remarks

In this paper, we explored the neurologic basis of language and discussed the phonological considerations for the creation of an international auxiliary constructed language. We highlighted the interconnectedness between language and the brain, and the subsequent choices regarding the phonetic, phonological, and morphological elements.

Language is a complex system that is not merely a

R 2. PHONOLOGY – NIEK ELSINGA

communication but a reflection of our culture
tity. While constructing a language that is com-
evoid of bias may be challenging, striving for
y and inclusivity is a worthy endeavour. The
of a neutral and accessible language has the po-
o enhance global communication, foster cultural
e, and promote inclusivity. While language will
arry cultural influences, our efforts to create a
utral language reflect our commitment to open
and mutual understanding in an increasingly
nected world.

Chapter 3

Writing Atlan Stijn Janssens

Writing system – Jarno Smets

TLAN'S writing system is a natural application of our philosophy: start with elementary parts, and every complexity shall be a mere combination of these parts. Our glyphs (as we shall call them) each denote one syllable. This is always the case: they will always stand for the *same* syllable. Unlike English: in the words “tone” and “to”, the “to” is pronounced respectively [tə] and [tə].

That is the rationale behind our writing system; let us dive into the details. As told, Atlan has a set of basic

R 3. WRITING ATLAN STIJN JANSSENS

ey are:

Consonants		Vowels	
Line	In I.P.A.	Line	In I.P.A.
	/t/	⊂	/u/
—	/k/	⊃	/i/
/	/n/	⊔	/a/
\	/m/	⊆	/o/
⌒	/j/	○	/e/
⌑	/s/		
⌐	/f/		
⌑	/l~r/		
∅	/p/		
◦	<nil ¹ >		

The basic lines of Atlan's writing system.

e lines all represent a single vowel (V), or a sin-
onant (C). We can combine them to make sylla-

you see this hollow circle, the other line is combined with
Don't panic if you don't yet understand this; it will be ex-
ortly.

3.1. WRITING SYSTEM – JARNO SMETS

. By combining two consonant lines, you get a CVC-syllable, such as *loj*, *pas* or *mup*. You can also make a VC-syllable, such as *mu*, *po*, or *ji*. The vowels don't have separate lines in a CVC or VC-syllable; instead, the vowel is determined by the position of the two consonant-lines. We will go deeper into that below. First, we give the rules for the order of the consonants and vowels: what determines whether two lines make e.g. *poj* or *jop*, *mu* or *um*? This order is determined by the manner in which the lines combine. There is always a "bigger" line, and a smaller one. Rule of thumb: the bigger line usually is the most vertical of the two. These lines fit inside an imaginary box. The position of the smaller line relative to the bigger line, determines the order of consonants. A visual of this rule helps:

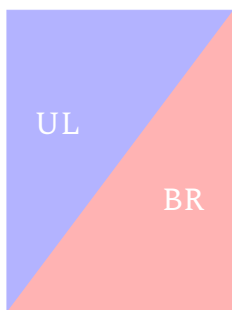


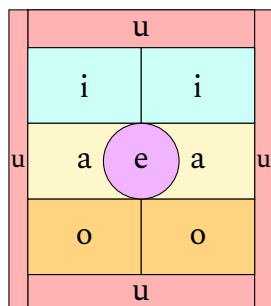
Figure 2: Box for determining consonant order.

R 3. WRITING ATLAN STIJN JANSSENS

smaller line is in the upper-left triangle (UL), the
nt it designates comes first. If it is in the bottom-
e, it comes second. For the rest of the explana-
s advised to keep this box in the back of your
example:

As you see here, the smaller
line is found on top. Hence,
it is placed inside the upper-
left triangle. The consonant
for which the smaller horizontal
line stands (the *k*), comes before
the other consonant, the *s*.

id, the vowel emerges from the position of the
ine in relation to the bigger letter-line. Again, a
ll help:



3.1. WRITING SYSTEM – JARNO SMETS

The vowel is...

- *u* if the smaller line is found at the edges. The smaller line is in its whole above, under, left or right of the main line.
- *i* if the smaller line is found on the upper-left or upper-right side of the main line. It is usually smaller than the line made for *u*, to avoid confusion.
- *a* if the smaller line is found left or right to the middle of the line.
- *o* if the smaller line is found on the bottom-left or bottom-right hand of the main line. Again, this line is smaller than the line for *u*.
- *e* if the smaller line is placed in the middle. Or, if the small line intersects with the main line at the middle. In some instances, the small line is then split up by the main line.

Then we have three exceptions to these rules. The first: you can combine two of the same letter-lines, to make syllables such as *pop*, *mum*, or *lol*. The order of the lines doesn't matter; hence we choose to place the smaller line to the upper-left of the main line in such

R 3. WRITING ATLAN STIJN JANSSENS

For the vowel *u*, there are two small lines, split center. For *e*, there are either two or three small lines, at least one of those lines crosses through the center of the imaginary box.

The second exception has to do with the *k* (/k/). That's the base line – , right? It's a horizontal line. Of that, we have to think of a different box than 2 and 3 to figure out the consonant-order and the vowel-order. The solution is simple: we flip the boxes. They look like this:

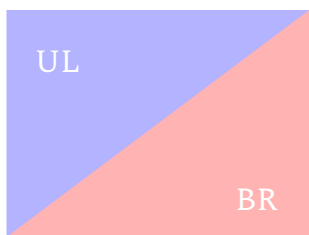
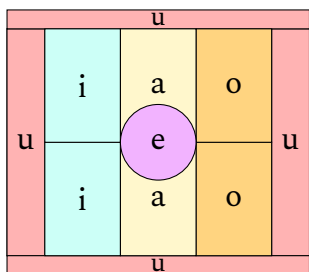


Figure 4: the imaginary vowel-placement box for /k/. Figure

5: the imaginary consonant-order-box for /k/.

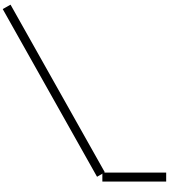
3.1. WRITING SYSTEM – JARNO SMETS

The third exception: Remember that the *p* is represented by the dot •. For clarity, we couldn't combine two dots to make a full syllable. Hence, two *p*'s combine a bit different from the rest of the lines. *p* also can't combine well with the circle (which designated “nothing”). They combine in the following way:

Basic line		u	i	a	o	e
	•	◊	⊘	◌◁	◌▷	⊗
•	•	◊	⊘	◌◁	▷◌	⊗
•		⊘	◊	▷◌	◌◁	⊗

These were the rules for the script of Atlan. It might be a bit cryptic, so let's discuss some examples. If you still feel uncertain whether you understand the rules, go through them again. Personal experience tells that, after some time, recognizing letters gets more intuitive.

Let's dissect this letter. This is the letter “mok”, or (/mɔk/) phonetically. First step is to discover the main line, which is the long diagonal here. This is the consonant *m* (/m/). Then there is a smaller line, found in the



R 3. WRITING ATLAN STIJN JANSSENS

right corner. This is the k ($\langle /k/ \rangle$). The horizontal line is in the bottom-right of our imaginary square. The m comes before the k (see also figure one). The two consonants, now rests the vowel. Feel back at figure two. The smaller line is found bottom-right corner, hence the vowel here is an o . The full syllable is mok ($\langle /mok/ \rangle$).

let's look at another
if you can determine
yourself first. The
is obvious: it's the big
this big curve is a j^2 ($\langle /j/ \rangle$).
The smaller dot is a p ($\langle /p/ \rangle$). The
and inside the quadrant



figure one. Hence, the dot comes first. The dot is
to the left of the centre of our imaginary box.
The vowel here, is the a ($\langle /a/ \rangle$). The full syllable is
($\langle /a/ \rangle$).

you feel if you got the hang of it? If you don't,
I'll do a few more. To spice things up a bit, we'll
syllable with the vowel e .

tip: the curve for j looks alot like the j itself, doesn't it?
more of these similarities in our writing system; they help!

3.1. WRITING SYSTEM – JARNO SMETS

Remember that this vowel
smaller lines be placed in
centre. Alternatively, the
smaller line could intersect the
line, or be split up by it. In
the example, the smaller line is
split up by the bigger line. The



larger line, the l ($/l/$) splits up the line for j ($/j/$). Because
of this, the vowel is e ($/e/$), and the syllable is lej ($/lej/$).



Now the last example. This,
we think, is the best-looking
glyph in our catalogue. What
does it stand for? There aren't
two, three, or four separate lines
here, as should be. Instead,
there is a triangle with a circle

inside. What do we do? Well, remember the p ($/p/$),
which was a dot. And remember that “nothing” also has
a line: the circle. There was an exception for when two
dots combined, or a p -dot and a circle. The exception
was explained a few pages back. If you go there, you en-
counter the same glyph. This syllable is the pe ($/pe/$). A
tip for remembering these glyphs: if you see a glyph with

R 3. WRITING ATLAN STIJN JANSSENS

e and a circle, think of the *p*.

ope the examples have made clear how our writing
m works. This concludes the explanation of our
system for syllables. Upcoming is our writing
or numbers, and for names. Before we get to the
t, a few words of advice for learning the writing

the next pages, a full list of our glyphs is added.
ey are 490 in number; as many glyphs as we have
lables. Don't be intimidated by the list; instead,
it wisely. Look through the list, and try to grasp
pattern of formation. Read the explanation above,
d try to get a feel of how our glyphs are formed.
ain: after some while, you'll have a stronger in-
tion.

y drawing some of the glyphs. It helps for get-
g used to the glyphs. You don't need a ruler to
w them; just make sure they can be distinguished
m each other.

y blank space is the "nothing" we talked about
ove. You can see the circles appear at the rows
those blank spaces. If the half of an entire row

3.1. WRITING SYSTEM – JARNO SMETS

is empty, it means that a combination's reverse is unnecessary. Think of syllables like *lol*, *nun* or *juj*. They are symmetric, so we don't need a full row for them.

in, on the next page is a table containing all our glyphs. The left two columns contain the base lines, placed in order of combination. If you've forgot which base line is for which consonant, return to the table at the beginning of the chapter. It's a great tool for using as a reference. So come back when you need it.

R3. WRITING ATLAN STIJN JANSSENS

$\neg \neg \neg \neg \neg \neg \neg .$ | \circ $\searrow \nearrow \cup \cap \cup$

$\neg \wedge \vee \cap \cup \setminus \cdot \circ$ | $\backslash \checkmark \ell r u$

┐ ㄐ ㄑ ㄒ ㄓ ㄔ ㄕ ㄖ ㄗ ㄘ ㄙ ㄚ ㄛ ㄜ ㄝ ㄞ ㄟ ㄠ ㄡ ㄢ ㄣ ㄤ ㄥ ㄦ ㄧ ㄨ ㄩ ㄣ ㄤ ㄦ

--- + \n / \r \l \g \k \j

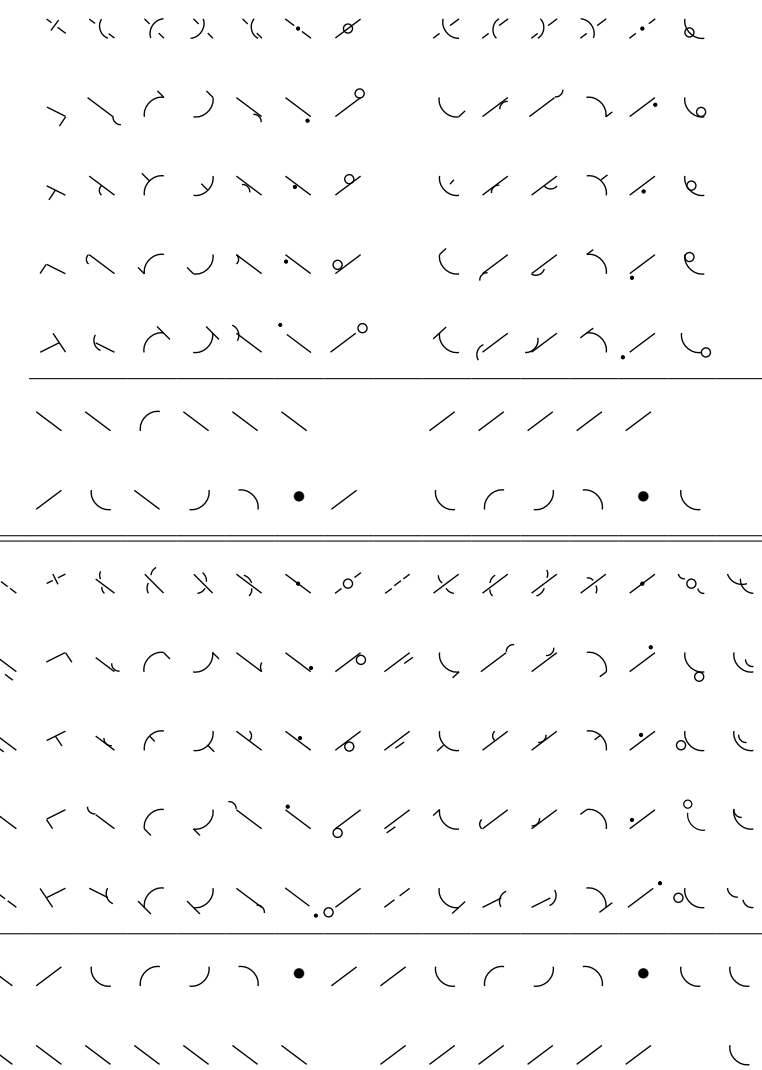
— | → ↘ ↙ ↘ ↙ ↘ ↙ — • 9 | | ↘ ↗ ↘ ↗ ↘ ↗

A collection of 20 geometric symbols and shapes, including various lines, angles, and curves.

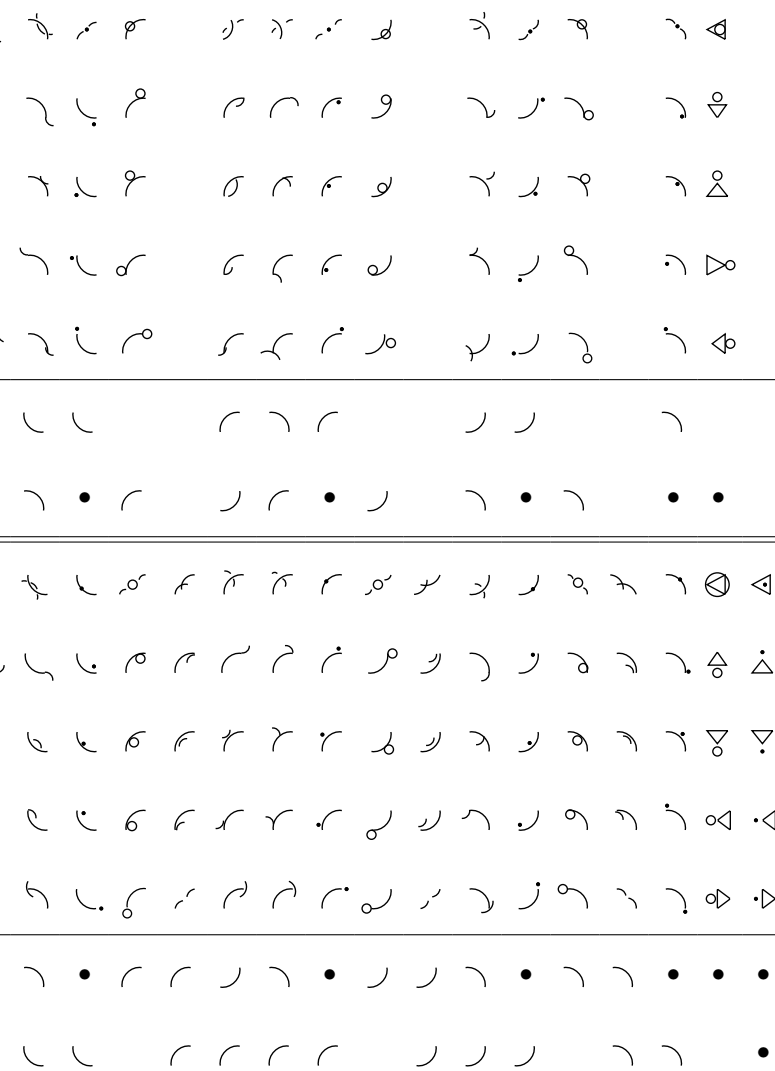
-- + \ / - () ~ = . ! | < / \ | r p

[illegible]

3.1. WRITING SYSTEM – JARNO SMETS



R 3. WRITING ATLAN STIJN JANSSENS



3.2. DIRECTION OF READING

Direction of reading

of the roughly 6.500 languages worldwide, only 12 languages read right to left. The biggest among these is Arabic, with about 1,7 billion speakers. Preferred reading direction has to do with the materials on which the language was historically written, correlating with the technique used to write or carve letters. However, since Latin languages form the majority, and writing materials don't form an obstacle anymore in our modern world, Atlán will use this reading direction as well. Nearly all human languages on earth are written in rows, stacked from top to bottom, and thus Atlán will do the same. This will also make it hospitable to digital environments and graphic design. Because the syllable glyphs can be read all at once, they could also be stacked in vertical columns, reading from top to bottom, for example when employed in calligraphy, or writing along pillars or vertical ridges.

R 3. WRITING ATLAN STIJN JANSSENS

Punctuation

is minimal punctuation, only having dedicated for a comma and a full stop, and spaces are the in any other orthography. A comma is notated as half circle which is open at the top: ~ symbolizing ' continuation of the sentence, and the full stop d at as a half circle which is open at the bottom:

⌒ , symbolizes a closed sentence.

r punctuation will be marked by Atlan's semantics: question sentences start with the interrogative particle E, and so this eliminates the need for a question mark. Exclamative sentences start with the particle exclaiming the need for an exclamation mark. Other words would be '&' being 'AN' ('and'), '%' being 'EP.NO' ('hundred'), ':' being 'I' ('relative clause') etc.

Transliteration

words will be transliterated into the Roman alphabet using the archetype letters U, I, A, O, E, P, T, K, M, N, . Dots are used in between each syllable, in order to prevent confusion about where syllables are broken

3.5. NUMERALS

since this could create ambiguity in meaning. Dots between two of the same consonants (eg. AK.KA) or vowels (e.g. KA.AK) are pronounced as a glottal stop or a, respectively (see chapter 2.2 and 2.3).

Atlan's syllables are all (C)V(C). Some loanwords or names, however, might have two or more consonants in a row within the same syllable. In such cases, the individual letter lines that exceed the CVC limit, will stand on their own next to the syllable glyph. The name 'Stijn', for example, will then become 'S.TEJ.N'.

Numerals

Atlan also has a numeric system distinct from the familiar Arabic-numerals. They look like this:

1	┐	10	┐	100	┐	1000	┐
2	┐	20	┐	200	┐	2000	┐
3	┐	30	┐	300	┐	3000	┐
4	┐	40	┐	400	┐	4000	┐
5	┐	50	┐	500	┐	5000	┐
6	┐	60	┐	600	┐	6000	┐

R 3. WRITING ATLAN STIJN JANSSENS

70 \sqcap —	700 — \sqcap	7000 \sqcap —
80 \nearrow —	800 — \nearrow	8000 \searrow —
90 \searrow —	900 — \searrow	9000 \nearrow —

bottom right corner will be the first order of magnitude (below 10), the upper-right corner the second order (tens), the bottom-left corner the third (hundreds) and the top-left the fourth (thousands). This way, when reading a single numeral, one would read from left to right and then from top to bottom, first the thousands, then the hundreds, then the tens and then the below tens, like for example the number 2023 ($\begin{smallmatrix} \text{---} \\ \text{---} \end{smallmatrix}$). An empty line is equal to zero (---), and having one of the corners empty but with a number attached means that that order of magnitude is zero (such as the third order of magnitude being zero). Decimals can be made by using a comma and a number behind it: in this case the orders of magnitude are flipped: one behind the comma is a tenth, two a hundredth, three a thousandth and four a ten thousandth. For example, 4321,4321 would be $\begin{smallmatrix} \text{---} \\ \text{---} \end{smallmatrix} \sim \begin{smallmatrix} \text{---} \\ \text{---} \end{smallmatrix}$. The added benefit of this numeral system, besides taking less space, is that addition could be more visually

3.5. NUMERALS

itive for some numbers: $\overline{\text{---}}^{\text{I}}$ plus $\overline{\text{---}}^{\text{II}}$ equals $\overline{\text{---}}^{\text{III}}$,
 plus $\overline{\text{---}}^{\text{IV}}$ equals $\overline{\text{---}}^{\text{V}}$, $\overline{\text{---}}^{\text{I}}$ plus $\overline{\text{---}}^{\text{VI}}$ equals $\overline{\text{---}}^{\text{VII}}$.

Atlan's numeral system also allows for a duodecimal
 notation, with the addition of unique numerals for
 and 11:

10 = $\overline{\text{---}}^{\text{X}}$, 120 = $\overline{\text{---}}^{\text{XX}}$, 1440 = $\overline{\text{---}}^{\text{XXX}}$, 17280 = $\overline{\text{---}}^{\text{XXXX}}$

11 = $\overline{\text{---}}^{\text{XI}}$, 132 = $\overline{\text{---}}^{\text{XXI}}$, 1584 = $\overline{\text{---}}^{\text{XXXI}}$, 19008 = $\overline{\text{---}}^{\text{XXXXI}}$

The added benefit of a duodecimal system is that cer-
 fractions divisible by 2, 3, 4 and 6 are more straight-
 ward to calculate by heart, whereas 10 can only be di-
 d easily by 2 and 5. Another benefit is that it is more
 ed for many numerical systems, such as time divided
 60 seconds per 60 minutes per 24 hours, the twelve
 nths and zodiac signs, as well as rotation being di-
 d into 360 degrees.

R 3. WRITING ATLAN STIJN JANSSENS

duodecimal system can easily be counted on a single hand, by using the thumb of the hand as a pointer to the segments of the 3 finger segments of the remaining fingers (see Atlans number syllables system as well: 1, 2 and



Image taken from West (2015).

P' \triangle 'OP' \triangleright 'UP'

ce these all end in P

grouped together on

finger. 4, 5 and 6 are 'IK' $\underline{\quad}$, 'OK' $\underline{\quad}$, 'UK'

owing again the same vowel pattern, but with K,

g them on the second finger. Similarly, 6, 7, and

K' \circ — 'IM' \searrow 'OM' \searrow , and 9, 10, 11 are 'JI'

\searrow 'JU' \circ .

ently, Atlan does not have a standardised sys-

larify beforehand whether decimal or duodeci-

erals are used, other than to spot the usage of

erals $\overline{\quad}$ and $\underline{\quad}$. Frankly, current duodeci-

ems in Arabic notation don't have this either,

uld be easily stated verbally beforehand.

3.6. MATHEMATICS

Mathematics

As with punctuation, mathematical symbols can be approximated by semantic atoms. For example, plus + could be 'AN' \mathcal{A} ('and'), minus - could be 'NE' \neq ('negative'), divided by : could be 'EP' \triangleq ('per'), equals = could be 'ME' \backslash ('equal'). This way, speakers will not be required to learn many new mathematical symbols, rather the glyphs could function as these, as well as carrying their own pronunciation and meaning. More complicated mathematical symbols or notations might need to be formalized and standardized by mathematics, which might require more than one syllable.

Typography

Since Atlan's writing system is comprised of a set of baselines, a great degree of artistic freedom is possible in creating different fonts and calligraphy styles to write a language. Atlan typography should make sure to remain faithful to the specific orientations of the different lines as to not cause confusion between them. Since a glyph contains a minimal amount of lines, usually

R 3. WRITING ATLAN STIJN JANSSENS

ographic ornamentation can be added to glyphs causing much confusion. Here we provide four s of typographic variations on the word ‘Atlan’: New Roman font, a *Comic Sans*-style font, Asian-igraphy and Arabic-style calligraphy.



Figure 1: *Atlan* in different typographic styles.

Names, loanwords, and cartouches

at denote certain names, such as personal names, nes, names for institutions, as well as loanwords y certain culture-specific objects or concepts, like neals, for example, are incorporated into Atlan tic approximations of these words in their source e. Because every possible syllable in Atlan al- s its own designated semantic value, this could nfusion, if the syllables are interpreted as Atlan stead of loanwords. For this reason, names and ds always follow the following structure:

noting category, e.g.:

3.8. NAMES, LOANWORDS, AND CARTOUCHES

person ꞑ (EJ)

place ꞑ (LU)

town ꞑ (TOS)

food ꞑ ꞑ (MAJ.KOS)

- NA ꞑ (name)
- phonetic approximation

Using this formula, ‘Paris’ would become: ‘TOS.NA. LI’ ꞑ ꞑꞑꞑ, and ‘curry’ would become ‘MAJ.KOS. KA.LI’ ꞑ ꞑ ꞑꞑꞑ. If the conversational context makes it explicit enough that the subject matter concerns a loanword, and not a literal Atlan word, the mark-denoting category + NA may be omitted for the sake of fluent speech.

In (formal) typography, a cartouche may be employed to encircle the phonetic approximation in order to enhance intelligibility. Cartouches originate from Ancient Egyptian hieroglyphic orthography, where they were used to encircle the names of pharaohs (see Fig. 1) (Chrisholm, 1991).

R 3. WRITING ATLAN STIJN JANSSENS



uche of Hapshep-
f the few female
Egyptian history.
en referred to an-
y because of a lack
e royal nomencla-
s-Brown, 2010).

The cartouche originates from the hieroglyphic ‘shen’, a stylised loop of a rope, which literally means ‘to encircle’, but had come to symbolise eternal protection. Therefore, it was also believed to have apotropaic powers (warding off evil). The word ‘cartouche’ comes from the French word for ‘paper bullet cartridge’, first applied when Napoleon’s soldiers encountered the frequently occurring glyph and noticed its resemblance to such cartridges (White, 2002). Interestingly, cartouches played an impor-

in deciphering Egyptian hieroglyphics. The Egyptologist Thomas Young, following a sug- from the decipherment scholar Jean-Jacques my, compared the cartouches that appeared on tta Stone to the proper names that appeared in k text (Robinson, 2009). This way, he was able

3.8. NAMES, LOANWORDS, AND CARTOUCHES

decipher the name ‘Ptolomy’ (Young, 1823). If Atlan were ever to become widely used, and inscriptions were made employing cartouches, perhaps the archaeologists of some distant future might employ them to decipher a script in the same manner that Young did.

The 2001 conlang *Toki Pona* by the Canadian linguist John C. Lang also employs cartouches in these contexts, for the same reason that Atlan does so (Lang, 2016). Just as in *Toki Pona*, Atlan’s cartouche shall be a simple oval enclosing the name or loanword, without the additional straight line as used in Egyptian cartouches. One might object that employing a typographic element that derives from a specific cultural tradition, namely that of Ancient Egypt, is in conflict with Atlan’s constraint of cultural neutrality. To this we object with the following justifications: the cartouche serves a practical function in preventing orthographic ambiguity; the Ancient Egyptian culture is currently extinct and thus does not advantage current day Egyptians over any other culture; the historical origin adds an extra layer of symbolism and cultural depth which in no way interferes with practicality, and frankly, we are of the opinion that it looks quite

R 3. WRITING ATLAN STIJN JANSSENS

e will see in the next part, the glyphs can be
the computer using T_EX³. This also is true of
es. A cartouche can be made by simply saying
For example, the T_EX code `\cartouche{Utrecht}`
s **Utrecht**⁴.

n Dyslexia

to 7 percent of the population has some form of
(Peterson & Pennington, May 2012),(Kooij, 2013),
up to 20 percent of the population experiences
degree of its symptoms (National Institutes of Health,
or Atlan's writing system, this means that a mi-
group faces these problems, therefore not inter-
with it's universality / majority constraint. How-
m the perspective of the world population, this
unts to a significant group of people that might
vantaged by a writing system that is hostile to
Atlan's writing system has some aspects that
h and others that work against dyslectics.
in language psychology, different orthographies

³ I'll talk more of T_EX shortly.

⁴ Utrecht = UT-LEK-T

3.9. ON DYSLEXIA

classified by their orthographic depth. Shallow orthographies are mostly phonetical, encoding all the necessary information for pronunciation in a straightforward consistent way. Deep / opaque orthographies are on the other end of the spectrum, often deviating from literal, phonetic spellings, or omitting certain phonemes in written language (Besner & Smith, 1992). Examples of languages with shallow orthographies are Hindi, Finnish and Turkish, while examples of languages with deep orthographies are English, French and Tibetan (whose spelling reform took place 800 years ago).

Ideally, an IAL has an orthography that is as shallow as possible, making it more accessible for dyslectic people as well as optimising learnability for non-dyslexics. In this regard, English is quite a poor IAL candidate. Because of great phonetic shifts and many etymological spellings (see chapter 8.2.0 on language variation), English orthography is highly inconsistent, which can be shown by pushing it to some quite absurd extremes. The word ‘church’ could hypothetically be spelled ‘tolot’, combining the ‘t’ as pronounced in the word ‘picture’ and ‘olo’ as pronounced in ‘colonel’. The word ‘fish’ could be spelled ‘ghoti’, combining the ‘gh’ from ‘enough’, the

R 3. WRITING ATLAN STIJN JANSSENS

‘women’ and the ‘ti’ from ‘nation’. The word could be spelled ‘oed’ by combining ‘o’ as in ‘one’ from ‘hacked’, and ‘why’ could be spelled ‘ho’, the ‘ho’ from ‘choir’. Different spelling proposals have been suggested during the past centuries, however no reforms have been established because of low acceptance (Wolman, 2009).

Because Atlan is spelled phonetically, it might be seen as a shallow orthography. However, because of its nature, vowels are not directly notated but rather by the place of conjunction of consonant lines. More technically speaking, vowels are notated, however they might require more focus and attention to be read. More concerningly, many of Atlan’s glyphs have symmetrical counterparts, or are composed very simply by differing in orientation or place of conjunction. For instance the Roman letters p, q, p and b are often confused. One might expect that such symmetries and similarities might also pose a problem for Atlan. A counterargument could be that the specific locations of line conjunction in Atlan is very regular and straightforward, requiring readers to identify these relative positions in order to differentiate them from one another. Since

3.9. ON DYSLEXIA

Atlán does not have any fluent speakers at the time of writing, its actual effect on dyslexia can only be guessed. Therefore, we shall briefly discuss three example orthographies that share some similarities with Atlán, with respect to dyslexia.

Most research has been done on alphabetic orthographies, most often English, this tendency is also known as *alphabeticism*. Therefore, it is hard to come with concrete data on the effects of other types of writing systems on dyslexia. Hangul, a *featural* writing system used for writing Korean, has many letters that look alike, or are an other's mirror image: ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅈ, ㅊ, ㅋ, ㆁ, ㄷ, ㅌ, ㅍ, ㅍ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅡ, ㅟ, ㅠ. Hangul's orthography functions by combining letters into syllable blocks, therefore resulting in blocks that are sometimes barely distinguishable from one another at first glance, like 반 & 번, and 본 & 분. Surprisingly enough, many Koreans are unaware of the existence of dyslexia. A possible reason for this might be that Koreans read by identifying each syllable block all at once, rather than reading all the different letters within each block. Therefore, native speaking Koreans who learn to read as children might not experience any symptoms of dyslexia, because they learn to recognise many different

R 3. WRITING ATLAN STIJN JANSSENS

block combinations as single images, while allowing the phonetic cues present within the block if they read it right (jreidy17, 2014). This would create a strong association between visual cues, pronunciation and meaning. Atlan might have the same benefit when learning the language at an early age, where children might learn to recognise each glyph as an individual, directly associating its pronunciation with its meaning as information is encoded in the glyph. Since each syllable glyph has a unique meaning assigned to it, this might even result in quasi-logographic reading where people might recognise the word, its meaning and pronunciation from its visual shape. People who learn Atlan at a later age might not have this benefit and struggle more with learning the language, as they might not be used to reading each glyph letter for letter, making the learning process more confusing and confusion more likely.

The Inuktitut language of the Inuit people employs a writing system which, like Atlan's writing system, might be classified as a featural syllabary, or an *abugida*. It was developed by Christian missionaries in the 19th century, and like Atlan, it consists of basic geometric shapes, with different orientations or symmetries in-

3.9. ON DYSLEXIA

ting different vowels, e.g.: \wedge = pi, $>$ = pu, $<$ = pa, \vee =
Research into dyslexia among Inuktitut speakers is
ce, but the small existing body of research suggests
such writing systems are not any more difficult to
n for children than other writing systems, and that
y might even be easier to learn as a first writing sys-
because of the clear parsing of syllables as linguistic
s (Donovan & Tulloch, 2022).

The Amharic language of Ethiopia uses an abugida
ned ‘Fidel’. This writing system requires speakers to
norise many different basic shapes, some of which
very similar to one another, which are then system-
ally modified with small grapheme variations to in-
te different vowels. E.g.: \mathfrak{R} = ts’ä, \mathfrak{R} = ts’a, \mathfrak{R} = ts’o,
dä, \mathfrak{R} = da, \mathfrak{R} = do, \mathfrak{R} = d. Again, research is limited,
indicates that because of the nature of the writing
em, glyph-naming is crucial for all other indicators
teracy (Mekonnen, 2023). When this type of writing
em maximises legibility, this creates a positive feed-
k loop, but when it limits reading, it does the oppo-
making overall legibility harder.

R 3. WRITING ATLAN STIJN JANSSENS

\TeX and Atlan – Jarno Smets

precise when needed, but not forcingly rigid. nitty-gritty is possible, but not demanded. Hence, typesetting language \TeX ⁵ is a great fit for this book, Atlan’s writing system. Here I will quickly guide you through the uses of \TeX in this book.

There are our glyphs. They are hand-programmed using Tikz. It was a strenuous effort, but worthwhile glyphs are high-resolution, scaleable, and they are the part. The usage is straightforward as well. To print any of the Atlan glyphs in \TeX you load in the package `Atlan.sty`. Most glyphs are simply transliterations of syllables, with a backslash in front. E.g. `\mum` prints \mum . Some commands were already occupied, hence the commands are named differently, e.g. `\Atlanpi`, instead of `\pi`. Next up, I plan to make more glyphs available on other typesetting platforms. For our numeral glyphs rely on $\text{\LUA}\text{\TeX}$, a more modern version of \TeX . The command, again, is straightforward. You simply state

⁵modern forms \LaTeX and $\text{\LUA}\text{\TeX}$, as used here.

3.10. T_EX AND ATLAN – JARNO SMETS

`\numbr{<number>}`⁶. An example of the `numbr` command: `\numbr{321}` produces $\overline{321}$. To produce a duodecimal number, you type `\numbrdd{1435}` to get $\overline{1435}_{12}$ ⁷. Then, of course, this book is typeset in T_EX. We could have made it easier for ourselves. But, typesetting in T_EX was worth the effort. We are proud of what we have made; both content- and appearance-wise.

⁶Due to the nature of our numeric system, the biggest number you can fill in decimals, is 9999.

⁷The highest duodecimal number you can fill in, is 20735.

Chapter 4

Morphosyntax

Ambiguous Syntax – Jarno Smets

AMBIGUITY is of all times and places, and natural language is rife with it. *Goal, purple people eater, lies with Mary*; these words and expressions can be interpreted in multiple ways. Some despise ambiguity, while others wallow in it. Whatever one thinks of it, it is a part of natural languages.

In our constructed language, we want to minimize ambiguity. This for the sake of clarity and communication. In this chapter,

In this chapter, I will cover a specific type of ambi-

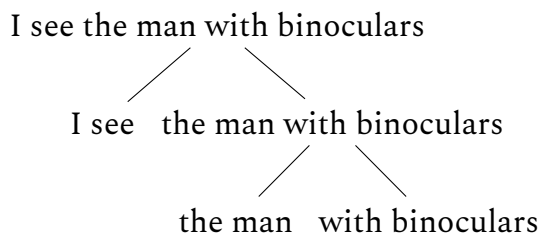
4.1. UNAMBIGUOUS SYNTAX – JARNO SMETS

y, namely *syntactic ambiguity*, also known as *structural ambiguity*. A sentence that can be interpreted in multiple ways due to its syntax, is structurally ambiguous.

My aim in this chapter is twofold. First, I want to show why syntactic ambiguity is a problem, especially in the context of the goals of our project. Then, I will propose a strategy to minimize this form of ambiguity, and argue for a specific strategy choice.

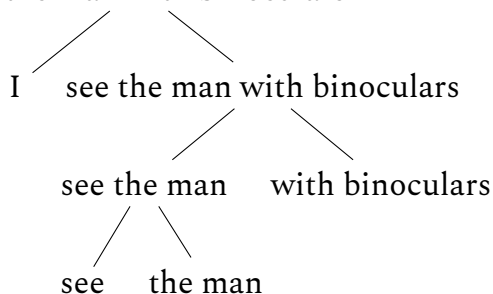
What is syntactic ambiguity?

Syntactic ambiguity occurs when word-order gives rise to multiple interpretations (Oaks,2012, p.16) . The sentence “I see the man with binoculars” could be parsed (i.e. broken down into grammatical parts) in two ways:



R 4. MORPHOSYNTAX

see the man with binoculars



different syntax trees for “*I see the man with binoculars*”

In the above syntax trees, the difference in interpretation hinges on the (de)coupling of the words *man* and *binoculars*. You could make *with binoculars* modify *see*. One could also modify *see* via *with binoculars*. The structure of the sentence doesn’t give preference to either interpretation.

Further illustration: one common type of syntactic ambiguity, is *scope ambiguity*. Scope ambiguity commonly occurs with logical operators such as quantifiers (there exists), negation, and coordinators (and, or, etc.). The *scope* is the part of a sentence over which such an operator ranges. Other in-

are all operators in propositional logic. Quantifiers: \forall = for all, \exists = there exists, \neg = negation, “not”. \wedge = and/but, \vee = or.

4.1. UNAMBIGUOUS SYNTAX – JARNO SMETS

ences of scope ambiguity are seen with modifiers, which I will briefly discuss below. Scope and scope ambiguity can best be explained by example:

(1) *My cat is not grey or black*

Two readings for (1): my cat is neither grey nor black, he is black, for example. Alternatively, my cat is not grey, but black of colour. The scope for negation is ambiguous. The *not* either or it has scope over *grey or black*, or only has scope over *grey*, .

Where lies the origin of such structurally ambiguous sentences? Yang (2014) discerns five major causes of structural ambiguity in English:

- A Negation scope
- B Words with special syntactic functions
- C Improper abbreviation
- D Unclear word-characteristics
- E Unclear modifier-relations

I have just discussed an instance of A above already. With B, I now refer to words that generate *subordinate clauses*;

R 4. MORPHOSYNTAX

ences. These sentences could either be the ob-
e bigger sentence, or be a truly subordinate clause.
girls reported to me when they came. Did the girls
me after they arrived? Or did they report their
arrival? it is unclear, due to the meaning of the
en.

on to cause C Yang mentioned. *Improper abbrev-*
the improper shortening of a sentence. Again,
the sentence *Mary trades cards with Joe.* I could
d *Mary trades cards together with Joe* if I wanted
y that message. But I didn't; I left out the word
making it ambiguous.

, an example will elucidate cause D: *drinking wa-*
afe. Is *drinking* a verb in itself, or part of the
phrase *drinking water*? The word characteristics
ing are unclear. *Drinking* can either be seen as a
as a noun together with *water*.

ly, cause E refers to a modifier. A *modifier* is a
c element that changes the meaning of another
c element. For example, *grey* modifies *dog*. With
modifier relations, it is not apparent which mod-
ifies what. In the phrase *purple people eater*, it
r whether *purple* modifies *people*, or *eater*.

4.1. UNAMBIGUOUS SYNTAX – JARNO SMETS

I propose we bring these causes down to two. Firstly, structural ambiguity is caused by unclear semantic roles. The *semantic role* of a word or sentence-part is the role it plays in the meaning of the sentence. For example, the semantic role of *the grey dog* is the same in both sentences underneath:

- (2) *The cat attacked the grey dog*
The grey dog was attacked

In the example given above, *I see the man with binoculars*, the semantic role of *with binoculars* is indeterminate. Is *with binoculars* how I see the man? Or does the man have binoculars? It is precisely this indeterminacy that seems to generate the ambiguity.

The second cause I propose, is unclear word-grouping or unclear scope. To get rid of the ambiguity in phrases like *purple people eater*, or *lesbian vampire killer*, it needs to be specified which words modify which.

issue for Atlan

In the previous part, I examined syntactic ambiguity. Now, is this a problem for Atlan?. I will here propose the reasons for that goal. First, I will argue that structural ambiguity inhibits the parsing of language by computers. Computer-parsing could boost the spread of At-

R 4. MORPHOSYNTAX

ondly, I will show that some forms of syntactic
y would endanger the communicative function
onstructed language. Atlan should be a bridge
two languages. Syntactic ambiguity can make it
ficult for two speakers from different languages
communicate. Lastly, I will argue that, in some high-
circumstances, syntactic ambiguity could be a great

of all, syntactic ambiguity is a problem for com-
Computers need a so-called *parser* to understand
uage: The machines pick apart a sentence, in
fully understand it (Schubert,2020). Syntactic
y is a true roadblock for such parsing. Because
e ambiguity gives rise to multiple parsing op-
computer can't give a definite parsing of a syn-
ambiguous sentence. To circumvent, or to (par-
ercome it, multiple algorithms have been cre-
it remains a difficult problem (Chowdhary, 2020,

ur constructed language, computer parsing and
ng could be of help to the language learner. Trans-
ould be more accurate, and practice materials
enerated more quickly. The presence of syntac-

4.1. UNAMBIGUOUS SYNTAX – JARNO SMETS

Ambiguity is troubling for computers to analyze natural language.

Besides, structural ambiguity endangers universality. Since our language is intended as an auxiliary constructed language, people learn our constructed language as a *second language*. Hence, learners all approach our language from the perspective of their mother tongues. Now here is the problem: different languages have interpreted scope in different ways.

This has been shown, for example, in Scontras et al. (1997). This team of researchers found out that Mandarin does not have *inverse scope*. Inverse scope can best be explained with an example: “A badger dug every hole”. In English, two readings are available for such sentence:

Surface scope - *There was one badger such that it dug every hole.*

Inverse scope - *For every hole, there was a (different) badger that dug it.*

Scontras et al. found out that the inverse scope reading is simply not available in Mandarin Chinese. Furthermore, they found out this lack of inverse scope is found in the English of native Mandarin speakers. Another

R 4. MORPHOSYNTAX

showed similar results: Korean learners of English strongly preferred the surface-scope reading, and left inverse-scope reading out (Seon & Shin, 2022).

When learning new languages, speakers have the tendency to bring their native scope-reading preferences into the new language. This endangers the communicative function of the language. If our constructed language has certain ambiguities in it, miscommunication can occur. Consider speaker X, in whose language both scope-readings are available. She communicates such a scope-ambiguous sentence to speaker Y. X wants to bring across the surface-scope-reading. To speaker Y, *inverse* scope-readings are *not* available. Then X fails to bring across the intended interpretation of the sentence; a communication failure has occurred. Hence, structural ambiguity endangers the communicative clarity of our constructed language.

Building further on communicative clarity: some environments strictly demand that there be no ambiguity. Hazardous environments, such as nuclear power plants, weapon factories and the like, should communicate in a clear, unambiguous manner. Also law practice should be ridden free of ambiguity. These are high-stake-environments. Any

4.1. UNAMBIGUOUS SYNTAX – JARNO SMETS

communication mistake could have far-reaching consequences.

Say an English nuclear-power plant has the following instructions etched into an important control panel:

In case of emergency: pull the horizontal striped lever

Now, there are two levers in the control room. One is a vertical lever that you pull from north to south, and it is marked with horizontal stripes. The other lever is horizontal, but has vertical stripes instead. Which lever do you pull? I hope this example makes it clear how dangerous syntactic ambiguity can be.

Of course, this was a fabricated example. A real-life example, can be found in (Layman,1962):

Serbian subjects in the United States, shall enjoy the rights which the ... laws grant ... to the subjects of the most favoured nation.

Example (4) elicits two interpretations: Serbian subjects who reside already in the United States enjoy the rights, or Serbian subjects, independent of where they remain, enjoy the rights when they are in the United States. This illustrates syntactic ambiguity in law. Here it can have grave consequences for a large number of citizens.

R 4. MORPHOSYNTAX

these few examples, I have shown why syntactic ambiguity is best left out in our constructed language. It would make it hard for computers to parse our language. That while computers generally help to spread the word faster. Secondly, structural ambiguity in a language can cause miscommunication within a language. Any language allows multiple scope readings, for example. Scope ambiguity can then lead to miscommunication in a language. Thirdly, syntactic ambiguity can be a real danger. It could cause communicative issues in high-stakes environments, such as infrastructure and

It must be noted, however, that syntactic ambiguity is not only a *bad* phenomenon. It can also serve poetic and other artist endeavours. For example, the structurally ambiguous sentence

Time flies like an arrow; fruit flies like a banana

is remarkably witty². Does the fruit fly similar to a banana? Or do fruit flies love a banana? The first part of (5) is designed to prime the reader for the first reading.

Minimizing syntactic ambiguity

² in (Cornish-Bowden, 2015).

4.1. UNAMBIGUOUS SYNTAX – JARNO SMETS

Now I will look at the efforts of other constructed languages to minimize syntactic ambiguities. I will examine the benefits and downfalls of their approaches. From this examination, I will aim to distill the strategy for our constructed language to bring structural ambiguity to a minimum.

One of the main origins of structural ambiguity is the distance between sentence-parts. In a structurally ambiguous sentence, it becomes unclear how the words are grouped into phrases, and then how phrases fit in a sentence. For example, in the noun-phrase *purple people eater*, does *purple* belong to *people*, or to *eater*? Solving structural ambiguity is then making clear which words modify what so that only give one interpretation of a phrase or sentence. The constructed language Lojban³ indeed does this. It has two ways of specifying which words belong together. The first manner comes in the form of the structure word *bo*. *Bo* enforces scope (The Lojban Reference Grammar, 2023). To see how, let's take the English sentence "That is a big bug catcher". In English, you could

³Lojban [loʒban] is a constructed language, created by a group of people wanting to improve another constructed language, *Loglan*. One of its spear points is having an ambiguous syntactic structure. Find out on: <https://mw.lojban.org/papri/Lojban>, may 23rd, 2023.

R 4. MORPHOSYNTAX

this either as a big catcher of bugs, or a catcher of big bugs. In Lojban, the word *bo* makes this difference

- (6) That is a bug-catcher that is big.

*Ta barda miptera bo kavbu*⁴.

- (7) That is a catcher of big bugs.

Ta barda bo miptera kavbu.

might have guessed from the above examples, the word *bo* “pulls” two words together, to combine them. Since the combination of words is made explicit, the ambiguity is resolved.

There is a second way of coupling words in Lojban. The designers of Lojban decided to make rules for groupings, so-called *brivla*. *Brivla* is an umbrella term for nouns, verbs, adjectives and adverbs (The Lojban Reference Grammar, 2023). The *left-grouping-rule* states that the leftmost *brivla* are grouped together. So, the sentence *ta barda miptera kavbu*, is automatically parsed equivalent to the second reading above (The Lojban Reference Grammar, 2023).

⁴That is”, *barda* = “big”, *miptera* = “bug”, *kavbu* = “catcher”, the structure word. English translation found in (Jbovlaste: Dictionary, 2023)

4.1. UNAMBIGUOUS SYNTAX – JARNO SMETS

It seems Lojban got structural ambiguity under control with these two restraints. What are the advantages and disadvantages of this approach?

As already mentioned above, the word-groupings are made explicit, effectively removing structural ambiguity from the language. This increases the clarity of Lojban, and thereby makes the language more universal. There are some downsides however. As we saw above, some possible readings are not even available in the mother-tongue of some speakers. The left-grouping rule described above would enforce a reading upon the language learner, which the language learner is far from familiar with. Lojban might sometimes give rise to miscommunications. Another constructed language with the intent of minimizing (syntactic) ambiguity, is *Ithkuil*. Ithkuil marks semantic roles explicitly in noun cases (Ithkuil, Case Morphology, 2023). This is relatively similar to German, where the case *der* usually marks the (male) subject of the sentence, or *des* marks the possessor. Ithkuil has more cases, including the ones we all know (subject, object, possessive, dative). Examples are *instrument*, *force*, *agent*, and *much* more⁵.

⁵Readers interested in more should visit Ithkuil's website:

R 4. MORPHOSYNTAX

il specifies the exact case of every noun. Due to clear which word plays what role in a sentence. *people eater*, for example, *eater* could be nominative *purple people* would be marked as accusative. way, ambiguity is brought down to a minimum. ;, there is one big downside to this approach: complex. Ithkuil is very complex, and hard to even the creator, John Quijada, can't speak it fluently (Quijada, 2023). Thus, the ubiquitous presence of cases do more harm than good; it eliminates ambiguity at the cost of learning-ease and fluency.

have seen how Ithkuil and Lojban deal with syntactic ambiguity. Taking this in account, how will Atlantean deal with it?

structure of Lojban was the explicit word-coupling structure word *bo* . The word directly made that words formed a separate noun-phrase. However, it is an extra word to remember. We believe it is a good idea to couple words in the most direct sense possible: literally connect them to each other. This is a unique feature of, for example, Dutch: *grijze hondentemmer* (grey hound-tamer) versus *grijze-honden temmer*

4.1. UNAMBIGUOUS SYNTAX – JARNO SMETS

Both in English and Dutch, the words “dog ”and “her ”are joined to indicate that they belong together. In speech, words that should be separated, are separated by a pause.

Now, what about scope ambiguity? For negation, for example, we will include two types: sentential and predicate negation. Sentential negation is a form of negation that spans over a whole sentence. For this we put NE in front of the sentence. E.g. *I have **not** been to school today.*

Predicate negation on the other hand, only spans over a predicate. For this we put NE in front of the predicate (or noun). For example, *I’m very **un**happy at the moment.* This would fix negation scope ambiguity. Take the mentioned example *my cat is not grey or black.* The two readings can be separated using the distinction between types of negation:

(8) *My cat is ungrey or black*

it is not the case that my cat is grey or black

Sentential negation will take the form of a distinct particle, whereas the predicate negation will be an affix. This has the following reasons. Sentential negation spans over a whole sentence. To make it immediately

R 4. MORPHOSYNTAX

that a sentence is negated, it would be convenient to have a loose particle to place at the beginning of the sentence. Predicate-negation occurs within a sentence and binds to predicates. Hence, it will be an operator projected to the predicate it negates.

Our approach to negation doesn't make it more difficult to learn. Most languages are familiar with it: the common types of negation are negative particles, adverbs (Martin et al., 2005, p. 454). Even if, for a native mother-tongue, there is a mismatch between meaning (sentential and negation) and form (particle or affix), the forms are very likely familiar. This likely makes our approach to negation somewhat intuitive for a language learner. Moreover, predicate-negation is present in a majority of languages (Martin et al., 2005, p.467).

What about scope ambiguity outside of negation? Consider the sentence *The dog and the cat and the bird made a mess*. Here, we borrow operator strength from Classical Logic. Negation comes first. Then comes conjunction ("and"). Last comes disjunction ("or") (O'Donnell et al., 2007, p.120)⁶.

⁶What about the conditional ("if...then", \rightarrow) and the biconditional ("if and only if", \leftrightarrow). As far as I can tell, they don't seem to have syntactic ambiguity, hence I leave them unmentioned.

4.1. UNAMBIGUOUS SYNTAX – JARNO SMETS

In the above example, the sentence is read as: (the dog or cat) and (the bird) made a mess. That the bird made a mess, is certain. Whether the dog or the cat made a mess is uncertain.

Now it is worth noting a few *caveats* about my approach. Firstly, I reasoned mostly from syntactic ambiguities in English and Dutch. This could leave room in solutions for syntactic ambiguities not thought of by me. Hence, I talked primarily of *minimizing* syntactic ambiguity. Besides, it is worth noting that context will disambiguate as well. I have mostly examined structurally ambiguous phrases and sentences in isolation. Some of these phrases or sentences would not be as ambiguous in context.

In this chapter, I have shown two things. First, I argued that syntactic ambiguity should be avoided when constructing a language. This because syntactic ambiguity troubles computers, endangers communicative function, and can be potentially harmful.

Secondly, I have proposed several general recommendations for battling syntactic ambiguity. This I distilled from previous attempts at constructing structurally unambiguous languages, such as Lojbans and Ithkuil. Lo-

R 4. MORPHOSYNTAX

de its structure clear, but had a redundant syn-
Ithkuil explicitly specified the semantic role of
d, but became extremely hard to learn and speak
sequence.

won't be as specific as Ithkuil or Lojban. It is a
we need to find between preciseness and learn-
Both Ithkuil and Lojban are extremely precise,
fice learnability. I am confident that Atlan will
od balance, and that the learner will profit from

Atlan's grammar

grammar has the challenge of steering a middle
between minimalism of complexity, yet simulta-
allowing for unambiguity. It tries to be mini-
descriptive in its structure, allowing for more free-
individual and cultural expression while remain-
ligible. Atlan will do this in the following way:
mmatical function that can be expressed within
age, has its own unique assigned syllable. Verbs
onjugated in complicated arbitrary tables, and
e not endlessly modified by cases, but rather

4.2. ATLAN'S GRAMMAR

Specific grammatical functions are conjoined together, as legoblocks, in an entirely regular way. This allows a lot of freedom in choosing specific grammatical functions without having to know the details of grammar. The grammatical markers are added in the order in which they are listed in the word list provided in chapter 6.2.

Atlan's word order is both SVO (subject – verb – object, I eat fruit) and SOV (subject – object – verb, I fruit eat). This means that in every case, the first word of a sentence (apart from mood markers such as interrogative or exclamative) is the subject of the sentence. From here, the speaker is free to choose either the verb or the object to follow, depending on, for example, highlighting words, concept constructions, stream of consciousness &c. According to Kemmerer (2012), the total amount of SOV and SVO dominant languages, or in other words, languages that always put the subject first, amounts to 89% of all languages on earth. However, most languages still allow a basic degree of freedom in word order, the dominant word orders are merely tendencies, not hard rules. Therefore, having the flexibility of SVO and SOV ensures that most people on earth will be intuitively capable of formulating sentences in Atlan.

R 4. MORPHOSYNTAX

rent cases can be marked by adding their corresponding syllables as prefixed to the designated word. Subject is marked with the accusative marker 'EK' , objects with the verb marker 'TU' ᵀ , possessives with the genitive 'TA' ᵇ &c. Plural is always marked at the end of the word, as the only exception.

Verbs can be given tense, aspect and mood. Unmarked verbs are always present tense or infinitive, depending on whether it has a subject. A word can be made past by adding the prefix 'PA' ᵀᵀ , future by adding 'FE' ᵀᵀᵀ , conditional by adding 'PO' , passive by adding 'PI' &c. For a complete list of grammatical markers, see the list in chapter 6.2.

Relative clauses, in which something is said of something, such as 'the fruit is sweet', are marked with the predicate marker 'FUT' ᵀᵀᵀ . Where the noun (fruit) comes before, and the predicate (is sweet) would come after. This would make 'FUT' ᵀᵀᵀ ᵀᵀᵀ . An adjective can also predicate something of a noun, meaning that the very same construction, without spaces, can create 'fruit which is sweet', which can be reformulated as 'sweet fruit'. Since sweet describes the fruit, it is placed behind the word for fruit,

4.2. ATLAN'S GRAMMAR

...e this is the basic rule of thumb for word hierarchy
...atlan.

Because usually, Atlan words are interpreted literally,
...aphoric speech may be indicated by the prefix 'MU'

Gender is not marked obligatory; purely gender-neutral
...uage is entirely possible, and very straightforward in
...n. If the speaker still desires gendered language, the
...icles for 'masculine', 'MA' ʘ or 'feminine', 'FI'
...can be added.

Atlan has three separate markers for so called 'de-
...es of removedness from speaker'. This means that
...first degree refers to the here and now of the person
...ring the language: the first person 'I' 'EJ.AM' ʘ, the
...e 'here' 'LU.AM' ʘ ʘ, the demonstrative 'this'
...AM ʘ ʘ, the time 'now' 'JA.AM' ʘ ʘ. The
...ond degree is the second person, once removed from
...speaker: 'you' 'EJ.UN' ʘ ʘ, 'there' 'LU.UN'
...ʘ, 'that' 'ES.UN' ʘ ʘ, 'then' 'JA.UN' ʘ ʘ,
...the third degree is 'them' 'EJ.AJ' ʘ ʘ, 'yonder'
...AJ' ʘ ʘ. The demonstrative 'ES' ʘ without

R 4. MORPHOSYNTAX

or removedness can be understood to be equivalent to 'it'.

ly, Atlan uses a scale degree of 'negative' - 'neutro-positive'. These markers can be added as pre-words to create relative terms, such as cold - body temperature - warm. The possibilities with expression are endless as you can combine many words and functions together, allowing for the on of thought that might go beyond the lexical y of natural languages.

Greenberg's universals

merican linguist Joseph Greenberg (1963) compiled a set of cross-linguistic grammatical principles, being an IAL, should ideally comply with as many of these universals as possible, such that its grammar is as intuitive as possible to as much people as possible. Below is the full list of Greenberg's universals. If a principle is indicated with a plus '+', this means that Atlan follows the principle. If it is indicated with a tilde '~' this means that it does not apply to Atlan's structure, but therefore also does not break any universal. If it is indicated with a mi-

4.3. GREENBERG'S UNIVERSALS

'-', however, this means that Atlan does not follow principle, while it would have to apply. Only 4 out of the total 45 universals are not obeyed by Atlan, and 41 do not apply. This means that Atlan complies with Greenberg's universals to a satisfying degree, and in the cases in which it doesn't comply, this is for the sake of consistency and simplicity of its rules.

ology

- . + "In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object."
- . + "In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes."
- . ~ "Languages with dominant VSO order are always prepositional."
- . - "With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional."

R 4. MORPHOSYNTAX

If a language has dominant SOV order and the relative clause follows the governing noun, then the adverbial likewise follows the noun.”

All languages with dominant VSO order have OSV as an alternative or as the only alternative basic order.”

If in a language with dominant SOV order, there is no alternative basic order, or only OSV as the alternative, then all adverbial modifiers of the verb likewise precede the verb. (This is the 'rigid' subcategory of III.)”

When a yes-no question is differentiated from the corresponding assertion by an intonational pattern, the distinctive intonational features of each of these patterns are reckoned from the end of the sentence rather than from the beginning.”

With well more than chance frequency, when question particles or affixes are specified in position by reference to the sentence as a whole, if ini-

4.3. GREENBERG'S UNIVERSALS

tial, such elements are found in prepositional languages, and, if final, in postpositional.”

- . + “Question particles or affixes, when specified in position by reference to a particular word in the
- . + “Particles do not occur in languages with dominant order VSO.”
- . ~ “Inversion of statement order so that verb precedes subject occurs only in languages where the question word or phrase is normally initial. This same inversion occurs in yes-no questions only if it also occurs in interrogative word questions.”
- . ~ “If a language has dominant order VSO in declarative sentences, it always puts interrogative words or phrases first in interrogative word questions; if it has dominant order SOV in declarative sentences, there is never such an invariant rule.”
- . + “If the nominal object always precedes the verb, then verb forms subordinate to the main verb also precede it.”
- . + “In conditional statements, the conditional clause

R 4. MORPHOSYNTAX

cedes the conclusion as the normal order in all languages.”

In expressions of volition and purpose, a subordinate verbal form always follows the main verb in the normal order except in those languages in which the nominal object always precedes the verb.”

In languages with dominant order VSO, an inflected auxiliary always precedes the main verb. In languages with dominant order SOV, an inflected auxiliary always follows the main verb.”

With overwhelmingly more than chance frequency, languages with dominant order VSO have the adjective after the noun.”

When the descriptive adjective precedes the noun, demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise.”

When the general rule is that the descriptive adjective follows, there may be a minority of adjectives which usually precede, but when the general rule is that descriptive adjectives precede, there are exceptions.”

4.3. GREENBERG'S UNIVERSALS

- . + “When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite.”
- . - “If some or all adverbs follow the adjective they modify, then the language is one in which the qualifying adjective follows the noun and the verb precedes its nominal object as the dominant order.”
- . + “If in comparisons of superiority the only order, or one of the alternative orders, is standard-marker-adjective, then the language is postpositional. With overwhelmingly more than chance frequency if the only order is adjective-marker-standard, the language is prepositional.”
- . ~ “If in apposition the proper noun usually precedes the common noun, then the language is one in which the governing noun precedes its dependent genitive. With much better than chance frequency, if the common noun usually precedes the proper noun, the dependent genitive precedes its governing noun.”

R 4. MORPHOSYNTAX

If the relative expression precedes the noun either as the only construction or as an alternate construction, either the language is postpositional, or the adjective precedes the noun or both.”

If the pronominal object follows the verb, so does the nominal object.”

ogy

If a language has discontinuous affixes, it always has either prefixing or suffixing or both.”

If a language is exclusively suffixing, it is postpositional; if it is exclusively prefixing, it is prepositional.”

If both the derivation and inflection follow the root, or they both precede the root, the derivation always comes between the root and the inflection.”

If a language has inflection, it always has derivation.”

If the verb has categories of person-number or if it has categories of gender, it always has tense-aspect categories.”

4.3. GREENBERG'S UNIVERSALS

- . ~ “If either the subject or object noun agrees with the verb in gender, then the adjective always agrees with the noun in gender.”
- . ~ “Whenever the verb agrees with a nominal subject or nominal object in gender, it also agrees in number.”
- . ~ “When number agreement between the noun and verb is suspended and the rule is based on order, the case is always one in which the verb precedes and the verb is in the singular.”
- . ~ “No language has a trial number unless it has a dual. No language has a dual unless it has a plural.”
- . + “There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed only by zero.”
- . + “If a language has the category of gender, it always has the category of number.”

R 4. MORPHOSYNTAX

A language never has more gender categories in plural numbers than in the singular.”

Where there is a case system, the only case which never has only zero allomorphs is the one which includes among its meanings that of the subject of an intransitive verb.”

Where morphemes of both number and case are present and both follow or both precede the noun base, the expression of number almost always comes between the noun base and the expression of case.”

When the adjective follows the noun, the adjective expresses all the inflectional categories of the noun. In such cases the noun may lack overt expression of one or all of these categories.”

If in a language the verb follows both the nominal subject and nominal object as the dominant order, the language almost always has a case system.”

All languages have pronominal categories involving at least three persons and two numbers.”

4.3. GREENBERG'S UNIVERSALS

- . ~ “If a language has gender categories in the noun, it has gender categories in the pronoun.”
- . + “If a language has gender distinctions in the first person, it always has gender distinctions in the second or third person, or in both.”
- . + “If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also.”

Chapter 5

Ontology – Stijn Janssens

The goal of this is to expound on the ontology of the language, which concerns its semantics and syntax-inventory. I will achieve this by discussing the evidence on various attempts and classifications proposed within the topics, and then dividing all linguistic phenomena into its irreducible, unambiguous and unique components, as to respect constraint 2 (unambiguous). I will use some principles of natural bifurcation (5.1.2), and built off of the universal substrata-phenomenology and qualia (5.2 and 5.3). It will be linguistically grounded in the culturally universal elements of human life (5.4).

5.1. PARSIMONY IN SEMANTICS

Parsimony in semantics

1 *Oligosynthesis*

ides constraint 1 (cultural neutrality), which will be discussed in chapter 5.4, Atlan's lexicon has to follow constraint 2 (unambiguity) and 3 (parsimony), which will be discussed in this section. Ideally, the language should contain as little basic words as possible, as to reduce the time required to learn the language. Therefore, it should be as sparse with its words, only adding new words when they carry a meaning that is not already covered by another word. Complex concepts should not get their own separate words, for this would add an inestimably large number of extra words, but rather be composed of more simple and universal words that constitute its meaning. Atlan shall achieve this by using a semantic system that is oligosynthetic, meaning that it has a very limited number of semantic atoms¹(*oligo* = few), from which more complex meaning is built by combining different atoms (*thesis* = combining). Each semantic atom (or 'root') shall be covered by a unique one-syllable word. Atlan's syllables can take four shapes (C = some consonant, V = some vowel): V, CV, VC, CVC. This is abbreviated to

R 5. OUR ONTOLOGY – STIJN JANSSENS

There are 9 consonants and 5 vowels in its phonetic inventory, yielding a total of $5 + 5 \times 9 + 5 \times 9 + 5 \times 9 \times 9$ possible combinations, however all syllables ending in *-i*, so all *(-ij)* will not be included. This gives a total number of 490, and thus the challenge posed in this chapter is that of reducing all meaning to 490 possible combinations of these.

The term *oligosynthetic* was first coined by the linguist Benjamin Lee Whorf and is defined as having at most a few hundred-word roots. However, this seems extremely rare among natural languages. A possible example might be the Kalam language of the Highlands of New Guinea (Pawley A. , 1993). Two other languages previously regarded as oligosynthetic by Whorf are the Aztec language Nahuatl and the Native American language Blackfoot, but these are now commonly regarded as polysynthetic (using many roots to synthe-

Throughout the book we refer to Atlans syllable words as 'sems / primes', but this definition is roughly equal to the term 'morpheme', which means as much as 'the smallest unit of meaning'. Since the atoms play a crucial role in Atlan's ontological structure, and this is not covered by the term 'morpheme' (which also doesn't necessarily have to be a single syllable), we opt for this specified terminology.

5.1. PARSIMONY IN SEMANTICS

more complicated meaning). Oligosynthesis is more popular among constructed languages, such as Sona, Ro, Ygyde and Kali-sise (FrathWiki, Oligosynthetic language, n.d.). These all have different numbers of semantic primes and methods of synthesising them, but they commonly have the following problems (Watson, n.d.):

- . Complicated combinatorial systems
- . Unclear word-parsing
- . Vagueness of composite meaning

Atlan will overcome problem I by using an extremely simple manner of combinatorial synthesis: the most semantically essential prime comes first and is followed by primes that hierarchically specify the meaning of the word. Grammatical functions always come in front of the semantic root as prefixes (except for the plural¹), and semantic specifications are appended as suffixes. Atlan's word composition is as follows:

grammatical.function – main.semantic.root –
semantic.specification – plural

R 5. OUR ONTOLOGY – STIJN JANSSENS

will also solve problem II, because syllables can be defined to be semantic when they are CVC, and non-semantic otherwise. This way, grammatical syllables can be distinguished from semantic ones. It is therefore always audibly clear where a word begins and ends. In text this is aided by the fact that in Atlan's own system, CV and VC syllables always consist of a vowel attached to a line, while CVC syllables are always two lines connected to each other.

CV and VC syllables will be restricted to mood-markers for general sentence structuring, because of the onomatopoeic quality of these basic vocal sounds. These, however, can also be used grammatically to modify the meaning of nouns, verbs and pronouns, for example to turn 'where' into 'where?'.

Exclamative (prosody), imperative, vocative = o \subset

Interrogative (question, prosody) = e \circ

Stress marker (prosody) = a + stress \cap

Relative clause = i (+ pronoun) \supset

Conjunctive (wish) = u \cup

5.1. PARSIMONY IN SEMANTICS

The 89 CV, VC syllables will be mostly restricted to morphological and abstract functions, indicating grammatical, syntactical or logical functions and relationships between words. Logical functions will be derived from logical operators of predicate logic and possible worlds semantics (Priest, 2011), and grammatical functions are derived from the Universal Networking Language (UNL) Portal, sd), and syntax-semantically reduced where possible (see chapter 6.2). UNL was initiated by the United Nations University in 1996 and continued by the international non-profit organisation UNDL from 2001 onwards. Its goal is to function as a formalised pivot language between natural language and interlingual machine translation, therefore having formalised all grammatical functions, together with a large-scale ontology of all concepts contained within all its source languages (Universal Networking Language Portal, n.d.). Atlan does not use the latter, however, because it would break constraint 3, parsimony.

The remaining 396 CVC syllables will cover the semantic primes, and these will be systematically selected and ordered in the remainder of the current chapter and chapter 3 and 4. Because Atlan's writing system is syl-

5. OUR ONTOLOGY – STIJN JANSSENS

and each syllable has a fixed semantic value assigned to it, individual glyphs can be read both phonetically as well as ideographically/logographically. This would make it to be a kind of interlingual orthography.

Item III, that of vagueness of composite meaning, is tackled in several ways. Most importantly, some of the most universal ‘semantic molecules’ (see chapter 4) have their own assigned syllable. These molecules are simple notions that could be reduced to more fundamental ones but are often used to build more complex meanings, therefore being condensed into a molecule as to avoid the unnecessary complexity of compound words. However, this is not an all-encompassing solution for words that require very specific and context derived definitions.

Circumlocution is the phenomenon where concepts which do not have a specific word for them in a language are described by giving a circuitous description of the intended meaning. An example of this in English would be ‘the day after tomorrow’. Atlan will never be able to eliminate some forms of circumlocution in its lexicon, mainly because of it being an oligosynthetic language. However, confusion around these instances can be analysed in the following ways.

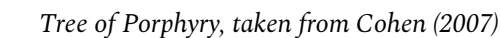
5.1. PARSIMONY IN SEMANTICS

A standardized set of compound definitions can aid speakers by being a guideline to using compound words. Someone who learns the language, would have to learn 490 semantic primes, and then be familiarized with the standardized compound words. Because the meaning of the compound word will be derived from the syllables it contains, learning these compound words will be intuitive and require less mnemonic effort than learning a completely new word would. This way, when a speaker encounters a word they have never heard before, they will be able to derive, or at least estimate the intended meaning just by recognizing its syllables.

Additionally, neologisms could be created during improvised language use, as a sort of generative etymology, which can be directly perceived by the listener, possibly allowing for freer linguistic cultural- and self-expression. Finally, some compound words will be systematically connected in a taxonomical fashion, when the word's definition allows for this, such as is the case for all living creatures. Each compound word is sorted along the axis of importance, with the most fundamental semantic prime in front, and followed by other primes that hierarchically add layers of semantic precision. This creates a univer-

in the composition of compound words that allow us to identify the ontological category in which a word falls and refine the definition by the refining words appended to it, as if ‘zooming in’ with a series of lenses.

reme genus:



philosophical constructed languages that came before were suggested or employed a so-called taxonomy. Many of these are inspired by a diagram by the 3rd century Neoplatonist Porphyry when

5.1. PARSIMONY IN SEMANTICS

aining Aristotle's Categories (Porphyry, 300), called the Porphyrian tree (Franklin, 1986). In the Categories, Aristotle outlines a broad ontology of human apprehension, classifying everything that can be the subject or predicate of a statement into 10 categories: substance, quantity, quality, relation, place, time, position, state, action, affection (Aristoteles, 40). The Porphyrian tree (shown below) shows how a human falls into these categories, by showing different bifurcations from the first categories (Cohen, 2007):

Before Aristotle, similar ontological categorisations were made by the Vaisesika school (Padārtha) (Stanford University, 2019) and the Stoic school, and after him numerous other thinkers, including Plotinus (430, 2019), Kant (categories) (1781, 1998), Hegel (1812, 1975), Peirce (1867), Husserl (1900, 1993) and Whitehead (categorical scheme) (1919, 2010). Not to mention folk ontologies, such as the traditional nominal classes (Bleek, 1862–69) and the common distinction between animate and inanimate. All of these were not taxonomical, however, and have been discussed extensively, but it is not the purpose of this paper to further investigate these discourses.

5. OUR ONTOLOGY – STIJN JANSSENS

There is, however, one specific notion within these categories which Atlan will incorporate, namely that of the *radical categorisation* of Being. In Hegel, this comes in three degrees: 1) Being (mind, consciousness, sensation), 2) Essence (Quality) and 3) Notion (synthesis, reference). In Atlan's system, these equate to Firstness (Quality, feeling, consciousness), Secondness (Reaction) and Thirdness (Meaning, representation). Atlan incorporates this through three degrees of removing from the speaking subject (see chapter 4). The first degree can be combined with the second degree to form a prime for 'person' ('EJ' ∅) to mean 'I' ('EJ.AM' ∅) or the second degree to mean 'you' (EJ.UN ∅ ∅), and the third to mean 'he / she / they' ('EJ.AJ' ∅ ∅), as well as other primes like place, time and demonstratives. This is an explicit taxonomical ontology makes use of a hierarchical classification. The idea of a 'perfect' philosophical language was a popular idea during the Enlightenment, being discussed by Bacon, Descartes, and others as a part of a widespread desire for a language that does not confuse the speaker's understanding or distort the natural order present in it (Eco, 2009-227).

5.1. PARSIMONY IN SEMANTICS

The core idea of such a language is that it would build words by adding different letters for every hierarchy meaning. This way, related words would sound similar. There is, however, always a necessary degree of arbitrariness to such a system, since the arbitrary choice has to be made of what categories are specified by adding a limited set of different letters. These would have to be made from scratch, not based on previous languages, as to avoid confusion caused by natural language. These languages are known as a priori, and Atlan will be a priori as well, even though its contents are sourced from words from many natural languages but recombined in an attempt to circumvent natural language confusion.

The first serious attempts at this ideal were made by Pierre Dalgarno in his *Ars signorum* (1661, 1968) and John Wilkins in his *Essay Towards a Real Character and a Philosophical Language* (1668, 1968). Initially, the two collaborated on a philosophical oligosynthetic language, but they couldn't agree on whether to make the taxonomy cyclopaedic or build compound words from a small set of primes. Wilkins published his own version based on the former and Dalgarno the latter. Dalgarno's language never caught on, perhaps because the explanation

R 5. OUR ONTOLOGY – STIJN JANSSENS

linguistic working of the language was shrouded in mystery which explained the structure².

Linnaeus's language found a bit more recognition, being eventually taken seriously by the Royal Society, with the intent to finish the language after his death by a despatch committee. It too, however, slowly lost interest and descended into oblivion (FrathWiki, Ars Linnaea, n.d.). Its taxonomical classification was structured to encompass every animal, plant, mineral and artefact. He achieved this by setting up an ingenious taxonomic tree to indicate the relations and bifurcations between things, along with a system of hierarchically adding prefixes and consonants to specify differences and species within different categories.

Linnaeus regarded the language presented in his chapter as a draft, although he provides 2,030 different words, as well as a 15,000-word list for different English words but admits that it should be worked out by different teams of scientists to work out different conventions within their respective disciplines. His collaboration with the Royal Society was largely part of that atmosphere. Later on, Wilkin's taxonomy went on to inspire

² I hope this is not the case for this book as well.

5.1. PARSIMONY IN SEMANTICS

et's Thesaurus (1805) and later on Diderot and d'Alembert's Encyclopaedia (1759).

The idea was met with criticism as well. Voltaire criticised the optimism of the people attempting to create a language in the form of the character Dr. Pangloss in his satire *Candide* (1759, 1963). Jorge Luis Borges wrote an chapter criticising taxonomical categorisations of General and Wilkin's language specifically (1942) in which he mocks different instances of arbitrary classification by mentioning a fictional Chinese taxonomy called *Celestial Emporium of Benevolent Knowledge*. This list contains some very culture-specific, arbitrary and absurd categories such as 'those belonging to the Emperor', 'those who have just broken the vase' and 'those that from afar look like flies'. This criticism seems like a bit of a stretch, because Wilkins put in a systematic effort to make a coherent classification, and this is not as arbitrary or absurd as Borges' fictional classification.

The linguist George Lakoff supports this claim by arguing that many non-western cultures use classifications similar to European ones (Lakoff, 1987). Borges also point out that a successful execution of the idea would in theory have many benefits: "Mauthner points

5. OUR ONTOLOGY – STIJN JANSSENS

children would be able to learn this language knowing it be artificial; afterwards, at school, would discover it being an universal code and a se-yclopaedia” (Blevins, n.d.).

Walt was inspired by Borges’ essay to write his *The Order of Things* (1966, 2010, preface) in which he discusses the social grounding of epistemic assumptions. He argues that implicit norms within intellectual communities determine thought and influence which topics are researched, and which are not, and how the establishment influences the interpretation of the data that is collected. These assumptions and norms are bound to cultural and historic settings, and periodically go through changes as a result of paradigm shifts. This is a strong critique of the aspiration of a universal classification of the world. Borges claims that this is because: “we do not know what sort of thing the universe is”. Metaphysicians and phenomenologists might differ on this, however, as discussed in chapter 3 of this chapter.

For this reason, Wilkins’ system has the disadvantage of requiring a limited number of differences and species to be specified because of the limited phonology. It will be more similar to Dalgarno’s language, in that

5.1. PARSIMONY IN SEMANTICS

will not drive its words through a hierarchical process taxonomy, but rather by combining primes at will, allowing for exponentially more semantic combinations than are possible in Wilkin's system.

Another problem of Wilkin's language is that words with similar meanings have very similar pronunciations, leading to the point of confusion. Modern information theory warns of this (Norman, n.d.), and Eco even identified that Wilkins himself made such a mistake, confusing *Gade* (tulip) for *Gape* (tulip) (1994, p. 249). This would hinder the language's intelligibility when mishearing can easily change important nuances in definition, as well as making it harder to speak fluently, because any speaker would have to work through tables and flowcharts in their heads while simultaneously talking, without making any mistakes. The language would also be very intolerant of subtle shifts in pronunciation and phrasing that tend to occur naturally within languages over time, because this would cause the whole encyclopaedic house of cards to be crashing down.

Atlan will not have this problem, because its semantic primes are syllables instead of phonemes, and Atlan's phonemic inventory is built to accommodate variation

R 5. OUR ONTOLOGY – STIJN JANSSENS

etic approximation and sound shifts within its type letters, without causing confusion or am-

philosopher Deleuze and psychoanalyst Guatarri and an alternative to arborescent (tree-like, hierarchical) epistemic networks like employed by Wilkins (1980, 2019), namely that of a *rhizome*, an analogy with a decentralised plant root network (1980, 2019). Such a model does account for bifurcations and conceptual reconnections but is more modal and allows for more complex interlinking than mere hierarchy. It would also fit philosopher Quine's idea of an interrelated epistemic 'web of beliefs' (Ney, 2014), as well as Wittgenstein's claim that concepts are not clearly delineated, but rather surrounded by a 'corona' of associated concepts (1953, 2010,

Because of this, the main ontology seems to be better described as a combinatory system, which would allow for free recombination and web-like relationships between words. This however doesn't mean arborescent ontology should be completely abandoned. The most modern case of taxonomical classification is that of bacterial species, although individual species don't have

5.1. PARSIMONY IN SEMANTICS

er demarcations and are loosely defined by their ability to produce fertile offspring (Nature Publishing Group, 2018). Genetic diversification happened through bifurcations, known as *speciation*. The reverse, different species merging into one through hybridisation, called *despeciation*, does sometimes occur (including among early hominins), but is exceedingly rarer (Junior, 2018). Because of this, the evolutionary tree of life is primarily arborescent.

Modern biological taxonomy employs the following hierarchical classification: life – domain – kingdom – phylum – class – order – family – genus – species (Biological Dictionary, 2017). Atlan's biological lexicon is connected along this framework, using semantic primes to describe different bifurcations, inspired by the Latin etymologies employed in binomial nomenclature (a hangover from Latin being the academic *lingua franca*). These binomial nomenclatures only mention genus and species names of an organism, and words to designate species that re-occur in other genera to identify other species, leading to the parsimony of terms required to name all organisms within this system.

Atlan shall have separate primes for the categories:

R 5. OUR ONTOLOGY – STIJN JANSSENS

bacteria, archaea, plants, amoebas, fungi, animals. In the rest of this chapter, a few culturally universal animals are identified, and these will be reduced to: *fish, bird, worm, reptilian, insect*. Other animals will be reduced to their descriptions: sponges could be described as foam-animals, starfishes to star-animals, legless-reptiles, amphibians to mucus-reptiles, shellfish to shell-animals, jellyfish to mucus-animals etc.

Chemical molecules could be named by formalising a version of the IUPAC nomenclature of organic chemistry in the originally Greek roots (IUPAC, 2021). Just as in Einstein's language, Atlan will be dependent on scientists and specialists of different kinds of professions to formalisations of their respective jargon nomenclature in the lexicon in order to fully flesh out its lexicon.

One reason why such a systematic description of reality, claiming to be universal might be problematic, could be that it claims to have an objective image of reality, while human thought, individual or collective, will always be fundamentally sourced from subjective experience. The philosopher Thomas Nagel described this as the claim that there is no *view from nowhere* (1989), while Atlan's taxonomy appears to claim this anyway. The fact

5.1. PARSIMONY IN SEMANTICS

that human concepts always have a necessary degree of arbitrariness because of the limited resolution of our conceptual boundaries. A taxonomical ontology implies the possibility of grasping some fundamental reduction-principle inherent to reality, while failing to see that many human concepts are emergent phenomena. The reductionistic concept of an 'organism' cannot be reduced to a collection of organic molecules, because it is their coordinated interplay that generates the multiple properties that pertain to an emergent system of self-preservation that humans call a single organism (Brigandt & Love, 2007). On a microscopic level however, the clear boundaries of any single individual become fuzzy. It is because human life does not generally take place within a microscopic paradigm, that our concepts don't have this level of detail. Reality is never described objectively, but always relative to the individual(s) observing and describing their reality. Humans appear to be 'at the centre' of their own language and understanding of reality.

the anthropocentricity of language

Language and ontology are strongly entwined with one another. An ontological system is dependent on the words used to name its parts, and likewise a language is dependent on the set of concepts, relations, abstractions and other things that are captured by its lexicon (Moltmann, 2017) (Moltmann, 2019). Though by far not being the only elaboration on this idea, the Sapir-Whorf hypothesis is the most well-known inference that has been drawn from this support, this idea postulates that the range and limits of a person's thought are determined by the language they speak. The strong version of this claim, linguistic determinism contends that all of human thought is fundamentally determined by language (linguistic determinism resulting in some thoughts being lost and modified through translation or even untranslatable. It treats language as a fixed set of cognitive tools that acts as a constraint on the individual. This view, however, doesn't have a scholarly consensus (Whorf, 1956).

However, this conception of language seems to be static and extra personal instead of dynamic and dependent on individuals and their fluid interactions. It

5.2. THE ANTHROPOCENTRICITY OF LANGUAGE

challenged when presented with the fact that people within the same language can have vastly different ontologies, philosophies and vocabularies, depending on their individual personalities, interests and social environment, and the fact that individual people can learn multiple different languages and express their thoughts through them, nonetheless. Multilingualism is, however, used to create within an individual, multiple linguistic 'personæ' for the different spoken languages, where each is a way of formulating thoughts and uttered sentences, altered by the individual characteristics of the different languages (Pavlenko, 2006).

Perhaps the influence that language has on the thoughts of a speaker can be likened to how putting on different types of glasses can alter one's perception but does not change the fundamental scene being perceived through them. Clear glasses block out UV light, tinted glasses block out certain colours, different lenses shift the focus to what is near and other to what is far etc.: they all suppress some elements and amplify others, but they never change the basic composition of what is being perceived (if we discount virtual reality glasses).

Using this metaphor, the purpose of Atlan is some-

R 5. OUR ONTOLOGY – STIJN JANSSENS

ilar to being a clear, untainted, undeformed, un-
pair of linguistic glasses, for as far as this is pos-
every single person's eyes are different, and the
of their native language might be more or less
o Atlan's. The question then becomes: what
es this clear human experience that becomes
y language?

we must realise that language itself is onto-
dependant on the total sum of living speakers.
forgotten or undeciphered language cannot be
currently exist in the same way that a living lan-
ke English exists. It might be revived in the fu-
ough the reconstruction of its linguistic infor-
but it only comes back into being when living
are again able to read, write or speak the lan-
urthermore, Wittgenstein's private language ar-
states that a language is a fundamentally social
d that a purely personal language is therefore by
n impossible (1953, 2010, §243-271). Language
licated system of communicating all kinds of
nformation, like thoughts, feelings, intentions,
data &c., for all kinds of different purposes, like
ion, social bonding, problem solving . In indi-

5.2. THE ANTHROPOCENTRICITY OF LANGUAGE

Animal growing up in solitude or alongside animals never has the need nor possibility to learn and use a language, because there are no other humans around to converse with. After having passed the critical period of language acquisition without ever having learned a human language, an individual will never again be able to do so for the rest of its life (Robson, 2002).

Therefore, language is an inherently human thing, which emerged from the transferring of information from one person's individual experience to another's. Phenomenal cues, like the sound of words, the rhythm of speech, facial expressions and gestures are used as an interpersonal bridge between the private mental worlds of separate individuals. Someone can both hear themselves talking, as well as someone else: language exists in a shared phenomenal space, whereas inner thought is private. Language then becomes a highly codified system of phenomenal metaphors. The sound of a specific word is not the same as the information it codifies, but is consistently associated with the referred phenomenon, in the form of an abstracted 'concept'.

Atlan should thus have an ontology that is built off subjective human experience, when regarded in a so-

R 5. OUR ONTOLOGY – STIJN JANSSENS

text and in direct contact with its physical environment. This immediately brings a degree of anthropocentricity with it, because words relating to the human mind, the body, daily life, social environment etc. will have a higher priority than the myriad of concepts and terms within specialized disciplines that are less directly related to the everyday human experience. Moreover, we should be able to think, talk and understand fluently in our language, and not be required to consciously perform complicated linguistic computations in their heads while using the language.

Phenomenological ontology and qualia

Our experience is ultimately prior to any claim, observation, connection &c that can be communicated about reality. Anything in the world has to first present itself to us humans through phenomenal, subjective experience, before we can abstract it and understand it as an ‘objective’ phenomenon. However, the mainstream scientific metaphysical framework has, up to now, been materialist, physicalist and reductionist. Even when it does acknowledge the existence of mind,

5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA

often does so in a greatly unsatisfactory manner by employing some version of dualism, with the mind being metaphysically separate from the physical, but somehow miraculously still having epistemological and sensory access to it and the ability to manipulate it (moving one's body at will) and be manipulated by it (physical alterations to the body or the ingestion of physical substances alter the mental perception). When confronted with these problems, science will often try to explain the connection through a functionalist account of the neural network and a materialist explanation of the composition of neurons, but always failing to close the explanatory gap to how this material process constitutes phenomenal experience. This is most painfully brought to light by the Hard Problem of Consciousness, the insurmountable chasm between a mechanistic description of neurology and the subjective experience of what it 'feels' like to be a conscious entity. Heidegger, building off the phenomenological philosophy of Husserl, already concerned of this in his own time halfway the 20th century, called it *Seinsvergessenheit*, the 'forgottenness of Being' (1962, 2019). Somehow the abstractions of reality that were derived from experience have gotten a higher

R 5. OUR ONTOLOGY – STIJN JANSSENS

cal priority than the original experience itself, 'reflective' is regarded with a higher esteem than the 'original'. It is beyond the purposes of this chapter to explain why this happened and why it is metaphysically contradictory. Therefore, building off the premise of historic anthropocentricity established in the previous chapter, I shall relate the relevance of the subjective phenomenal experience to the construction of a universal ontology in the current chapter.

A famous thought experiment regarding the irreducibility of experience is called 'Mary's room' (Jackson, 1982). It describes a hypothetical scientist named Mary who (disregarding ethical concerns for the sake of the thought experiment) is raised in an exclusively black and white environment for her whole life, and educated about the physics of colour perception, without ever seeing colour. She would have learned all there is to know about the physics of light, the biology of light receptors in the eye, and the neural processing of visual information in the brain. The thought experiment then asks us: if Mary is suddenly taken out of her black and white environment and allowed to see colours for the first time in her life, will she learn anything new from experiencing, for ex-

5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA

ble, the colour red for the first time? Could she have known its qualitative experience before she left the black and white room? Philosophers generally agree that she could not have known (the ‘knowledge argument’) (Nida-Rümelin & Conaill, 2019).

The same thought experiment could be extended to other subjective sensory perceptions like smell, taste, touch and sound and by extension even emotions and altered states of consciousness. Therefore, these ‘subjective’ qualitative aspects of experience appear to be fundamental and irreducible, modern philosophers call them ‘qualia’ (Nida-Rümelin, 2021). Since the coining of the term qualia in 1929 by C.I. Lewis, the concept has remained mostly confined to a longwinded debate within the philosophy of mind.

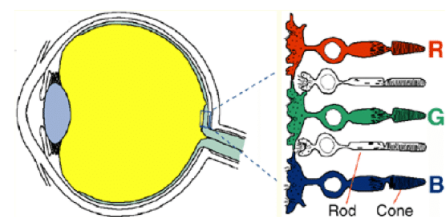
The formalisation of qualia is done by taking an ‘objective’ scale such as the spectrum of light frequencies, and then mapping phenomenal experience onto this by regarding as a fundamental unit the smallest perceptible difference. Classically, qualitative experience is divided into five physical senses: *vision, hearing, smell, taste and touch*. In this chapter I will supplement these with affective emotional experience and altered states of consciousness (the subjective experience of being stoned,

CHAPTER 5. OUR ONTOLOGY – STIJN JANSSENS

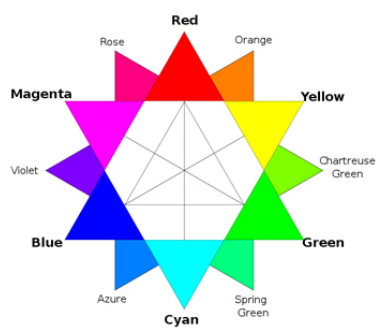
If tripping seems to be irreducible, they can be described when compared to sober consciousness. To know the qualia of the experiment, one must experience the substances personally).

Since the language is oligosynthetic (see chapter 1), the main purpose of this chapter is identifying the basic building blocks of the different types of experiences (vision, sound, taste, scent, physical sensation, and consciousness states), which can then be combined into more nuanced qualitative descriptions.

Vision



5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA



Wheel of colour. From Judge (2012)

tly, colour is the most straightforward, because it is
ady commonly divided into primary, secondary and
ary colours. In psychophysical colorimetry the pri-
y colours red, green and blue are regarded as being
plete, that is, constituting all colours perceivable by
nans when combined along an axis of light to dark
litive light mixing), yet also being *imaginary*, meaning
y existing subjectively as qualia, while not being dis-
uishable as primaries through ‘objective’ measure-
at (Mac-Evoy, 2007). A standard human eye contain
e types of light receptor cones, one for each of the
e colours.
All qualia that exist as polarities on a spectrum will
duced to a semantic prime for the spectrum, com-
ed with the particles for *positive/high*, *neutral* and *neg-*

R 5. OUR ONTOLOGY – STIJN JANSSENS

Translating this into an oligosynthetic semantic inventory, we would get the following (capital letters representing some as of yet undetermined assigned syllable):

DES:

black/dark: JE.LAS ʃ ɳ (= negative + brightness)

white/light: FO.LAS ɸ ɳ (= positive + brightness)

MARY:

red: EL ʔ

green: OS ɔ

blue: UL ɛ

PRIMARY AND TERTIARY (RGB is chosen as fixed

orange: EL.OS ʔ ɔ

yellow: OS.UL ɔ ɛ

magenta: UL.EL ɛ ʔ

5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA

- Orange: EL.OS.EL ു ഓ ു
- Chartreuse green: EL.OS.OS ു ഓ ഓ
- Spring green: OS.UL.OS ഓ ു ഓ
- Azure: OS.UL.UL ഓ ു ു
- Violet: EL.UL.UL ു ു ു
- Rose: EL.UL.EL ു ു ു

In 1969, anthropologist Brent Berlin and linguist Paul Kay published their book '*Basic Color Terms*' in which they proposed their research concerning the prevalence and development of different colour terms in languages around the world (Berlin & Kay, 1969). They proposed a chronological scheme of seven evolutionary stages through which languages generally add colour terms to their lexicon. These are as follows:

- Stage I: dark-cool (>'black') & light-warm (>'white')
- Stage II: red
- Stage III: green OR yellow

R 5. OUR ONTOLOGY – STIJN JANSSENS

ge IV: green AND yellow

ge V: blue

ge VI: brown

ge VII: purple, pink, orange or gray

, needing to conform to constraint 1, cultural
y, should thus contain these colours in its lex-
age I-V have already been accounted for, and
and VII can be covered by the combination of
urs from the earlier stages (following the order
XYZ and primary-secondary-tertiary):

own: red + green + blue = 'EL.OS.UL' ʔ ɔ ɛ

nk: white + red à = 'FO.LAS.EL' ɓ ɔ ʔ

ay: colour + brightness + neutral = 'KAL.UJ.LAS'
ɔ ɔ

ther colours and shades can be achieved using
binatory system. One might argue that not all
s have the same lexical colour inventory, and
ke more or less distinctions than English, but it

5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA

It should be noted that having a word for a specific colour is not the same as being able to perceive these different colours and their (subtle) differences (weak linguistic determinism, see chapter 1). Within the line of thought of linguistic determinism, one could argue that learning to speak Atlán, as having this colour system, would gift the speaker with an intuitive understanding of the composition of phenomenal color.

Besides colour, three-dimensional shape is the other primary irreducible element within vision, constituting what in cognitive science is known as Gestalt (Rollinger et al., 2019). This term is also applicable to proportional ‘shapes’ or patterns within other qualia, like a musical melody. Visual shape can be geometrically reduced to lines/sides (1D), corners, surfaces (2D), angles and volumes (3D), making use of numerals to specify the quantities of these elements, as well as spatial prepositions to indicate relative location. This, however, will be further expounded upon in chapter 4 of the book, on Atlán’s numerals and mathematics.

2 Sound

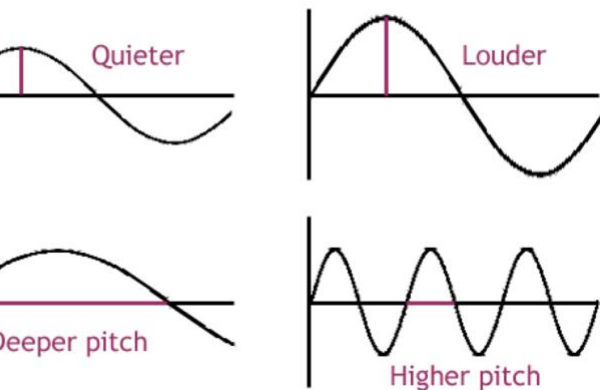
Sound is commonly divided into three elements:

R 5. OUR ONTOLOGY – STIJN JANSSENS

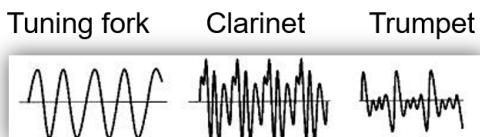
the frequency of the soundwave, phenomenally corresponding with *pitch*;

the amplitude of the soundwave, phenomenally corresponding with *volume*;

the shape of the soundwave, phenomenally corresponding with *timbre*.



Frequency visualised. From Mata (2015).



Timbre. From SimplifyingTheory (n.d.).

5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA

Music theory almost universally describes pitch into a scale of notes, which repeats at the octave, where the ratio to the first note of the previous scale is 2:1. Western music uses the *diatonic* scale, comprised of specifically alternating whole and half

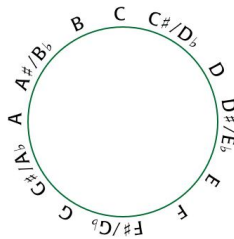


Figure 3: The Chromatic Circle. From Wikimedia (2022).

steps, and which is *heptatonic* because it is comprised of seven notes, before repeating at the octave (8th). There are a total of 12 half note steps within an octave, which comprise the *chromatic* scale. Chromatic notes are specified relative to the *natural* (♮) heptatonic scale, by indicating a half step lower by a *flat* (♭) and a half step higher by a *sharp* (♯).

The diatonic scale is not the most prevalent musical scale in the world, but rather the *pentatonic* scale (five notes) (Encyclopædia Britannica, inc., n.d.). However, a pentatonic scale fits neatly within a heptatonic scale, for example: C–D–E–G–A is a pentatonic scale which only omits the notes F and B. Because of this, Atlan shall use a system based on the Western heptatonic scale, while remaining culturally neutral because this system also ac-

5. OUR ONTOLOGY – STIJN JANSSENS

ates the more common pentatonic scale. If it
se the pentatonic scale as the standard, the hep-
scale would be an extended modification, which
quite clumsy seeing as Western music is so vast
alised.

s will be indicated with a context-dependent op-
semantic prime for *note/pitch* in front of the note
the combination *note* + the first 7 numerals (see
5.2) can be taken to represent the 7 notes (hep-
IP, OP, UP, IK, OK, UK, IM, pentatonic: IP, OP,
UK). This will also make calculating musical in-
more intuitive because it will only require simple
subtraction.

o will be indicated by *pitch-positive* (also desig-
igh in the context of sound frequency) and flat
icating *low* in the context of sound frequency)
negative. Double sharps and double flats could
d by reduplicating *positive/negative* respectively.
nal notes (which fall in between the chromatic
uld be accounted for as follows: half-sharp -PN,
-NP, three-quarter-sharp -PNP, three-quarter-
N. Major and minor could be designated by the
c primes for *happy* and *sad* (see the subchapter

5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA

tions), and the seven scale modes could simply be numbered. Harmonic music theory is way too elaborate and complicated to cover fully in this paragraph, however it could be fairly easily constructed from the building blocks presented here.

Volume is a lot simpler: within music theory it is known as dynamic, and divided into loud (Italian: *forte*, and soft (Italian: *piano*, *p*), further nuanced by the pre-medium (Italian: *mezzo*), and by introducing a three-free scale of intensity (*ppp*, *pp*, *p*, *mp*, *mf*, *f*, *ff*, *fff*). Adopting this system, Atlan can specify volume by combining the semantic prime for *volume* with *positive*, *medium*, *negative*, and a *comparative/superlative* system, which will be applied in other places of the language: *X*, *more X*, *less X*, *most X*. Changes in dynamic (growing louder, *crescendo*, or softer, *decrescendo*) can be described by the semantic primes for *becoming* combined with *more-volume-positive/negative* and *comparative/superlative*.

Finally, timbre, corresponding with the specific shape of the soundwave pattern, has an almost infinite range of possible combinations. This is why in language, terms that denote timbre are always metaphoric approximations, describing the sound with words that denote phe-

R 5. OUR ONTOLOGY – STIJN JANSSENS

unrelated to sound when taken literally, but have a phenomenal quality to the sound (e.g., piercing, bright), or are related to the origin of the sound (e.g., al, metallic). P. Sesuni analysed 45 studies on different terms in English, Japanese, French, Czech, Dutch, Finnish, Spanish and German, and from these studies identified 59 different descriptors (see appendix 1) (Carron, Rotureau, Dubois, & Misdariis, 2017). Because these are all semantically reducible to non-sound-related terms, Atlan will not have any semantic markers specific to timbre, but rather use these and other descriptors, preceded by the semantic prime *de-sound*, combined with an adjective-marker. When used in a clearly sound-related context, the *sound* marker may even be omitted, as there might not be any risk of confusion when it is already obvious that the term refers to a sound.

Taste and Scent

Taste and scent are strongly correlated because both rely on the same receptors (molecule-detectors) (Reina, 2022) , and the main difference between the two is that taste is detected by the tongue and concerns solid and liquid

5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA

ter, while smell goes through the nose and pertains to gaseous matter. Because many tastes have an olfactory counterpart, tastes may be marked by adding the semantic prime *taste* and scents with *smell*.

The five main taste categories are *sweet*, *sour*, *bitter*, *salty*, and *umami* (Deutsch, 2019), which will be separate semantic primes in Atlan. These can double as scents when marked with *smell* instead of *taste*. The true range of scents and tastes, just like timbre, is overwhelmingly complex, thanks to the many different possible molecules and combinations between these. Thus, further nuances in flavour and aroma may be constituted in similar fashion to timbre, referring either to the source of the smell or taste (e.g., floral, alcoholic, vanilla-like), or a comparable quality (e.g., harsh, sharp, mellow). Spice is not a taste, but rather a form of phenomenal pain, because it is registered by chemical nociceptors (molecule-pain receptors). Pain can thus be constituted by combining the semantic prime for *pain* with *taste/smell*. Of course, different taste and scent designators can be combined at will to create an infinite number of more nuanced descriptions.

Physical sensation

Physical senses cover a broad range of different sensations in different parts of the body (Reina, 2022). All physical sensation shall be preceded by the sense prime for *feeling*. This prime might also be used logically/affectively in other contexts. Mechanoreception sense physical deformation like pressure, touch, vibration and motion but also sound, corresponding to vibrations of the ear drum. Because sound has already been covered, it will not be counted among physical sensation. Combined with the range *positive-neutral-negative*, this will yield the following terms:

Contact: pressure – touch – barely touch

Extension: stretched – relaxed – contracted

Texture: rough – normal – smooth

More detailed textural descriptions can be made by combining with different material primes mentioned in the conclusion of this chapter. Terms for motion of touch are mentioned in chapter 5.4.

Thermoreceptors report *temperature* and nociceptors report pain, as stated earlier regarding the taste of spice. *Plea-*

5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA

does not have a specific receptor but is rather constituted by neurotransmitters fired in the brain in reaction to certain perceptions, however, is often seen as the opposite polarity of pain (Kringelbach & Berridge, 2009). We will thus be constituted as follows:

- *Temperature*: hot – tepid – cold
- *Valence*: pleasure – neutral – pain
- *Wetness*: wet – moist/damp – dry

Somatic sense refers to the outside of the body, and visceral sense to internal organs. These will be codified as the semantic primes for *outside* and *inside* respectively.

- Internal *tension*: bloating/swollen – normal pressure – cramp

Proprioceptors sense the relative position of body parts, and the vestibular system registers the orientation of the whole body in space, perceived as *balance* (Proske & Gandia, 2012).

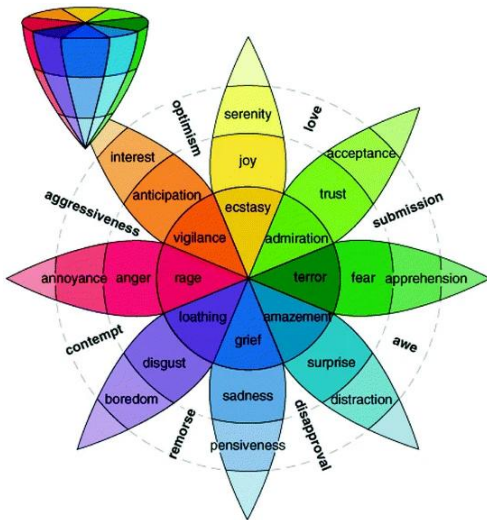
- *Balance*: grounded – balanced – out of balance.
- *Dizziness*: *balance-negative + turning*.

R 5. OUR ONTOLOGY – STIJN JANSSENS

otion sickness: *balance-negative + motion*

asickness *balance-negative + sea.*

motions



Plutchik's wheel of emotions. From Gkonou et al. (2015)

the psychology professor Robert Plutchik published a geometric model of the different basic emotions (Plutchik, 1980). Each pair of geometrically opposing emotions constitute emotional antipoles. The justification of this classification is psychoevolutionary and traces

5.3. PHENOMENOLOGICAL ONTOLOGY AND QUALIA

origins of the different affects to behavioural responses would emerge naturally in reaction to different challenges and attractors encountered by humans and other animals (see appendix 2) (Plutchik & Kellerman, *Theories of emotion*, 1980).

Again, using the grid of positive – neutral – negative, eight basic emotions and their respective three degrees of intensity can be covered:

- . Anger
- . Anticipation
- . Joy
- . Trust
- . Fear
- . Surprise
- . Sadness
- . Disgust

Intermediary emotions can be achieved by combining other emotions or semantic primes (see appendix

R 5. OUR ONTOLOGY – STIJN JANSSENS

consciousness states

erent consciousness states of the human psyche
ardest qualia to pin down. The NOBA Project,
e Psychology teaching platform, identifies the
g (Teener & Teeny, n.d.):

consciousness

wareness

pnosis

ssociation

ance / depersonalisation

ep

llucination

pression

mulation / excitation

lete the pallet of consciousness, I shall add the
g:

ason / abstract logos / logic

5.4. UNIVERSAL SEMANTICS

- . Memory
- . Desire / will
- . Conscience (moral)
- . Intuition / instinct
- . Imagination / creativity
- . Understanding

These can be combined with other semantic primes to constitute various psychological terms (see appendix 4). To sum up this chapter, to cover most all human qualia, we will employ a set of qualia-specific atoms combined with some more general semantic atoms (see appendix 5).

Universal Semantics

Now that the basic constituents of human psychic experience have been accounted for, the different concepts within mankind's material understanding of the world have been largely neglected. This is what this chapter

R 5. OUR ONTOLOGY – STIJN JANSSENS

rectify. In needing to satisfy Atlan's first con-
human universality / cultural-linguistic neutral-
semantic inventory would need to contain se-
primes or syntheses of these that cover a set of
shared cross-linguistically.

comparative linguist Morris Swadesh published
only used list of 207 lexicostatistically univer-
cepts in 1952 (see appendix 6) (Swadesh, 1952),
making a series of revised versions. It is used to
relatedness between languages by analysing the
tive overlap of their words for different Swadesh
Atlan builds on this list as a guideline, grouping
concepts, reducing some and adding other terms
possible (see chapter 6 of the book).

her, more abstracted summary of the basic se-
lements of human language is given by the Nat-
antic Metalanguage developed by the cross-cultural
Anna Wierzbicka (Goddard & Wierzbicka, Mean-
Universal Grammar: Theory and Empirical Find-
02). Though different languages might express
in different ways, the semantic content of NSM
d into 65 semantic primitives, spread over 16
es (see appendix 7) (Levisen & Waters, 2017).

5.4. UNIVERSAL SEMANTICS

Atlan uses this classification as a guideline, and the semantic primitives will be synthesised with the Swadesh

Within NSM, semantic ‘molecules’ are terms that can be reduced to the 65 primitives, or ‘atoms’, but are then used to build more complicated meanings. Where atoms are abstract, molecules are more concrete. There is therefore, an added value of listing such molecules to minimise complexity, because molecule-composite words would be a lot more longwinded when all their semantic atoms had to be stated individually. The research on this topic is still underdeveloped, but a few sets of supposedly universal semantic molecules have been proposed. The website of the university of Griffith mentions the following universal semantic molecules (see appendix 9) (Griffith University, n.d.).

This lexicon already bears a striking similarity to the Swadesh list. Cliff Goddard identifies a few additional semantic molecules in English (Goddard, 2012). These can be incorporated because of the globalised distribution of Anglo-American culture and language. In a video lecture he adds several other molecules, mostly culture- and language-specific (NSMLab, 2021). Atlan will synthesise these lists into the semantic inventory (see appendix 8).

including remarks

Chapter, I have introduced the problem of an oligosyn-
ontology to constitute Atlan's semantics in a way
pects the constraints of cultural neutrality, un-
ty and form from function. I discussed previ-
mpts at a system like this and other systems of
ation, and the critical discourse around these. I
d the role of the human perspective within hu-
guage, and the relationship between language
and subjective thought. I dove into the philo-
movement of phenomenology and argued for
ucibility of qualitative experience (qualia). I set
ate-space of qualitative experience, and mapped
different qualia could be reduced into an oligosyn-
combinatory system of word-generation. I then
ed different academic projects that mapped out
al account of cross-culturally universal irreducible
, which will be added to the language in order
ct the constraint of cultural neutrality and the
of linguistic anthropocentricity. Chapter 6.2 of
contains the final semantic inventory of Atlan's
antic syllable-primes, sorted into conceptually

5.6. APPENDIX

ted semantic categories. These are sorted into hav-
maximally similar initial letters and modelled by AI
e as similar as possible to various different natural
guages, weighed by their linguistic genealogy and to-
amount of speakers, making the lexicon semi-a priori
rced.

Appendix

imbre descriptions in natural languages.

<i>Occ.</i>	<i>Eng. word</i>	<i>French word</i>	<i>Occ.</i>	<i>Eng. word</i>	<i>French word</i>	<i>Occ.</i>	<i>Eng. word</i>	<i>French word</i>
29	Soft	Doux	9	Light	Léger	6	Uneven*	Irrégulier*
28	Dull	Sourd, mat	9	Noisy	Bruité	6	Deep	Profond
21	High	Aigu	9	Muffled	Feutré	6	Narrow	Etriqué
21	Loud	Fort	9	Large	Large	6	Tonal	Tonal
19	Low	Grave	9	Strong	Puissant	6	Cold	Froid
19	Sharp	Aiguisé, incisive	9	Resonant*	Résonant*	6	Near	Proche
19	Rough	Rugueux	8	Thin	Mince	5	Piercing	Perçant
18	Bright	Brillant	8	Long*	Long*	5	Strident	Strident
16	Smooth	Lisse	8	Continuous*	Continu*	5	Irregular*	Irrégulier*
15	Clear	Clair	8	Dark	Sombre	5	Vibrating	Vibrant
15	Round	Rond	8	Quiet	Calme	5	Constant*	Constant*
15	Rich	Riche	8	Clean	Net	5	Aggressive	Agressif
14	Nasal	Nasal	8	Calm	Calme	5	Heavy	Lourd
14	Full	Plein	8	Harsh	Rêche	5	Complex	Complexe
13	Hard	Dur	7	Shrill	Criard	5	Dynamic*	Dynamique*
11	Weak	Faible	7	Short*	Court*	5	Natural	Naturel
10	Slow*	Lent*	7	Powerful	Puissant	5	Empty	Creux
10	Fast*	Rapide*	7	Metallic	Métallique	5	Far	Lointain
10	Even*	Régulier*	7	Open	Ouvert	5	Edged	Tranchant
10	Warm	Chaud	6	Ringing	Sonnant			

psycho-evolutionary classification of animal emotions

R 5. OUR ONTOLOGY – STIJN JANSSENS

Obstacle	"Enemy"	Anger, rage	Biting, hitting	Destruction
Potential mate	"Possess"	Joy, ecstasy	Courting, mating	Reproduction
Loss of valued person	"Isolation"	Sadness, grief	Crying for help	Reintegration
Group member	"Friend"	Acceptance, trust	Grooming, sharing	Affiliation
Gruesome object	"Poison"	Disgust, Loathing	Vomiting, pushing away	Rejection
New territory	"What's out there?"	Anticipation	Examining, mapping	Exploration
Sudden novel object	"What is it?"	Surprise	Stopping, alerting	Orientation

- Disgust
- Faith / belief = trust + know
- Extravagance = ecstasy + distracted
- (In)security = (not +) trust + self
- Discouraged = passive + not + trust + anticipation
- Jealousy = desire + annoyed

Combinatorics of Atlan's consciousness and semantic atoms

- Unconscious = not + consciousness
- Sub-conscious = below + consciousness
- Ego = feeling + self
- Self-consciousness = consciousness + self
- Narcissism = admiration + self + self
- Selfishness / egotism = interest + self
- Depersonalisation = feeling + not + self
- Ego-death = feeling + self + death
- Derealisation = feeling + not + contact + reality
- Libido = desire + sex
- Arousal = stimulation + sex
- Orgasm = ecstasy + sex
- Deep sleep = sleep + not + consciousness
- Dreaming = sleep + consciousness
- Lucid dreaming = dreaming + consciousness + awareness
- Enlightenment = consciousness + light / bright
- Bliss = ecstasy + peace
- Mystical experience = consciousness + God
- Sensory overload = feel + excitation + positive
- Peace = excitation + neutral
- Numbness = feel + excitation + negative
- Euphoria = feeling + good
- Dysphoria = feeling + bad
- High = feeling + cannabis + excitation + positive
- Stoned = feeling + cannabis + excitation + negative
- Tipsy = feeling + alcohol + neutral
- Drunk = feeling + alcohol + positive
- Understand = reason + grasp
- Alas-Erlebnis = feeling + understanding
- Empathy = feeling + other
- Social awareness = awareness + social
- Intelligent = reason + positive
- Dumb = reason + negative
- Guilt = conscience + bad
- Know-how = understanding + action
- Wisdom = understanding + life

Overview of qualia-related atoms

- Negative
- Neutral
- Positive
- Colour
- Brightness
- Red
- Yellow
- Blue
- Sound
- Volume
- Become, transform
- Note/pitch
- 7 note-names corresponding with the numbers 1-7
- More (comparative)
- Most (superlative)
- Snell
- Taste
- Sweet
- Sour
- Bitter
- Salty
- Unami
- Feeling / affect
- Contact
- Tension
- Texture
- Temperature

5.6. APPENDIX

- Disgust
- Faith / belief = trust + know
- Thoughtfulness = serene + interested
- Extravagance = ecstasy + distracted
- (In)security = (not +) trust + self
- Discouraged = passive + not + trust + anticipation
- Jealousy = desire + annoyed

Combinatorics of Atlan's consciousness and semantic atoms

- Unconscious = not + consciousness
- Sub-conscious = below + consciousness
- Ego = feeling + self
- Self-consciousness = consciousness + self
- Narcissism = admiration + self + self
- Selfishness / egotism = interest + self
- Depersonalisation = feeling + not + self
- Ego-death = feeling + self + death
- Derealisation = feeling + not + contact + reality
- Libido = desire + sex
- Arousal = stimulation + sex
- Orgasm = ecstasy + sex
- Deep sleep = sleep + not + consciousness
- Dreaming = sleep + consciousness
- Lucid dreaming = dreaming + consciousness + awareness
- Enlightenment = consciousness + light / bright
- Bliss = ecstasy + peace
- Mystical experience = consciousness + God
- Sensory overload = feel + excitation + positive
- Peace = excitation + neutral
- Numbness = feel + excitation + negative
- Euphoria = feeling + good
- Dysphoria = feeling + bad
- High = feeling + cannabis + excitation + positive
- Stoned = feeling + cannabis + excitation + negative
- Tipsy = feeling + alcohol + neutral
- Drunk = feeling + alcohol + positive
- Understand = reason + grasp
- Ala-Erlebnis = feeling + understanding
- Empathy = feeling + other
- Social awareness = awareness + social
- Intelligent = reason + positive
- Dumb = reason + negative
- Guilt = conscience + bad
- Know-how = understanding + action
- Wisdom = understanding + life

Overview of qualia-related atoms

- Negative
- Neutral
- Positive
- Colour
- Brightness
- Red
- Yellow
- Blue
- Sound
- Volume
- Become, transform
- Note/pitch
- 7 note-names corresponding with the numbers 1-7
- More (comparative)
- Most (superlative)
- Snell
- Taste
- Sweet
- Sour
- Bitter
- Salty
- Unami
- Feeling / affect
- Contact
- Tension
- Texture
- Temperature

R 5. OUR ONTOLOGY – STIJN JANSSENS

- Disgust
- Faith / belief = trust + know
- Extravagance = ecstasy + distracted
- (In)security = (not +) trust + self
- Discouraged = passive + not + trust + anticipation
- Jealousy = desire + annoyed

Combinatorics of Atlan's consciousness and semantic atoms

- Unconscious = not + consciousness
- Derealisation = feeling + not + contact + reality
- Libido = desire + sex
- Arousal = stimulation + sex
- Orgasm = ecstasy + sex
- Deep sleep = sleep + not + consciousness
- Dreaming = sleep + consciousness
- Lucid dreaming = dreaming + consciousness + awareness
- Enlightenment = consciousness + light / bright
- Bliss = ecstasy + peace
- Sub-conscious = below + consciousness
- Ego = feeling + self
- Self-consciousness = consciousness + self
- Narcissism = admiration + self + self
- Selfishness / egotism = interest
- Depersonalisation = feeling + not + self
- Ego-death = feeling + self + death
- Drunk = feeling + alcohol + positive
- Understand = reason + grasp
- Alas-Erlebnis = feeling + understanding
- Empathy = feeling + other
- Social awareness = awareness + social
- Intelligent = reason + positive
- Dumb = reason + negative
- Guilt = conscience + bad
- Know-how = understanding + action
- Wisdom = understanding + life

Overview of qualia-related atoms

- Negative
- Neutral
- Positive
- Colour
- Brightness
- Red
- Yellow
- Blue
- Sound
- Volume
- Become, transform
- Note/pitch
- 7 note-names corresponding with the numbers 1-7
- More (comparative)
- Most (superlative)
- Snell
- Taste
- Sweet
- Sour
- Bitter
- Salty
- Unami
- Feeling / affect
- Contact
- Tension
- Texture
- Temperature

5.6. APPENDIX

Chapter 6

Icon – Jap Antonisse

ACCORDING to research by First Education (2022), the Dutch are positioned as the most proficient speaking population globally. Other countries achieved a position in the top 10 ranking were Denmark, Belgium, Sweden, Finland, and Germany. While the Dutchman may attribute this success to hard work and dedication, it is worth considering other factors that could be at play here. Notably, these countries with commendable English proficiency are all located in Europe and speak very similar languages.

Linguistics offers a unique perspective on the relationships between different languages. By comparing

vocabulary, grammar and sound systems of various languages, researchers have identified related language families and have constructed language family trees to illustrate the evolution and divergence of languages over time. Thanks to that research, we know that many countries with notable position in the English Proficiency Index share a common linguistic background. Such a linguistic background, or language family, thus may provide a foundation for proficiency in a new language.

This would imply that certain languages are easier to learn for certain population groups. For Atlan, it is deemed important that it will become a language that is easy and quick to learn for *everybody*. This is a challenging task but might be achievable if we find some sort of shared background between almost every natural language. If it is possible to find words that look similar in different languages, which are known as cognates, the translation for those words in Atlan can be designed to resemble them as much as possible. With a model that can do this on a large scale, Atlan will become easy, neutral and global.

To achieve this, it is first key to create some understandings of what methods are used to compare different

R 6. LEXICON – JEP ANTONISSE

es. Therefore, we will take a closer look at the cosine similarity. Thereafter, it is necessary to have an examination of the existing language families. In that way, we gain a deeper insight into the connections between existing natural languages. In addition, with that gained understanding, it is possible to determine which language we will make available in connection to the process of cognate finding. The most relevant languages are weighed against each other to create a dataset that is representative of the real world. All the pieces of the puzzle come together in the final part of the chapter, where the computer program that we will use to generate words in Atlan will be discussed.

Comparison methods and language families

Cosine Similarity

According to research by First Education (2022), the Dutch are considered as the most proficient English-speaking country globally. Other countries that achieved a position in the top 10 ranking were Denmark, Belgium, Sweden, and Germany. While a proud Dutchman might attribute this success to hard work and dedication,

1. COMPARISON METHODS AND LANGUAGE FAMILIES

It is worth considering other factors that could be at play here. Notably, these countries with commendable English proficiency are all located in North Europe and speak very similar languages.

Linguistics offers a unique perspective on the relationships between different languages. By comparing vocabulary, grammar and sound systems of various languages, researchers have identified related language families and have constructed language family trees to illustrate the evolution and divergence of languages over time. Thanks to that research, we know that many countries with notable position in the English Proficiency Index share a common linguistic background. Such a linguistic background, or language family, thus may provide a foundation for proficiency in a new language.

This would imply that certain languages are easier to learn for certain population groups. For Atlan, it is deemed important that it will become a language that is easy and quick to learn for *everybody*. This is a challenging task but might be achievable if we find some sort of shared background between almost every natural language. If it is possible to find words that look similar in different languages, which are known as cognates, the

R 6. LEXICON – JEP ANTONISSE

on for those words in Atlan can be designed to e them as much as possible. With a model that is on a large scale, Atlan will become easy, neu-global.

hieve this, it is first key to create some under- s of what methods are used to compare different s. Therefore, we will take a closer look at the cosine similarity. Thereafter, it is necessary to an examination of the existing language fami- exist. In that way, we gain a deeper insight into ections between existing natural languages. In , with that gained understanding, it is possible e which language we will make available in con- g to the process of cognate finding. The most languages are weighed against each other to cre- aset that is representative of the real world. All ces of the puzzle come together in the final part chapter, where the computer program that we generate words in Atlan will be discussed.

text	Frequency “Merry”	Frequency “christmas”
christmas ”	1	1
stmas”	0	1

1. COMPARISON METHODS AND LANGUAGE FAMILIES

Table 6.1: Word-appearance in “Merry” and “Merry christmas”.

This table can be visualized in a two-dimensional array, where on each axis the count of a word is represented. Now both texts can be placed as a dot on this grid accordingly. Drawing two lines from each point to the origin of the grid creates an angle between those lines. This angle at the origin can be calculated, in this case it would be 45° . To finish the cosine similarity, all that is needed is to take the cosine of this angle, in this example $\cos(45) = 0.71$.

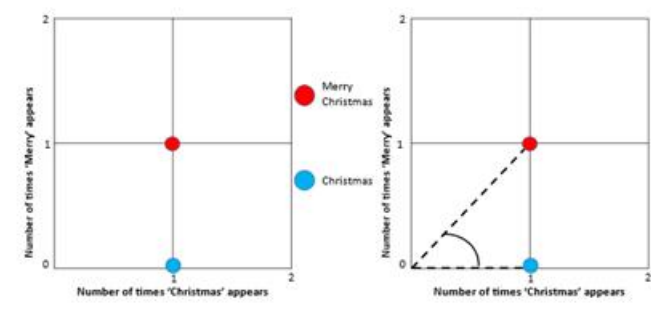


Figure 6.1: Graph depicting word-appearance

If the compared sentences are identical, the two dots would be placed on the same place in the grid. Thus, the lines towards the origin would fall precisely over each

6. LEXICON – JEP ANTONISSE

Therefore, the angle between both lines would result in a cosine score of $\cos(0) = 1$. On the other hand, if both sentences have not a single element in common, the lines would be perpendicular to each other. With an angle of 90° , the cosine similarity would be $\cos(90) = 0$. Thus, in any case, the cosine similarity is a score between 0 and 1, showing the degree of similarity.

For another example, let's compare two words: 'Bert' and 'Ernie'. Instead of words, the vector now can be made of letters. In a table, this would look like this:

Name	E	R	N	I	B	T
Ernie	2	1	1	1	0	0
Bert	1	1	0	0	1	1

Figure 6.2: Letter frequency in the words "Bert" and "Ernie".

With different letters occurring, it would be possible to place both 'Bert' and 'Ernie' in a six-dimensional grid with axes pointing to the origin. However, it is impossible for humans to visualize a six-dimensional graph. We need a new way to calculate the angle between vectors. Luckily, there exists a formula to compute cosine similarity.

1. COMPARISON METHODS AND LANGUAGE FAMILIES

$$\text{Cosine similarity } A, B = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

In less mathematical terms, what this means is that each element of the two vectors A and B are compared. The products of all the elements are then summed up and divided by the length of both the vectors together. If we fill in the numbers for our Bert-Ernie-comparison, it will look like this:

$$\frac{1 \times 2 + 1 \times 1 + 1 \times 0 + 1 \times 0 + 0 \times 1 + 1 \times 0}{\sqrt{2^2 + 1^2 + 1^2 + 1^2 + 0^2 + 0^2} \sqrt{1^2 + 1^2 + 0^2 + 0^2 + 1^2 + 1^2}} = 0.56$$

Nevertheless, converting the words based solely on letter frequency inadvertently results in losing vital information about the arrangement of the letters. This observation is of utmost importance for our project since we try to find patterns and therefore adjacent letter combinations. To address this concern, we introduced a slight adjustment to the regular cosine similarity, where each letter pair with the same character in both words also scores a full point. In this way, our cosine similarity tries to reward words that have the same letters on the same place.

Language families

of various comparison methods, similar to the similarity, allowed linguists to identify groups of languages. These groups, or language families, categorized based on common linguistic features shared common ancestor (Campbell, 2018). Such as Proto-Language allows researchers to trace the relationship of various languages to a single root. However, it is not necessarily need to be the case. There exist languages for which it is impossible to classify them as part of a language family, such as Basque (Campbell, 2018). Researchers speculate that it might be possible that these languages, known as language isolates, might have been related languages in the past, that went extinct. Therefore, these languages now form their own language family, with them being the only member. On the other hand, not only genetic proximity between languages is enough to be placed in the same language family. Languages that are constructed instead of naturally developed cannot be considered part of any language family, since they do not have a shared ancestor with any other language (Campbell, 2018).

Although this means that the total number of lan-

1. COMPARISON METHODS AND LANGUAGE FAMILIES

Language families in the world might be in the hundreds, not all are equally relevant today. To begin with, 94 language families are extinct, meaning there is a lack of any surviving speakers (Campbell, 2018). In addition, the number of languages and the number of speakers differ largely. There are five language families that can be considered the main language families of the world. Every single one of these languages contains at least 5

The most widely spoken language family, with over 3 billion speakers worldwide today, is Indo-European. When Sir William Jones first spoke of this family, he proved there were several branches with related languages (Watson, 2011). First, there is *Indo-Iranian*, spoken in the middle east, where the languages Sanskrit, Persian and Urdu are placed. Secondly, *Italic*, with languages such as Latin, Italian and Spanish. Third, the languages of the northern parts of Europe were placed in the *Germanic* branch. A fourth branch called *Celtic* housed the languages of the island of Great Britain, such as Irish and Welsh. Lastly, Jones portrayed one branch on its own, *Greek*. Only after thirty years would this division be corrected, when researchers added three more branches to the family. The largest new branch was called the *Balto-*

R 6. LEXICON – JEP ANTONISSE

branch, including languages such as Russian, Ukrainian and Czech. The two remaining branches both only had one language, *Armenian* and *Albanian*. This tree remained the same to this day, except for the addition of two branches with extinct languages discovered in the first half of the 20th century, called *Anatolian* and *Tocharian*.

The Sino-Tibetan language family is, even though it has more daughter languages than Indo-European, the second most spoken family, with around 1.3 billion speakers. The family can be split up into two major subgroups: *Sinitic* and *Tibeto-Burman* (Shafer, 1955). The Chinese language is like a family on its own, made up of different regional dialects. The largest and best-known ones are Mandarin and Cantonese. Tibeto-Burman can be split into three further branches: Tibetic (Tibetan), Burmese-Lolo (Burmese and various Lolo languages) and

Niger-Congo is the third most spoken family, containing the highest number of languages: over 1,500 languages. The known ancestors of the Proto-Niger-Congo language are. Although this number is nearly twice that of Indo-European, it is spoken by 600 million people, due

1. COMPARISON METHODS AND LANGUAGE FAMILIES

The immense language diversity in the Sub-Saharan Africa region (Heine et al., 2000). The largest branch of the Niger-Congo family is called the *Atlantic-Congo* branch. Herein are numerous languages spoken in West Africa, such as Yoruba and Igbo. Also, Swahili, mostly spoken in the Eastern part of Africa, falls into this category. The languages spoken in the Central and Southern parts of Africa are mostly from another branch, called *Bantu-Congo* branch. These are languages such as Zulu, Xhosa and Shona. Other branches are the *Kordofanian* branch (Katla, Moro and Talodi) and *Mande* branch (Mandinka, Soninke, Mbaraka, Mandinka and Soninke)

Austronesian, with a similar high diversity as Niger-Congo, covers the languages found in the region that stretches from Southeast-Asia to the Pacific Island. In total this family contains more than 1,200 languages and spoken by approximately 326 million speakers, mostly in countries such as Indonesia, Malaysia and the Philippines. The most important subgroup within this family is the *Formosan* branch, forming a total of nine distinct branches (Tryon, 1995). These branches are all made up of the different indigenous languages of Taiwan. None of them are, however, the most widespread or diverse

R 6. LEXICON – JEP ANTONISSE

of the family. This is namely the tenth branch, as the *Malayo-Polynesian* branch, encompassing Indonesian, Javanese and Sundanese.

Finally, the languages mostly spoken in the North and East of Africa and Southwest Asia are grouped in the Afroasiatic language family. This family consists of many branches (Huernergard, 2004). The branch that contains the best-known languages is the *Semitic* branch, including Arabic, Amharic and Hebrew. Another large branch is *Berber*, with languages such as Tamazight and Tachelhit. Smaller branches are the *Cushitic* branch, which includes languages such as Oromo, Somali and Afar, and the *Chadic* branch, with as largest language Hausa. The languages in the Afroasiatic family combined are spoken worldwide by almost 600 million people.

Using cognates to generate words

By Otwinowska and Szewczyk (2018) argued that words that are similar sounding words with the same meaning in different languages, are the easiest words to learn when learning a new language. The resemblance with your mother tongue makes the words much easier to re-

6.2. USING COGNATES TO GENERATE WORDS

number and use then non-cognates words. By designing Atlán in a way that it has a lot of these cognates, we try to keep the trouble of learning Atlán as low as possible. To achieve this goal, it is key to make new Atlán words resemble existing words, or patterns in existing words, as much as possible.

The idea of using cognates to generate new words in a vocabulary is also used in the creation of the constructed language Lojban (Cowan, 1997). Lojban produced new words, or ‘gismu’s’ and looked for words that sounded similar to it in the languages Chinese, English, Finnish, Hindi, Russian and Arabic. If three or more letters were the same and in the same order as a word in the source language, the gismu would score points. For resemblance with larger language a gismu could score more points, meaning that large languages were viewed as more important. The amount of influence each language had, in other terms the ‘weight’, was solely based on the number of speakers in 1985.

For Atlán we have built a similar program, which we will call *Lexi* from now on. To understand how *Lexi* works, we will try to split the process into three parts: the language selection, the weights, and the program itself.

R 6. LEXICON – JEP ANTONISSE

Language selection

Decision made by the developers of Lojban to use the most common languages was good in terms of significance. The language set closely resembles the set of UN languages: Chinese, English, Spanish, French, Russian and Arabic. These languages are already for 49,6% of all people their mother tongue or second language and official language for more than half the states in the world, according to Ethnologue. However, the developers failed to take language families into account. This is evident from the facts that four out of the six languages used, Chinese, English, Spanish and French, are a descendent of Proto-Indo-European, while other large families such as Niger-Congo or Austronesian are not represented at all. Distributions so far from the real world might make the result very European. This creates a large group of language learners who are unable to match any words to their native language. In order to improve on this, the number of languages used as a source must be increased.

If the desired distribution should resemble the distribution of the real world, we need to know what the distributions in the real world *are*. The frequency of each language family in the 100 most spoken languages ac-

6.2. USING COGNATES TO GENERATE WORDS

ling to Ethnologue (2022) can provide a target percentage of how big the part of each language family should be in our program.

Now we will create a *language set* or *data set*, with in all the languages we want to find cognates in. It is important that the cognate and the Atlan word has the same meaning in all these languages: otherwise, it might be similar looking words, but with different meanings in different languages, which are known as *false cognates*. These cognates are not a sign of a common ancestor but rather a display of randomness and luck. Also, these false cognates are the most difficult words to learn in a new language, even more difficult than non-cognate words (Otwinowska et al., 2018). Hence, we should avoid including those in Atlan. To do so, we must be able to control the meaning of the words in other languages.

Translation software can get us this control. We will use the public available library called Googletrans (3.0.0). This software supports translation into 107 different languages. Since we desire the same significance the language set of Lojban had, we can analyze which of these languages are present in the list of the 100 most spoken languages. The result can be viewed in these tables:

R 6. LEXICON – JEP ANTONISSE

er	Language	Number of Native speakers in Millions	Number of Total speakers in Millions	Language family of the language
	English	379	1132	Indo-European
	Mandarin Chinese	918	1117	Sino-Tibetan
	Hindi	341	615	Indo-European
	Spanish	460	534	Indo-European
	French	77	280	Indo-European
	Standard Arabic	108	274	Afro-Asiatic
	Bengali	228	265	Indo-European
	Russian	154	258	Indo-European
	Portuguese	221	234	Indo-European
	Indonesian	43	119	Austronesian
	Urdu	69	170	Austronesian
	Standard German	76	132	Indo-European
	Japanese	128	128	Japanic
	Swahili	16	98	Niger-Congo
	Marathi	83	95	Indo-European
	Telegu	82	93	Dravidian
	Western Punjabi	93	93	Indo-European
	Tamil	75	81	Dravidian
	Turkish	69	80	Turkic
	Korean	77	77	Koreanic
	Vietnamese	76	77	Sino-Tibetan
	Javanese	68	68	Austronesian
	Italian	65	68	Indo-European
	Hausa	44	63	Afro-Asiatic
	Thai	21	61	Kra-Dai
	Kannada	44	56	Dravidian
	Filipino	0.125	45	Austronesian
	Polish	40	40	Indo-European
	Yoruba	38	40	Niger-Congo
	Odia	34	38	Indo-European
	Malayalam	37	38	Dravidian
	Ukrainian	27	33	Indo-European
	Sudanese	32	32	Afro-Asiatic
	Zulu	12	28	Niger-Congo
	Igbo	27	27	Niger-Congo
	Amharic	22	26	Afro-Asiatic
	Uzbek	25	25	Turkic

6.2. USING COGNATES TO GENERATE WORDS

Number	Language	Number of Native speakers in Millions	Number of Total speakers in Millions	Language family of the language
39	Sindhi	25	25	Indo-European
40	Romanian	24	24	Indo-European
41	Dutch	23	23	Indo-European
42	Pashto	21	21	Indo-European
43	Xhosa	8	19	Niger-Congo
44	Malay	16	19	Austronesian
45	Khmer	17	18	Austronesian
46	Afrikaans	7	18	Indo-European
47	Sinhala	15	17	Indo-European
48	Somali	16	16	Afro-Asiatic
49	Cebuano	16	16	Austronesian
50	Kurdish	15	15	Indo-European
51	Azerbaijani	14	14	Turkic
52	Czech	11	13	Indo-European
53	Greek	13	13	Indo-European
54	Kazakh	13	13	Turkic
55	Swedish	10	13	Indo-European
56	Hungarian	13	13	Uralic

Table 6.3: Overview of the languages and their number of speakers according to Ethnologue (2022).

sume that this entire set of 57 possible languages be-
comes the dataset, called set α . Then it is possible to
get the frequencies of each family in the α set. Since
we want to compare these numbers relative to the total
count of languages, we need to convert these frequen-
cies to percentages by dividing them by the total num-
ber of languages in set α , which is 57. Now it is pos-

R 6. LEXICON – JEP ANTONISSE

compute the distance between the current percentage and the target percentage by taking the absolute value of the target number minus the current percentage. This gives us the error rate. So, for Indo-European, the error rate would be $|42 - 43.9| = |-1.9| = 1.9$, meaning that is 1.9 percentage points away from the target percentage. We can do this calculation for every language family, and the result can be found in the fourth column of table 6.5. The error rate averages to an average error rate of 2.76. Meaning on average each language family is either 1.58 percentage points too large or too little. This is not a bad score, but it is possible to make this error figure smaller by adding or removing some languages to counterbalance.

The Indo-European language family is quite very much over-represented, with almost a language from each branch of the family, and a very similar language present. Thus, we have 25 languages untouched. We want those 25 languages to make up for 42 percent of the set, thus we want 25% of the dataset to be around $(24/42 \times 100) \approx 60$ languages. With the current 57, we should be able to add 3 more languages.

However, there is one language family that is far too under-represented. Almost all the languages in the top 100

6.2. USING COGNATES TO GENERATE WORDS

in the Austronesian family made it into the database, while they should be less frequent than Afro-asiatic and Niger-Congo. Therefore, we remove one language from the language family: Malay. Even though there are several less spoken Austronesian languages, the older common ancestor between these languages (Tryon, 1995) ensures that these may contain more vital information about the group of languages not seen in the data. The only exception is Filipino, which is a language that is derived from the already present language Tagalog, meaning they are also very similar. The choice to let Filipino stay is due to the interesting fact that it has much more speakers than a lot of languages, even though it has a relative small number of native speakers. This aspect of the language might be a good contribution to the desired ‘easy-to-learn-aspect.’ This reduces the number of present languages to 56, so we can add four new languages.

The language-family with the largest error is Sino-Tibetan. There is only one language that could be seen in Sino-Tibetan, although not all linguists would agree. Vietnamese is classified as part of the Hmong-Mien languages. Most Chinese scholars have accepted that it is part of the Sino-Tibetan family (Matisoff, 1991). Although linguists

R 6. LEXICON – JEP ANTONISSE

of Europe have a narrower view of Sino-Tibetan, at least agree that the Hmong-Mien languages are heavily influenced by Chinese languages. Therefore, adding Hmong to the data set and counting it as a Sino-Tibetan language. Even with Hmong added the Sino-Tibetan family seems underrepresented. However, we must keep in mind that a lot of the languages in the 100 most spoken languages are a form of Chinese, being closely related to Mandarin, which is present in the

there are three languages to add left for the other underrepresented languages: Niger-Congo and Afroasiatic. Afroasiatic has a higher frequency in the 100 most spoken languages, but they are currently both equally underrepresented. This means we should give Afro-asiatic two exchanges and Niger-Congo only one.

For Afroasiatic we can translate into Hebrew and Maltese, which are Semitic languages. Thus, we don't need to make any choices here. In the Niger-Congo family we choose between Shona, Sesotho and Chichewa. Since they are all in the same branch, we choose the one with the most speakers, which is Chichewa.

6.2. USING COGNATES TO GENERATE WORDS

Language	Number of Native speakers in Millions	Number of Total speakers in Millions	Language family of the language
Mong	8	8	Sino-Tibetan
Hebrew	7	9	Afro-asiatic
Maltese	0.5	0.5	Afro-asiatic
Ichewa	9	9	Niger-Congo

Table 6.4: Information about the new languages chosen for the dataset.

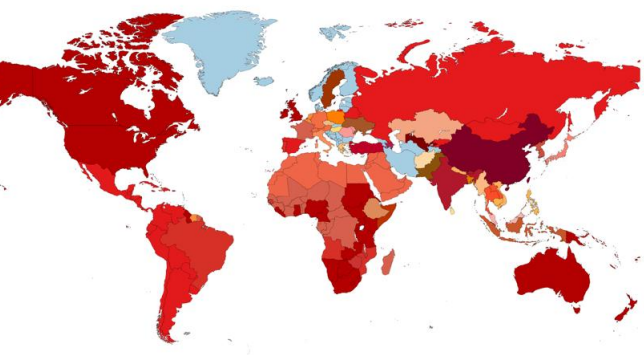
With this modification to set α , we have a new set of languages, we can call language set . By calculating the error again for each language family, we can see that the error measure now averages to an amazing 1.83. Meaning that on average a family is 1.1 language off from the real distribution. These are distributions that are very realistic to the real world.

Language family	Frequency in the 100 most spoken languages	Frequency in the α set	Percentage in the α set	Error in the α set	Frequency in the β set	Percentage in the β set	Error in the β set
European	42	25	43.9	1.9	25	41.6	0.4
Afro-Asiatic	15	5	8.8	6.2	7	11.6	3.4
Niger-Congo	12	5	8.8	3.2	6	10.0	2.0
Indonesian	9	8	14.0	5	7	11.6	2.6
Sino-Tibetan	9	2	3.5	5.5	3	5.0	4
Turkic	4	4	7.0	3	4	6.7	2.7
Dravidian	4	4	7.0	3	4	6.7	2.7
Japanese	1	1	1.8	0.8	1	1.7	0.7
Uralic	1	1	1.8	0.8	1	1.7	0.7
Coreanic	1	1	1.8	0.8	1	1.7	0.7
Austro-Dai	2	1	1.8	0.2	1	1.7	0.3

Table 6.5: Frequencies of languages families in different language sets

R 6. LEXICON – JEP ANTONISSE

ly represent the diversity within this set, we provide with the following map. Each non-blue country figure is associated with at least one official or language present in the dataset. The handful of countries indicate the absence of certain languages.



5.2: World map to visualize where languages present in the dataset are spoken according to Ethnologue (2022) and WALS (2013), made with mapchart.net

Further examination of the blue regions reveals that Scandinavian, Balkan and Baltic countries predominantly fall into this category. However, it is important

6.2. USING COGNATES TO GENERATE WORDS

notice that although the language self is not in the dataset, there might be one that is very closely related to it. The languages spoken in Denmark, Norway, Sweden, Finland and Iceland share such close relationships. Through cross-border communication, individuals from these regions often just continue using their respective languages (Gooskens, 2007). In the same way, languages in the Balkan region, such as Bosnian and Albanian, share a relatively recent common ancestor with Romanian (Kushlevich, 2015). This means they still share a lot of vocabulary and grammar. Similarly, the Baltic states exist, although there are fewer resemblances, notable similarities with languages such as Hungarian. Hence, these languages are not entirely absent of the data set.

We can do the same observation for the blue countries outside of Europe. Persian, as the official language of Iran, displays strong connections with Kurdish and Urdu and in lesser terms also with Hindi and Bengali (Watson, 2011). Additionally, due to the high trade during the Mongol empire, Persian has been largely influenced by Arabic and Turkic languages (Perry, 2005).

Consequently, the only countries in the world that show any indication of a Proto-Language connection be-

R 6. LEXICON – JEP ANTONISSE

their national language and our dataset are Greenlandic, Armenian. Greenlandic, being an Inuit-language, Armenian, being a language isolate, show too little similarities with other languages. However, the populations of these countries would be less than 4 million people, which accounts for less than 0.05

Weights

As discussed previously, the source languages of Lojban were native languages of native speakers were deemed more important (1997). Consequently, the scoring system slightly favors languages with higher weights, which seems justified as it benefits a larger number of individuals. However, the approach employed by Lojban only focused on the number of native speakers of the given language: Chinese has a larger weight than English. The existence of second language speakers was completely disregarded.

The obvious difference between a native and a second language speaker is the fact that a native speaker of the language learns while growing up. On the other hand, a second language speaker acquires a language later in life, motivated by factors such as work, tourism or recreation. This is why I am exclusively utilizing native speakers as a metric for

6.2. USING COGNATES TO GENERATE WORDS

Missing the importance of a language, the developers of Lojban missed the crucial information that the secondary speakers provide. The number of secondary speakers conveys a significant insight into both the language's political influence as its reach. (Saville-Troike, 2017). Besides, it can indicate a lot about vitality and the preservation of the language (Grenoble et al, 2005). Moreover, the secondary speaker count offers us more about the ease of learning the language. All of these are factors that are vital for the language Atlan will be.

As a result, we contend that secondary speakers should not be forgotten and should even be given relatively more weight than native speakers. As a result, the number of total speakers is calculated by 0.4 times the native speakers and 0.6 times the secondary speakers. Subsequently, the adjusted totals are then normalized to obtain the final weights. This way, the largest language contributes the most to the product.

3 Lexi and her workings

With consensus on the dataset and the weights, the time comes to explain how Lexi generates words from all the languages. Lexi is designed to take an English word

R 6. LEXICON – JEP ANTONISSE

to start the process. This is very important, because there is only a very select group of unreducible words that need to be generated this way. By taking a word as input, we prevent that it generates words we don't want and control the meaning of the new Atlan

word is then translated into all the languages we use up the trainings set that was just discussed. Now we can find a common pattern in this set of 60 languages. However, it is key to keep in mind that the same letters are often pronounced differently in different languages. For example, the sentence 'the *tear* in my new car brought a *tear* in my eye' contains two words that are spelled identical but sound very different. Such heterogeneity prove it is not enough to compare words only by their spelling. To investigate how words are pronounced, we transform every translation into their IPA for-

the constraints of comparison were a bit hard. With this modification, it is possible to soften them a bit. That is, because Atlan reduces the number of words drastically. This has as a result that certain words that are viewed differently in our languages, can

6.2. USING COGNATES TO GENERATE WORDS

refer to one and the same letter in Atlan. That means this sound can be viewed as the same in comparison. To achieve this, Lexi can map each IPA International Phonetic Alphabet sound to the letter it would represent in Atlan, according to the rules discussed in the phonology chapter. This creates a list with the transcriptions of how the translations would sound in Atlan. The only task that is left now is to find the patterns in this set. Since these patterns will form the new Atlan alphabet, it is certain they will always follow one of three patterns: either consonant-vowel-consonant (CVC), consonant-vowel (CV) or vowel-consonant (VC). We can look for possible candidates for this pattern by splitting each translation up into every possible combination that follows one of these patterns. This way, Lexi provides a subset of candidates of the patterns, made up from small parts of natural language translations.

If we calculate the similarity between every possible pattern and the transcribed translations, we can get the pattern that resembles the words best. However, to simulate the natural way of words generation a bit better, we experiment with some algorithms to make sure we pick the best word, such as evolutionary computing.

R 6. LEXICON – JEP ANTONISSE

is a problem-solving technique that uses the principles of natural selection and genetics to find patterns in data. Just as in nature, the individuals with the highest 'fitness' are selected to create new children. After generation after generation this will mean that the best traits survive, resulting in a solution. By using evolutionary computing, we are trying to simulate the way of one word evolving over time, with only the fittest words surviving. In this case, a 'fit' individual is a word very similar to all the translations. We calculate the fitness with the previously discussed cosine similarity.

The fittest individuals create children through a crossover process. Lexi divides both words in all possible points and swaps both halves. For example, if the words *tek* and *tep* are combined, it can recombine into *j-ek* and *tep*, similarly *jak* and *tep*. These children, combined with their parents, form again a large pool of possible Attributions. The best individuals are again selected, and they produce the next generation. After 50 generations the three individuals with the highest scores are selected and crowned as winners. In the end this process gives us the same outcomes as calculating the high-

6.2. USING COGNATES TO GENERATE WORDS

score for all of them.

In this way, we generated the five best options for every atom of language we needed to exist in Atlan. Another computer program assigned each word one of these options, keeping in mind that some syllables might appear in the list of options for several words. In the end, we are left with a lexicon of words that should be easy to learn and remember for almost the entire population.

R 6. LEXICON – JEP ANTONISSE

semantic atom vocabulary

MORPHOSYNTAX (V,CV,VC)		
MOOD MARKERS		
imperative/ imperative/	O	C
expletive		
negative	E	○
active	U	∪
prosodic)	A	∩
clause (+ pronoun)	I	⊃
LOGICAL OPERATORS		
conjunction (and, ∧)	AN	∧
conjunction (or, ∨)	OL	∨
conditional (if, \rightarrow , <i>ceteris paribus</i>)	IF	\rightarrow
inference (thus, therefore \therefore)	IS	\Rightarrow
negation (not, \neg) - sentential/predicate	NE	\neg
quantifier, -few +many)	SO	\sim
possibility (\diamond)	PE	Δ
necessary (\Box)	SE	\Box
SYNTACTICAL MARKERS		
active (affected thing/person, object, +	EK	\rightarrow
transitive		
have (possession, + verb = to have)	TA	\vdash
(receptive, benifice)	LO	\hookleftarrow
instrumental (tool, method etc.)	UT	\uparrow
nominal (noun, definite object,	NA	\nearrow
name (cartouche))		
(suffix) many	ON	\swarrow
to do <i>x</i>	TU	\downarrow
state (being, identity, attribute,	SI	\curvearrowright
(behind noun) adverbs (behind verb))		

6.3. SEMANTIC ATOM VOCABULARY

TENSE/ASPECT MARKERS		
Past	PA	▷◁
Future	FE	ℳ
Beginning	KA	⊖
Perfective	NI	↗
Progressive	PO	◁▷
Passive	PI	∇
PREPOSITIONS		
Coordinate: at (time/place)	ET	⊙
Left (+earth=west)	LA	ℓ
Right (+earth=east)	TI	ℓ
In front (time:before)	EN	↖
Behind (time:after)	IT	↗
Next to (right and left)	KE	⊖
Above	EF	↖
On	AF	ℳ
Under, below	OT	⊥
Inside (+time=during)	IN	ℳ
Via, through	LE	ℓ
Outside	AP	◁▷
Surrounding	AL	ℓ
In between	MI	↖
Near	KI	⊖
Far	FA	ℳ
Horizontal	IL	↗
Vertical	TE	⊥
Sagittal	SA	↖
Direction (of movement, combine coordinates x to y)	LI	ℓ
Range (until, up to)	TO	ℓ

R 6. LEXICON – JEP ANTONISSE

NUMBERS		
<p>t</p> <p>e</p> <p>red (base 10)</p> <p>12)</p> <p>and (base 10)</p> <p>e 12)</p> <p>ent</p> <p>al</p>	IP	⊖ ⊐
	OP	⊖ ⊐
	UP	⊖ ⊐
	IK	⊖ ⊐
	OK	⊖ ⊐
	UK	⊖ ⊐
	IM	⊖ ⊐
	OM	⊖ ⊐
	UM	⊖ ⊐
	JI	⊖ ⊐
	JO	⊖ ⊐
	JU	⊖ ⊐
	NO	⊖
	NU	⊖
	US	⊖
	OJ	⊖
SYNTAX-SEMANTIC		
line	MA	⊖
ine	FI	⊖
n	EJ	⊖
	JA	⊖
	LU	⊖

6.3. SEMANTIC ATOM VOCABULARY

1st removed: speaker	AM	↘
2nd removed: listener	UN	↗
making context		
3rd removed: beyond context	AJ	↙
speaker		
Part-to-whole relationship	PU	⊕
Final state	FU	↪
Intention	UF	↩
Cause, reason (effect = caused, passive)	KO	⊖
Inverse	EM	↯
Negative	JE	↱
Neutral	UJ	↲
Positive	FO	↴
Equative (same as)	ME	↵
Comparative (more,very)	MO	↶
Superlative (most)	AS	↷
Contrast (than, + relative clause = but)	KU	—○
Self	SU	↷
Other	OF	↶
SEMANTICS - (196-CVC)		
QUALIA		
See	SIK	↵
Colour	KAL	↵
Brightness, light	LAS	↙
Red	EL	↘
Green	OS	↘
Blue	UL	↙
Hear	TIN	↵
Sound	SAN	↵
Volume, loudness	LAT	↵

R 6. LEXICON – JEP ANTONISSE

me, transform pitch t mi ng, affect (+good=valence n) act, touch on re erature ce	PIN	↗
	PIT	˙
	SEN	↘
	TAS	└
	TIT	˙
	SAK	↘
	KIT	└
	SAL	↘
	MUM	˘˘
	SIN	↘
EMOTIONS		
r ipation ise ess st t	NAK	↗
	TIS	└
	SUS	˘˘
	FIS	┐
	TEF	↗
	SAP	↘
	NUS	↘
	KAM	↘
	FIP	┐
	LOF	┐
	KEF	↗
	LEJ	↘

6.3. SEMANTIC ATOM VOCABULARY

CONSCIOUSNESS		
Consciousness, mind	MAK	↖
Awareness, focus	FEK	↗
Hypnosis	NOS	↘
Dissociation	LIS	↙
Trance, depersonalization	JAS	↘
Sleep	SUL	↘
Hallucination	TAN	↑
Depression	TEJ	↗
Stimulation, excitation	JIS	↘
(negative = sedation)		
Think, reason, abstract, logic	SIN	↘
Memory, remember	LIM	↘
Desire, will	FAN	↗
Conscience (moral)	LEN	↘
Intuition, instinct	NIT	↑
Imagination, creativity	NIL	↘
Know (+ think = understand)	NEF	↗
Study	SUF	↘
EXPRESSION		
Say, speak	LAN	↘
Greet	MAL	↘
Ask	NUK	↘
Thank	JEK	↗
Word	SET	↑
Book	POK	↗
Read	LET	↑
Write	LIK	↘
True (yes)	TET	↑

R 6. LEXICON – JEP ANTONISSE

t	NOT	丿
ber	NOM	ㄣ
ure, amount	MEP	ㄣ
c (+ voice = sing)	NAN	ㄣ
	FAS	ㄣ
	LAJ	ㄣ
GEOMETRY AND TOPOLOGY		
ugmentative	PIK	ㄣ
	LAK	ㄣ
	LAM	ㄣ
	TOK	ㄣ
	PIF	ㄣ
, diminutive	SOT	ㄣ
	KOT	ㄣ
ow	NEJ	ㄣ
	PUT	ㄣ
icient (too, enough)	SEF	ㄣ
d	LOK	ㄣ
	FET	ㄣ
	PUT	ㄣ
ght	SIT	ㄣ
	TOL	ㄣ
	KAP	ㄣ
ce	SEM	ㄣ
al volume	NET	ㄣ
	KEJ	ㄣ
	POP	ㄣ
	LUJ	ㄣ

6.3. SEMANTIC ATOM VOCABULARY

PEOPLE		
Young, new	NUJ	/
Old	POL	ㄣ
Partner	JAP	ㄹ
Parent (reduplicate to add generation)	PET	ㅈ
Child (daughter, son)	TAT	ㅍ
Sibling	SIP	ㅅ
Acquaintance (friend, enemy)	TEN	ㅊ
FLORA AND FAUNA		
Nature	NIF	\nif
Creature, organism	NIK	/
Virus	SIF	ㄴ
Bacteria	JK	ㄹ
Archaea	NUL	ㄣ
Amoeba	PAM	ㅅ
Plant	TUP	!
Fungus	KON	/
Animal	NAF	/
Mammal	MAM	ㄴ
Fish	MIS	ㄴ
Bird	PAJ	ㄹ
Insect	KES	ㅈ
Reptile	LEP	ㄣ
Worm	LEM	ㅅ
Tree (+ substance = wood)	TEL	ㅊ
Ecosystem, biome (+ tree = forest)	PIM	ㅅ
Stick, branch	FAP	ㅅ
Fruit(+ fungus = mushroom)	FUT	ㅍ
Seed(+ fungus = spore, + animal = sperm)	NIP	/
Leaf	LIF	ㄣ
Root(+ fungus = mycelium, + shape = rhizome)	LUT	ㅍ

R 6. LEXICON – JEP ANTONISSE

ORAL VERBS		
consume (+ solid = eat, drink) spit, regurgitate smoke the (+ organ = lung, to smoke something) laugh (humor) cut	KOS	𐌺
	JAT	𐌵
	SOS	𐌱
	TUS	𐌸
	FOT	𐌳
	PUL	𐌹
	LUS	𐌽
	LAF	𐌶
	KAJ	𐌴
	NOJ	𐌺
ANATOMY		
(+ tree = bark, + material = leather) flesh, muscle breast (+ organ = breast) mucus (+ mouth = saliva, not)	POS	𐌱
	MAF	𐌶
	KAT	𐌴
	TIL	𐌰
	LIL	𐌶
	KOP	𐌶
	FAT	𐌳

6.3. SEMANTIC ATOM VOCABULARY

Feather	PEF	↘
Hair	PEL	↖
Head	SAT	↑
Face	LES	↙
Organ (+ hear = ear, + see = eye, nell = nose, + taste = tongue)	NOK	∠
Tissue	JUS	↷
Cell	SEL	↶
Mouth	MUT	↴
Lips (+ touch = kiss)	LIP	↷
Cheeks	KIS	↘
Tooth	TUN	↙
Nail	NEK	↗
Foot	PAF	↶
Leg	LUK	↵
Joint (+ leg = knee, m = elbow, + line = angle)	TUJ	↓
Hand	JAM	↘
Finger (+ foot = toe)	FEN	↗
Arm	MAP	↘
Wing	KIN	↗
Belly (+ inside = guts)	PEP	△
Neck	KEN	↗
Shoulders	TOP	↓.
Back	PUK	·-
Chest	SOP	↘.
Heart	JOT	↓
Liver	KAF	↶
Brain	NEL	↖
Genital	NEN	↗
Waste (+ solid = faeces, + liquid = urine)	FES	↖

R 6. LEXICON – JEP ANTONISSE

LIFE AND DEATH		
course, sex	SEK	↗
h (+ active = to birth, = to be born)	PES	↘
v	KEP	+
	FIN	/
ase	TIM	↑
h	MOT	↓
	KIL	↖
t	FAJ	↗
t	JAF	↘
MOVEMENT		
	MAT	↖
	KUT	↑
	TIP	↑
	SUT	↑
atch	SEJ	↗
	TON	↓
e, go	MEF	×
	KIK	—
n	NAS	↗
	FUL	↖






























6.3. SEMANTIC ATOM VOCABULARY

Walk	TOM	ㄴ
Come	KOM	ㄴ
Lie	LOM	ㄴ
Stand	TUF	ㄴ
Sit	SUP	ㄴ
Fall (+ transitive = drop, water = sink)	KOL	ㄴ
Happen, occur (+ transitive = do)	NES	ㄴ
Steer (a vehicle)	JIL	ㄴ
Jump	MOS	ㄴ
Block	FOK	ㄴ
HANDLING OBJECTS		
Give (+ passive = receive)	KIF	ㄴ
Hold	LEL	ㄴ
Squeeze	JIP	ㄴ
Rub	LOP	ㄴ
Wash, clean	SAS	ㄴ
Wipe	TAP	ㄴ
Pull	JAL	ㄴ
Push	PUS	ㄴ
Throw	NOL	ㄴ
Tie	TAL	ㄴ
Sew	SES	ㄴ
Shake, vibrate	JAK	ㄴ
Pick, take	NEP	ㄴ
Make, create	MEN	ㄴ
Find	FUN	ㄴ
Meet	MIF	ㄴ
Hang	NUN	ㄴ
Kick	KUP	ㄴ
Exchange	JEM	ㄴ
Sell	MEM	ㄴ

R 6. LEXICON – JEP ANTONISSE

ch	JOF	↷
ace	NUF	↶
er	LUP	↵
ch	SOJ	↷
	SOF	↶
age	JES	↷
ick, stay	MEJ	↷
n	PUJ	↷
rol	KOF	↶
er	JON	↷
v	MOM	↷
	MUL	↷
	MOK	↷
SUBSTANCE		
er, substance, material	MA	↷
g	MU	↷
ght, heavy	NE	↷
er	NA	↷
id	JI	↷
	FEL	↷
	PE	↷
	TA	↷
ze	LI	↷
	PI	↷
v	SE	↷
l	SU	↷
	JU	↷
	SO	↷
	TO	↷
e	JE	↷
	SA	↷
	TU	↷

6.3. SEMANTIC ATOM VOCABULARY

Air, wind	KE	
Gas	KA	
Sky	SO	
Cloud	NA	
Fog	TU	
Fire	FA	
Burn	PE	
Smoke	TU	
Ash	KA	
Metal	MI	
Glass	KU	
Paper	PE	
Textile	TI	
Rubber	LA	
Foam	MO	
Petrol-based substance (+ solid = plastic, liquid = oil)	PA	
Molecule	ME	
Atom (+ number = chemical element)	TA	
Particle	KI	
Explode	FO	
Shrink	FU	
Pure	PU	
GEOLOGY AND ASTRONOMY		
Sun	SON	
Star	SIS	
Planet	PAL	
Earth planet	TEM	
Night	NAJ	
Day	JAN	
Week	MIK	

R 6. LEXICON – JEP ANTONISSE

tal system (+ sun = solar system, laxy)	LOT	↓
n (+ time = month)	MUN	↘
ntain (+ sand = dune)	MAN	↘
peak	FON	/
n	PAS	↗
nd	NAM	λ
ntry	MES	↗
inent	KUN	/
d	PIL	↗
l	LOL	↘
r	LEF	↘
	LEK	↗
	SAM	↘
QUALITIES		
n	JOL	↘
d	MOL	↘
	POT	.
	PAK	÷
en	LAL	↘
y	JUT	↑
p (+ negation = dull)	TAK	↑
oth, soft	SAF	↘
l, tough	TOF	↓
cult	TIF	↑
	SIL	↘
	SUK	↘
ect	KEK	---
tiful	LIT	↓
	KUL	↘
ng	TUT	↓

6.3. SEMANTIC ATOM VOCABULARY

Quick	KIM	↘
Slow	NIM	↗
High	JEJ	↗
Low	MEK	↘
Real	LEJ	↘
Last	LUM	↘
Previous	POP	▷•
Relevant	FEJ	↗
Silly	NUP	•/
Extra	JEF	↗
CULTURE		
Home, domestic	KEM	↘
School	SUJ	↗
Town, city, village	TOS	└
Culture, customs	KUK	--
Job, profession	KAN	↗
Money	MON	✓
Politics	NIN	↗
Tradition	NOP	↗
Individual	FIT	└
Community, collective	SUM	↘
Group	KOK	—
Science	NIS	↘
Law	LON	↘
King, chief, president	TES	↘
Medicine	MET	↘
Machine	MAS	↘
Vehicle	FIL	↘
Wheel	MIL	↘
Building block	FUK	└
Layer	FEM	↘

R 6. LEXICON – JEP ANTONISSE

he	NUT	ㄷ
ricity, energy	MIN	ㄴ
outer	KUM	ㄴ
em	JIM	ㄴ
ing	NAP	ㄴ
	MUK	ㄴ
	MIP	ㄴ
ainer	NON	ㄴ
	JEL	ㄴ
e	PUN	ㄴ
lem, Conflict	POM	ㄴ
ake	KUF	ㄴ
truction	JUN	ㄴ
re, image	SIM	ㄴ
ool	MIM	ㄴ
y	FAM	ㄴ
	JOS	ㄴ
age	LUL	ㄴ
hes	KEF	ㄴ
iture	MUP	ㄴ
metry	FIM	ㄴ
ection	FEP	ㄴ
plexity	KOJ	ㄴ
ess	JUK	ㄴ

all syllables were used, since we have 490 op-
total, there are 42 additional words which weren't
l can be used when additional words are added

6.4. PROTOCOL FOR TRANSLATION

Protocol for translation

n's lexicon is composed by adding together different semantic atoms (see: oligosynthesis). Meaning is specified by following the principle of a "semantic lens" which "zooms in" with every atom that is added, reducing the possible meanings of a word to a more defined meaning. The order in which atoms are added is structured hierarchically. The first atom specifies the most basic classification of the intended word, to which other atoms are added which incrementally specify the exact definition. Two combinations of the same atoms that are ordered in different ways (e.g. AB and BA) will thus have different definitions. For example:

Bird + house = a pet bird (a bird pertaining to the domestic)

House + bird = a bird's nest (the house of a bird)

Flower + smell = a fragrant flower (a flower which smells)

Smell + flower = a floral scent (the smell of a flower)

R 6. LEXICON – JEP ANTONISSE

g + community = to sing in a choir (to sing communally)

community + sing = a choir (a community of singers singing)

Phosyntactic atoms can be recognised by the fact
y contain one one consonant and syllable (not
these are always put in front of the main semantic
again following the principle of hierarchy, and
to account the scope of each atom. Any atom
termines the scope of the atoms that come after
the ones that come before it. For example:

ssive + give = to receive X (to be given X: to receive a gift = to be given a gift)

ssive + give + (dative) = to be given to X (the gift is given given to them)

edicate + not + life = (to be) inanimate

t + predicate + life = not (to be) alive

6.4. PROTOCOL FOR TRANSLATION

ally, the plural marker comes at the very end of a word.
general order for compound words it thus:

morphosyntactic markers - main semantic root -
specifying semantic atoms – plural

It must be noted that translations are open to individual and cultural interpretation. As long as the hierarchical structure of word formation is obeyed, different approaches to specifying the same word are possible. For example:

Fish + fly = flying fish (a fish which flies)

Fish + wing = flying fish (a fish with wings)

Vehicle + sea = boat (a vehicle for the sea)

Vehicle + float = boat (a vehicle which floats)

Because of their sheer quantity and diversity, words for specific human artifacts can sometimes be more challenging to synthesize. Phonetically approximated loanwords can be employed when referring to specific cultural artefacts or concepts (see chapter 3.7), marked by

R 6. LEXICON – JEP ANTONISSE

the particle NA, together with a cartouche in writing language.

Expletives, better known as swear words, can be made more explicit by the prefix 'O' \subset to mark the exclamatory nature followed by a literal translation of the word, which may be culture dependent (e.g.: 'shit' = O.FES.TOJ (fess-toj)).

Jargon translation

An established lexicon will not be enough to cover the quantity of topic/discipline-specific jargon in a language. Specialists within their respective fields should formalize nomenclature systems which could then be adopted by their respective intellectual communities as a standard for communicating about their subject in Atlan as an auxlang. This could be a natural evolution within the arts and sciences away from often obscure and culturally imposed greco-latin terminology of the European academic tradition, as well as serving as an auxiliary language for scientists who often complain about ill-defined and arbitrary terminology within their subject to help design a rigorous reform towards more logical

6.5. JARGON TRANSLATION

and useful words in a more neutral lingua franca. We cannot dictate in this context how each discipline end up standardizing their terminology in Atlan, but could suggest taking inspiration from original etymologies and nomenclatures already in place as standardized systems. We will not further expound on how should be done here, since this is beyond our own expertise, and almost impossible to cover comprehensively.

Chapter 7

Semantics — Jonathan Roose

What exactly is meaning? How does language express meaning? These are some of the most central questions in the philosophy of creating a language. When creating Atlan, we tried to create a meaningful set of signs, meaningful in that they express people's intentions. We tried to do this by creating a vocabulary, words were created — but they were not generated — with a given, unchanging meaning. So, is that how meaning works? Do not words mean because other people understand them to mean those things. Meaning is inevitably tied to use. How can we then create meaningful words if nobody uses them?

These are very fundamental questions about language, though questions, and questions that inevitably, Atlan will have to deal with. We went into this project with the belief that meaning of words can be given from above, that words have meaning because the dictionary shows they do. In Atlan, to give meaning to signs, we have differentiated between language and speech; separated the linguistic code and the daily utterances. Following a structuralist understanding of language we separated meaning as depending on two things as Ferdinand de Saussure argued (Saussure 1959); language is (1) the linguistic code, this is the structure of grammar and syntax, the meaning of words as you find them in the dictionary; (2) how people use the language in a certain context — what people *do* with language, i.e. to order something or to begin a conversation (What does “hello” actually mean?) — what the linguist John Austin has called speech acts (Austin 1955)— In linguistics this second facet of meaning in language is the focus of the subdomain pragmatics. This subdomain hopes to answer how intention, speech and language interact and create meaning and understanding between speakers. This chapter will engage with pragmatics in the creation of Atlan.

R 7. PRAGMATICS – JONATHAN ROOSE

important to engage with Pragmatics when creating a conlang because, although certain structures in a language will indicate certain things grammatically, what a language does with language is eventually what makes the language a language. A conlang might have a very thorough structure of grammatical rules but how do you use it? Historical examples of conlangs that do not engage with the question of speech are legion, for example Leibniz's attempts of a perfect language based on a clear and logical structure in the end became calculus (Eco 271), a highly perfected 'language' but nearly impossible to use in conversation in. Also, we must know whether the conlang leads to clear and meaningful understanding and not that there is something of vital importance about that thing we call language that we have over-engineered. How meaning in language is expressed, in the end, depends on the speakers, what they intend to do with the language. Whether that be communicating information or emotional expression.

We cannot separate meaning from use, otherwise beautiful linguistic phenomena like metaphors, metonymies and idioms would only be false or incorrect. Poetry is a net of lies and falsehoods. It is clear there is

inherent link between the meaning of language, what linguists call; semantics, and the use of language, that pragmatics. As linguists like, for example, Gennaro Chierchia has shown these two facets cannot be understood separate (e.g., Chierchia 2012).

Esperanto, of course, is made with the goal of a language that can function as an international lingua franca. A language to assist speakers of different languages to communicate. It has a very practical goal. However, people do not seem to use language as the grammarians want. One might ask, how does Esperanto with phenomena like curse words; they depend not only on the semantic meaning but also on whether they *sound right*, express the feeling that is intended. “Holy cow” for example, does not mean what it means semantically, it expresses shock and confusion. In Esperanto we might have a way that to describe how to semantically describe a “holy cow” but how do we express the feeling of shock in the same way as in English? Even if we can make the meaning explicit, telling exactly what the feeling is but will that truly express the feeling? Is telling someone you are expressing the same thing? Another problem is that Esperanto is never finished, it is made anew by how its speakers use it every day. Every speaker is a language

R 7. PRAGMATICS – JONATHAN ROOSE

making a new language to express their experience and not merely to describe them. This is a problem for our ambitions with Atlan because, it will mean to connect with the grammatical rules and the daily use of language that way would soon fall apart, every language with its own version. This is what the writer Umberto Eco in his book about conlang called: “the inevitable Babel effect” (Eco 323), named after the biblical story of the tower of Babel. In the story humans in the beginning speak only one language and build a tower as a monument to themselves. However, soon they find that their speech is confused with each having a different way of speaking. The inescapable Babel effect is the seemingly unavoidable confusion that is a result of people using language in their own way. This effect puts a wrench in our ambitions with Atlan. The only thing we can do in this project is to create a linguistic structure that, if used correctly, would lead to easy and clear communication between speakers. The conlang searches for a *prescriptive* linguistic code. If people really use it, the conlang cannot create per-

Initially the linguistic theories of the iconoclastic linguist Roy Harris are used on this point

7.1. IMPLICATURES

speech. Pragmatics cannot be prescribed; it cannot be perfected; it is about how people *use* language not about how people *ought to* use language.

However, what we can do is to build implicit meanings into the semantics of Atlan to make the language as clear as possible no matter how people use it. To do this Atlan makes explicit what the intention is of the sentence, as far as this is possible. This chapter, on the one hand, will look at a linguistic theory about intention and speech. Afterwards it will look at how language changes depending on its use. How language relates to culture and its speakers. Whether Atlan can escape the Babel effect. I will finish with a synopsis of how language should be understood and if Atlan can be used practically.

Implicatures

In linguistics the term implicature means the implicit intention that a speaker has with an utterance. (Davis 1985) An implicature being what the speaker intends to say or say when speaking. In linguistics this term is useful as it describes the meaning that a speaker puts into a word. How meaning in language is achieved depends on what the speakers implicature is with an utterance.

Prosody

In all language the non-explicit markers of speech such as pitch, stress or rhythm might be used to indicate what the speaker's intention is behind an utterance (Wichmann 2009). These are called prosody; intonations, stress and rhythm can indicate the intentions of the speaker and can carry meaning. However, in Atlan the decision has been made that prosodic markers should have no semantic or pragmatic value². Although, this is not completely true as Atlan still uses stress markers, but prosodic markers are for the most part meaningless. Regardless to the confusion and ambiguity meaning that in English grammatical but based on prosodic markers are not grammatically marked. The five vowel syllables are used for these effects:

! = Exclamative (prosody), imperative, vocative

? = Interrogative (question, prosody³)

+ stress = Stress marker (prosody)

²More on this topic Niek's chapter will suffice.

³The letter *e* can also be used as a so-called 'filler word', like 'ehmm'

7.1. IMPLICATURES

≈ i (+ pronoun) = Relative clause

≈ u = Subjunctive (wish)

Atlan vowel that is most like the English vowel “a”
stress marker, indicating that the word is important
the *implicature* of the sentence. The equivalent in En-
h script would be to write a word in italics. (Take the
erence between the sentence “Are you going to *the*
ma?” and “Are you going to the cinema?” the first
questioning the location and the second questioning if
it is true or not that you are going to the cinema.)
ause italics is impossible in the Atlan script an extra
will be needed, likewise in the spoken language is
stress marker already used to indicate the core of the
d and thus a vowel marker can be used to for fill the
e function that stress has in a language like English.
What these prosody markers show is that when speak-
the message is more than just the sum of the words.
ry utterance has implicit information that the listener
only understand by placing it in the context of the
versation or by non-linguistic signs like laughter and
y language. How Atlan speakers communicate the
ormation that in a natural language like English would

R 7. PRAGMATICS – JONATHAN ROOSE

with prosody is though the grammar of the language. However, to communicate does not necessarily mean you are using language (It might also be said that using language does not necessarily mean you are communicating). By using these explicit markers Atlan can help speakers to make their intentions clear in a way that would otherwise be impossible or ambiguous.

Intentions in speech

As we have seen, Atlan incorporates much of the intention of the speaker into the grammar, making explicit what in natural languages was implicit. (This is not to say that these markers are not seen in natural languages; for example, English does have a stress marker “ $\gamma\epsilon$ ” similar to Atlan’s $\bigcirc \approx a$ (Liddell 1894:301)) Yet it is simply not viable nor desirable to make every meaning that a speaker might carry explicit into the grammar. Atlan is a language; it needs to be interpreted; not decoded.

Figures, by definition, cannot be incorporated into the semantic structure. Nevertheless, a language can make it crystal clear what is meant with an utterance. Ambiguity of meaning is something we are attempting to avoid. However, what a speaker might in-

7.1. IMPLICATURES

What we have done with Atlan is to divide the possible intention within speech into different uses of language. This way we can make sure that all the possible intentions of an utterance can be expressed in Atlan. The classification is based on the understanding of language of the linguist and literary theorist Roman Jakobson. In Jakobson's theory the meaning of speech depends on six possible uses that an utterance might have, based on six factors that are the most important in understanding speech. These factors are: (1) the speaker (ADDRESSER), (2) the listener (ADDRESSEE), (3) the utterance that carries (3) a MESSAGE, (4) a CONTEXT in which it is uttered, whether there is (5) CONTACT between the listener and speaker and (6) the linguistic CODE, meaning the grammar and lexicon the interlocutors can understand. These factors can be visualised like



Figure 1: Jakobson's factors of meaning in speech (Jakobson, 2018, 70)

R 7. PRAGMATICS – JONATHAN ROOSE

utterance all six factors are present. However, stating which factor has the most importance is what the recipient understands the intention of the utterance to be. Connecting the function with the most important factor we get a schema of the six different ways to use language:

REFERENTIAL

EMOTIVE POETIC PHATIC CONATIVE

METALINGUAL

2: *Six different uses of language (Jakobson, 2018, p.1074)*

the most important factor is the addresser (so the speaker), the intention behind an utterance becomes what means that what the speaker is feeling or thinking is most important to communicate. Take, for example, the sentence: “It is raining”⁴. When the intention is to use language in an emotive way “It is raining” is a dramatic expression like one might find in a romantic poem. An expression of sadness and gloom ‘it is raining in my soul.’ The reader must understand this sentence as expressing what

⁴ Examples are found in the lecture on Jakobson by Paul Hirsch (2009).

7.1. IMPLICATURES

speaker is feeling.

When the factor of the addressee becomes central the *conative* function is most important. This is when the utterance becomes an implicit command, for example a mother seeing a child go outside without a jacket might say “It is raining” meaning a command to put a jacket on. When the utterance is focused on the *context* around the speakers the function becomes referential. This is when the weatherman says, “It is raining,” merely saying the factual state of the nature around speakers.

When an utterance is intended to establish contact (the “Can you hear me?” and “Hello” of language) the function of speech is *phatic*. Take for example the scene of two awkward young people on a date, both are awkwardly silent and when one of them says “oh, it is raining.” the speaker does not really care whether it is raining or not, the utterance is simply meant to establish contact. The *metalingual* function is the capacity of language to talk about itself. It is Language to correct and explain language. Like how I have been using the sentence “It is raining.” for example but also, questions like “What do you mean with “it” when you say, “it is raining?”” (usually, a very puzzling question)

Lastly there is the *poetic* function, the function that targets the message of an utterance. This might mean the form or rhythm of an utterance or the combination of different

R 7. PRAGMATICS – JONATHAN ROOSE

in witty similarities. A good example our supervisor is: “it is raining bullets.” What is important in the function is the relation between speech and message, hence is calling attention to the how and why language instead of selecting the proper word a speaker commands. Similar to the metalingual function the poetic focuses of language as a semiotic system. However, the poetic function and metalingual function are in diametric opposition to each other; the metalanguage function is how the sequence of words is used to build an equation (‘it rains falling water’ for example) whereas in the poetic the equation is used to build a sequence (‘look how the drops of rain resemble bullets’)⁵.

are the six functions that language has. One way we have a crystal-clear language is by having the speaker exclaim that an utterance has one of these six functions. A marker to indicate the function. However, besides being inelegant, that would lead to ridiculous sentences. That, it is an international auxiliary language and was intended to replace natural languages, merely to work to easily communicate with speakers with different languages. These functions in language are unavoidable

son explains this with the difficult sentence of: “The position projects the principle of equivalence from the axis of content into the axis of combination.” (Jakobson 2018:1074)

7.2. CULTURE AND LANGUAGE

The aim of Atlan is to make a language that is semantically unambiguous and simple. Atlan, as it is now, is a language whose makeup is heavily tilted to utterances that referent the world as it presents itself for it creates words based on the speaker's empirical data of the world. How people can understand each other in Atlan is through understanding the words and collections of basic axioms, axioms that are experienced in the world. To say "it is raining" in the emotive sense in Atlan you would be better off by arriving to that emotion by the emotive axioms of Atlan. Saying "it is raining" in all the different ways as described above clashes with the aim to counter ambiguity in the language. You have to say what you mean to be able to form the words of Atlan. Thus, the language, in the way it forms its words, is always referencing the experience that such a concept entails.

Culture and Language

Max's introduction already discussed it is difficult to separate culture and worldview from language. Nevertheless, with Atlan we are attempting to do exactly that, in the name of a neutral language. In relation to culture Atlan set out to achieve two things: (1) Atlan needed to be independent to any dominant culture, otherwise it would be no better than English as a neutral lingua franca; (2) cultural expression and lived world

R 7. PRAGMATICS – JONATHAN ROOSE

to be able to be translated into and even be able to be expressed in Atlan. As Max also discussed the relation between language, country, politics and worldview is a very complex topic and a cause for problems in any auxiliary language.

On the one hand, Atlan will need to engage with the world within a specific cultural milieu. But on the other hand, Atlan hopes to be able to evade being tied up to any specific cultural expression as much as that is possible.

Unconsciously made Atlan as culturally neutral as we could manage. We tried to give the language enough scope to be able to express all kinds of different worldviews and culturally specific concepts and items. It is after all a language that hopes to transcend linguistic and nationalistic divides, this means that it cannot make different culturally specific expression understandable to all speakers. As we have discussed in 7.1 previous sections, Atlan is a language that should focus on the referential function of language. It is a language that tells you the facts as they are to the speaker. This might limit the speaker's cultural expression of the world. What is more important is that people can understand each other, even if that would limit their cultural expression ability. Here we see why culture is impossible to separate from language; already in Atlan there is a hierarchy where comprehension is more important than expressions. A language without culture is an impossibility and the world without speakers, if language has influence on that, will still

7.2. CULTURE AND LANGUAGE

influenced by Atlán just like any other language would do. If Atlán changes its speaker's way of life is evident by its concept as auxiliary language. To use an auxiliary language a speaker must be open to other cultures and other ways of speaking, especially in our language Atlán. The phonetic of Atlán is such that its speakers need to broaden their understanding of a specific sound more than their mother tongue would most likely do. Atlán speakers need to be very open and conscientious because words have many ways of being expressed and a concept might be expressed in many different ways.

Consequently, even though Atlán tries to separate language from the sociolinguistic context, it is very questionable that this can be achieved. As the linguist Alvino Fantini stated, language and speakers' values, beliefs and attitudes are mutually interdependent. (Fantini 2020) The symbols that make up a language can only be understood in a sociolinguistic context and are interrelated with the worldview and norms, values and beliefs of the speaker. (Fantini 2020:270)

It is inevitable that Atlán will create its own sociolinguistic context. We might then question whether Atlán is an improvement to a natural language like English as lingua franca. After all we had rejected English because it relates to one particular culture and not with all cultures and now it seems that Atlán will only create a new singular milieu: "Two things seem

R 7. PRAGMATICS – JONATHAN ROOSE

simultaneously: people attempt to fit their language to the situation or context that their language, in turn, helped to create in the first place (Gee qtd. in Kecskes 2008:146).”

As a problem for Atlan however, it we hope to avoid this problem by making Atlan a very open and loose language. For example, with the phonemes being very flexible and not having one correct way to say a word. We hope that this will create a culture around Atlan that is similar to a bilingual or multilingual interaction. What is called by Fantini ‘incipient’ multilingualism:

To simply put, this stresses an attitude of willingness to engage with others with no common tongue (in an uncommon situation) and attempting to communicate. In this view, bilingualism begins with this attitude, with a willingness to engage, even when no skill exists. (Fantini 273)

We would hopefully create a loose social linguistic milieu that makes it more likely for people from different ethnicities and cultures to try and understand each other. More so, the language would be a natural language like English. In English there is less space for different ways to say things. However, this openness might also be problematic because of the “inescapable Babel effect” that we discussed in the introduction of this chapter. More language variation will

7.3. IS LANGUAGE GRAMMAR?

te more confusion. Atlan does not avoid the Babel effect, the contrary, it amplifies it.

Is language grammar?

ar, we have looked at pragmatics as a source of problems Atlan. Whether Atlan can express all the functions of language, how pragmatic use confuses Atlan into incomprehensibility, whether or it is not it is impossible to separate language from culture and, lastly, whether language is only meaningful in a sociolinguistic milieu. The pragmatic use of language has been an obstacle to be overcome. The focus was to make a grammatical structure without ambiguities, easy to understand and simple in structure, yet when faced with the task of speaking Atlan it quickly becomes unwieldy and confusing.

What my colleagues and I hoped set out to achieve was to create a perfect language first and then see whether people would use it. We started with the grammar and machine algorithms and from there moved on to use. Leaving pragmatics until the end of the book. It is reasonable to wonder whether this is the best way to understand language. Does language exist without speakers using it? Do the rules of language form the basis for speech or does speech form the basis for the rules of language? The linguist Roy Harris attacks the notion of

R 7. PRAGMATICS – JONATHAN ROOSE

as the basis of language. For him, to understand you must place speech/ use first.first. (Haris 1987) Knapp and Benn Michaels Walter take this even further, arguing that there is only intention and words, there is no meaning without speakers. (Knapp and Walter 1985) In other words, a room full of chimpanzees typing at random on typewriters would never create a work of Shakespeare. It could not even be a piece of paper with letters on it that look like a work of Shakespeare. A true work of Shakespeare needs the intention of an author behind it to be language. “Neo-pragmatic” thinkers show that it is not obvious how to get to a perfect language (or any language at all for that matter) by creating an abstract grammatical structure.

Visual artists and poets are not so sure of our view on language. The French poet Mallarmé for example rebelled against the notion that things mean what the dictionary says. He put a lot of emphasis on the emotion that the words in a poem invoke. (The French word Jour for example, or Mallarmé to sombre to express “day” and Nuit to express “night”) The futurist project of Zaum is another example; this ‘language’ has no grammar or syntax rules and consists of word combinations. It was created by the Russian avant-Garde poet Vladimir Kruchenykh and Velimir Khlebnikov to show that language doesn’t depend on grammatical rules (Tynyanov

7.3. IS LANGUAGE GRAMMAR?

That language has an element of what I can only describe as the feel of the language is nicely exemplified in nonsense poetry. Utterance can create significance and meaning even when they should not like in Lewis Carroll's *Jabberwocky* (Carroll qtd. in Hofstadter 366):

Twas bryllyg, and the slythy toves Did gyre and
gymble in the wabe: All mimsy were the borogoves;
And the mome raths outgrabe.

Lewis Carroll has created a poem mostly made of non-existent words, yet everybody believes that he or she can understand it. 'It feels right.' This is also shown in the numerous translations made of the poem that reproduce the nonsense words but then in, among others, a French (Frank L. Warriner in ibid 366) and German (Robert Scott qtd. In ibid 366) language:

Il brilgue: les têtes lubricilleux Se gyrent en vrill-
lant dans la guave. Enmînés sont les gougebosqueux
Et le mômerade horsgrave.

Es brillig war. Die schlichten Toven Wirrten un-
wimmelten in Waben Uns aller-mümsige Burggoven
Die mohmen Râth' ausgraben

Why then would we be able to translate this poem and linguistic playfulness into English? We cannot. English is too phonetically

R 7. PRAGMATICS – JONATHAN ROOSE

It does not have one specific correct sound and its meaning is not ordered around associations or ‘family resemblance’ as in English or French. Pragmatic linguists have a different approach to language than we had when making Atlan. The understanding of language is shown in linguistic phenomena that Atlan cannot replicate.

Conclusion

In this chapter, we discussed the pragmatic use of language and what this would entail for Atlan. It discussed how it makes explicit what in natural languages is only implicit. Atlan can only go that far in incorporating the input of speakers into the grammar. Furthermore, the chapter ended with the worry that Atlan’s loose structure will fragment and confuse the semantic unity of the conlang and that Atlan avoids having a dominant culture as much as possible, it does not manage to create a completely cultural language. Besides Atlan also easily becomes conceptually splintered because of the high amount of variation in the language. Lastly the chapter engaged with the question of whether focusing on perfecting the grammar would create a better language. Whether the way we created an auxlang was the best way. In conclusion then we can say that, while Atlan is an interesting and potentially rich linguistic experience,

7.4. CONCLUSION

ce, the use of the language will pose some problems for its speakers. For one, the language is very phonetically loose, and the effort needs to be exerted when listening to other speakers. Secondly implicit messages need to be made explicitly. This creates a peculiar way of thinking that might be frustrating. Lastly, because of its focus on a perfect grammar, errors or mispronunciations would quickly cause confusion between speakers. Thus, it is highly doubtful whether Atlan will be able to reach its goals when pragmatically used. This does not mean that Atlan is a failure, far from it. The set of words we have created is interesting as a poetic project. Being the average human sound of one particular semantic meaning, a word in Atlan shows what one particular sound means for the collective average human being. Atlan is a project that shows how meaning can be created and shared from very basic human experiences. Communication might be blocked by language barriers; language will carry meaning for everybody and everybody meaningfully listens. Although much is still to be discovered, and much more needs to be thought about to make Atlan a complete language, Atlan has created a system to make simple and meaningful words from fundamental human life-experience. We now hope that more people will get excited about this project and help co-create Atlan on a daily basis. What Atlan misses most is speakers, people who can live in Atlan and express themselves in Atlan. Hopefully we can find many of

R 7. PRAGMATICS – JONATHAN ROOSE

reators.

Chapter 8

Further suggestions

Sign Language – Stijn Janssens

According to the World Health Organisation, anno 2015, some 8% of the world's population is deaf. Different countries have their own sign languages, but these are often mutually unintelligible. Currently, no standard universal sign language exists. In our project, we did not focus our attention on creating a sign language to accompany Atlan as an IAL, but we might suggest how others, who have more knowledgeable on the topic than us, might use Atlan to construct such a sign language.

A database of signs from different sign languages might be used, such as 'Spreadthesign' by the European Sign Lan-

R 8. FURTHER SUGGESTIONS

entre. Software such as Sign Language Processing might be used to build data models to formalize and signs from different languages. This way, 'universal' might be generated by identifying overlap or similarity between language, weighing each language by related amount of speakers. These could then be mapped on's lexicon, and from there the whole language might be copied into this sign language.

g such a signing system might have the benefit of deaf people from around the world to communicate with another. It might also make sign language more accessible to hearing people, who would only have to learn some words provided they already speak English. This would foster communication and mutual understanding between deaf and hearing people, as well as serving as an extra linguistic tool for communication when two speakers can see each other but are unable to hear what the other is saying due to noisy circumstances.

Language Variation – Niek Elsinga

yourself in the following situation: You are standing in front of a machine which will take you back in time to the year 1223, 800 years in the past. Peace had just returned to England after a hard-fought civil war which resulted

8.2. LANGUAGE VARIATION – NIEK ELSINGA

the signing of the Magna Carta, which limited the power of English kings. England in this time was a country that was still full of meadows, forests, and pristine nature, with a relatively small population of an estimated 4 million people, nearly 60 million fewer than today. You would be able to walk around, and enjoy a moment of tranquillity, peace, and sit along crudely constructed cobbled walls which indicated paths that led to small villages, towns, and cities. While these towns might have been humble and quaint, they still bustled with life. People buzzing in tightly cramped avenues, with the smells of fresh crisp sourdough bread, savoury stews brewing over campfires, and the pungent aromas of leather tanners, the fires of the bellows of blacksmiths must have all coalesced in the cacophony of the community. Shops, pubs, and small boutiques which sell clothes and other food stuffs are able to be found in dimly-lit alleyways. Market stalls with a variety of kinds of fruits, vegetables, loaves of bread, meat, and perhaps even a mystic stall with unique herbs and spices from faraway land are able to be found on plaza on a Saturday morning.

You walk up to a stall which sells different types of stew. While you are not entirely certain what it is in it, you are drawn to a certain type of stew which is simmering above a fire. Its aromas and smells are unlike you have seen thus far, and thus, you go ahead and order a portion of this seemingly tasteful

R 8. FURTHER SUGGESTIONS

on. “I would like to order a portion of this stew, please,”
d say. The salesman looks you in the eye, astounded
aps, suspiciously. He replies: “Hwæne canst þú ġecwides?”
dumbfounded at the vendor. With every single word
try to pronounce, it seems that his gaze turns more
eventually, you just point. “That one, please,” whilst
to a stew you did not even examine. You give him the
which he luckily accepts, and hands you a bowl full of
substance. This one smells significantly less refined
other one, but you cannot be bothered to go back and
r dissatisfaction. It was your fault either way, since
ed erroneously. You sigh, and begrudgingly eat your
ch still turns out to be somewhat alright.

happened here? How come that you were not able
tand each other? In this case, there are two factors
First and foremost, language changes and branches
time. This is normal, natural, and occurs organi-
cause of language change, the Vulgar Latin of the Ro-
ire diverged and evolved into the modern Romance
s of Spanish, Portuguese, French, Italian, over the
the last two millennia (Sala & Posner, 1999). The
pened with Vedic Sanskrit, which is the now-extinct
from which a plethora of languages on the Indian
ent are derived from (Burde, 2004).

second factor is a variable that has happened in the

8.2. LANGUAGE VARIATION – NIEK ELSINGA

ish language specifically, which is a shift in the pronunciation of English vowels. The standardization of the English language did not occur between the 15th and the 16th centuries (Denham & Lobeck, 2009), while the pronunciation of English vowels shifted during this time. This shifting-event occurred between the 15th and 18th centuries, and influenced the pronunciation of vowels of every single English dialect (Labov, 1995). Where the vowels in the word “boot” are currently pronounced akin to the Dutch diphthong /oe/ in “koe”, or just standard English [oo], in the 13th century it would have sounded more like the Dutch /oʊ/ as in “groot”, or the ⟨aw⟩¹ in the modern British English word “yawn”. This Great Vowel Shift (GVS), as it is called, resulted in a different pronunciation of vowels compared to the graphemic notation for the entirety of the English language (Denham & Lobeck, 2009).

The GVS likely occurred because of multiple reasons, however, there is no academic consensus for one single solution (Crystal & Silverman, 2012). Some theories include migrations towards the southeast of England from neighbouring regions following the population decline caused by the Black Death (Crystal, 2018). Another theory is the influx of French

¹These brackets are used for linguistic notations. ⟨...⟩ is used for graphemic notation (i.e., the letters as they are written down); [...] is used for the actual realized phoneme (i.e., the sound that is actually heard); and /.../ is used for the intended phoneme.

R 8. FURTHER SUGGESTIONS

s with differing pronunciation compared to the Anglo-
pronunciation of Old and Early Middle English (Mill-
ayes, 2011). Another theory is the complete opposite,
states that due to the wars with France in which Eng-
entangled at that period in time, anti-French senti-
sed a shift in pronunciation to make English phonemes
s French (Nevalainen & Traugott, 2012). It is more
t the GVS occurred due a combination of these fac-
er than that a single one resulted in the entirety of
es (Silverman & Silverman, 2012).

theless, it occurred, and English has not been the
ce. It is not unlikely that events like the GVS will
gain since language is fluid per definition. Schol-
on that language variation and change is both in-
unpreventable, and continuously happening (Lyons,
this chapter, I will elaborate on the specifics of lan-
ange, how it can occur, and how we have designed
age to be resistant to language variation and change
in degree.

language variation and change: inevitable?

variation refers to the different ways in which a lan-
vary based on factors such as geography, social groups,
periods, and individual speakers. These variations

8.2. LANGUAGE VARIATION – NIEK ELSINGA

manifest in various forms, including pronunciation, vocabulary, grammar, and usage (O’Grady et al., 2001). Take regional dialects, for example. Different regions within a country or even different countries that share the same language have distinct dialects. For instance, in Dutch, there are variations between the Dutch from the Netherlands and the Dutch from Flanders. These dialects can diverge in pronunciation (e.g., the pronunciation of the letter “g” and “r” in the Netherlands and Flanders (Verhoeven, 2005)), vocabulary (e.g., use of the second-person pronoun “uw” in Flemish contrasted with “jouw” in Dutch (Vandekerckhove, 2005)), and grammar (e.g., “moeten aan doen” in Flemish compared to “moeten doen” in Dutch (Haeseryn, 1990)). Another example is sociolects, which are variations based on social factors such as social class, education level, or occupation. There may be differences in vocabulary and speech patterns between a group of doctors and a group of construction workers, reflecting their professional backgrounds and the jargon they use in their respective fields (Bybee, 2015; O’Grady et al., 2001).

The main point in language variation is that variation is not the same as language change, however, language variation can serve as a precursor to language change (Chambers et al., 2004). When a language exhibits variation among its speakers or regions, it provides the foundation for changes to occur and spread throughout a language community. Lan-

R 8. FURTHER SUGGESTIONS

Change, in continuation of language variation, refers to the process by which a language undergoes modifications over time. There are multiple factors about language change, and they can occur at every linguistic level: Phonology and morphology, syntax, semantics, and pragmatics (Meecham and Miller, 2001). Phonological change involves alterations in the sounds of a language. Over time, sounds can shift in pronunciation, merge with other sounds, or split into distinct sounds. This happens more frequently if multiple sounds exist that sound similar, such as the /θ/ in ⟨thing⟩ being replaced by the /f/. This happened to me personally, and occasionally I still make the error of pronouncing the ⟨th⟩ as an /f/ instead of the /θ/. Lexical change refers to changes in vocabulary. New words are constantly introduced into a language, while others become obsolete or change in meaning. For example, the word “awful” originally meant “full of awe”, but it has shifted to its current meaning of “bad” or “terrible” (Awful, Adj. and Adv.: Oxford English Dictionary). Languages can also undergo changes in their grammatical structures. This includes modifications in verb conjugation, word order, and the use of grammatical markers. Take for example the distinction with the indirect object “aan” in English “moeten aan doen” compared to the Dutch “aan doen”, as stated earlier (Haeseryn, 1990). Semantic change occurs when the meaning of words or phrases evolves

8.2. LANGUAGE VARIATION – NIEK ELSINGA

time. Words can acquire new meanings, lose old meanings, or sustain shifts in connotation. An example is the word “gay”, which originally meant “happy” but has taken on the additional meaning of “homosexual” in modern usage (Hiskey, 1997).

2 Mechanics of variation and change

While changes in language as part of coincidences of linguistic levels, change can also be instigated by social factors such as group identity and language contact. Social factors play a crucial role in shaping language variation and driving language change. Certain speech styles or dialects may be associated with social prestige, power, or higher social status. Speakers who want to align themselves with certain social classes may adopt features associated with these groups. For example the use of certain lexical items, jargon, or slang on a semantic level. Using words associated with the desired group can give the illusion of being associated with that group. As a result, language change can occur as features from prestigious or standard varieties are adopted and incorporated into the speech of a wider population (Labov, 1972). Besides class and income, speakers may also associate themselves with certain social groups, such as skaters, punks, ravers, etc. Language is an important marker of social identity.

R 8. FURTHER SUGGESTIONS

Speakers may consciously or unconsciously modify their language use to identify with or differentiate themselves from other social groups. Language change can occur as speakers adopt certain linguistic features of the identity of the target group as a way to signal membership in a specific community or subculture. This happens oftentimes in groups of young adults, and as older individuals might not understand them (Coupland, 1986). Language change is also often observed between different generations. Younger speakers may introduce new linguistic innovations or modifications in their language use compared to older generations. Over time, as younger generations become the majority, their linguistic features may spread and become more widespread, leading to language change (Kerfoot, 2016). However, within these older populations, language change can occur as well, through social networks. People and elderly individuals create a certain lect during their lifetimes. Because speakers interact with others in their social networks, language change can be achieved through the adoption and diffusion of these linguistic innovations. Language change can occur when innovative linguistic features spread through social networks, especially if influential individuals or groups adopt and promote these features (Ke et al., 2018). More sinister causes of language changes can also occur. If a particular variation is stigmatized or associated with negative stereotypes, speakers may avoid using those features

8.2. LANGUAGE VARIATION – NIEK ELSINGA

modify their language use to conform to more prestigious socially acceptable forms (Maass, 1999). The opposite can occur, in that positive attitudes towards certain features promote their adoption and spread, leading to language change. This strikes back at the aforementioned options.

3 Implications for Atlan in language development

There are many reasons for both language variation and language change. Change and variation in language are inevitable (Chen, 1994). How does this fare against constructed languages then? Very few constructed languages have seen widespread implementations, or mass numbers of speakers. It seems that there is limited evidence for linguistic variation in Esperanto, the major constructed language (Sherwood, 1982). However, Sherwood (1982) solely found variation in the pronunciation of phonemes, and there was still no mutual unintelligibility whatsoever. This is also likely due to the fact that Esperanto has seen no official adoption globally, and its use is mostly by aficionados (Piron, 1989). This causes the spoken language to be more or less the same as when it was invented, approximately 150 years ago.

Treading the waters of future language variation can be a difficult subject, due to the fact that the future, simply put, cannot be predicted. Language variation and change is, of

R 8. FURTHER SUGGESTIONS

inevitable. However, we have taken steps in order to make the language more resistant to language change. This is mostly in the phonology: because there are cardinal groups of vowels and consonants in which similar phonemes are allophonic and grouped, variation will less likely occur at the phonemic level. The same is the case for morphology: because prepositions, referents, demonstratives, etc. have a fixed set and meaning, and syntactical variation is allowed to a certain degree. Furthermore, because the lexicon is naturally generated, but random by definition for other languages, variation is more likely to occur due to the implementation of lexical items of the mother tongue of a speaker. This is the case for L1-to-L2-transfer (Sparks et al., 2009), however, is not the case for Atlán. Because some lexical elements and words have complex meanings cannot be accurately translated due to cultural differences (House, 2010), speakers are encouraged to use the word literally, and perhaps elaborate on it to unknown speakers. A good example of a word that has no direct translation in English is the German word ‘Schadenfreude’. In Atlán, this word could be described as “joy (SUS) + other (OF α) + affect (SIN \curvearrowright) + bad (PAK $\dot{\text{---}}$) = SIN.PAK” $\curvearrowright \alpha \curvearrowright \dot{\text{---}}$. The use of these lemmas indicates that a negative occurrence caused another person, in the person speaking it, a certain degree of joy. By

8.2. LANGUAGE VARIATION – NIEK ELSINGA

cribing the source word in Atlan, it can be understood by a wider array of speakers who are not familiar with the term. Variation in this case then is more or less irrelevant unless the words themselves change meaning. However, because the terms are procedurally generated, variation can only occur if the pronunciation of a consonant or vowel is changed. And of course, is less likely due to the grouping of the consonants and vowels in their allophonic categories. Due to these considerations, we think that Atlan as a whole will likely experience a delayed progression of variation and change.

4 Conclusion

Anything is clear, it would be that language variation and change is inevitable, unpredictable in its course, and constantly occurring. Atlan, like every other language, will meet the same fate and changes will occur, be it regionally, socio-economically, or culture-related. Perhaps in the future, multiple different variations of Atlan will coexist, intelligible or unintelligible. Then, the decisions made for the mitigation of language variation and change will be in vain. However, is that not exciting? When language variation occurs, this means that it is dynamic and fluid. Being able to see a language flourish is, perhaps, a better outcome than rigid measures intended to keep a language intelligible for everyone.

pter 9

sample texts







The Story of Babel – Jonathan Roose

Figure 1: A sequence of six diagrams illustrating the construction of a path γ from a point x to a point y in a domain Ω . The diagrams show the path γ (solid line) and the boundary $\partial\Omega$ (dashed line). The path starts at x and ends at y , passing through a series of points x_1, x_2, \dots, x_n and y_1, y_2, \dots, y_n . The path is constructed by connecting these points in a sequence of line segments, with the final segment connecting x_n to y_n .

AF AT.TEM PA.TU.TA

ON.PLURAL ON ALL.PL-EARTH PAST.BE

y on Earth had

LAN.SUM AN EK.SET.ON.ME.

SAY.COMU. AND. ACC.WORD.PLURAL.SAME

9.1. THE STORY OF BABEL – JONATHAN ROOSE

same language and the same words.

it came to pass,

-CLAUSE.AND AT.TIME.THIS.PAST.HAPPEN

U.ET.JA.ES.PA.NES

𐤅 𐤋 𐤅 𐤁 𐤁𐤏 𐤁𐤏

𐤁𐤏 𐤁 𐤅 𐤅

PO.EJ.AJ.ON

-CLAUSE.PROGR.TIME THR.PERS.PLURAL

they

𐤁 𐤏 𐤏 𐤏 𐤏 𐤏

U.MEF.LU.KEM LI.JOL

3.MOVE.PLACE.HOME DIRECTION.ORIGIN

erated from

𐤏 𐤏 𐤅 𐤅

PES.SON EJ.AJ.ON

CE.BIRTH.SON PERSON.3RD-REMOVED.PLURAL

R 9. EXAMPLE TEXTS

hey

○ ㄣ ㄹ ㄱ

K.PAS.MI.MAN

D.ECOUNTER ACC.PLAIN.BETWEEN. MOUNTAIN.

n a valley

ㅇ ㄹ ㄣ ㄹ ㄱ ㄴ

LU.NA.[S.JI.NAL].

RY PLACE.NAME. [SHINAR]

d of Shinar.

/ ㄷ ㄹ ㄱ ㄴ ㄹ

N PA.TU.MEN.SUP.KEN

. PLURAL PAST. VERB. MAKE. SIT. HOME

settled

ㄱ ㄴ ㄹ ㄱ ㄷ ㄹ

AJ.ON PA.LAN

9.1. THE STORY OF BABEL – JONATHAN ROOSE

ONST.2ND.PLACE 3TH.PLURAL PAST.SAY

e. They said

ᵛᵛᵛ ᵛ ᵛ

U.OF.ON,

C.SELF.OTHER.PLURAL

ne another,

ᵛ ᵛ ᵛ ᵛ ᵛ

U.MEN.AM.ON

ERATIVE.VERB.MAKE.1ST.REMOVED.PLURAL

ne, let us make

ᵛ ᵛ ᵛ ᵛ ᵛ ᵛ

FUK.JET.ON I.AN

BUILD-BLOCK.STONE.PLU REL.CAUSE.AND

ks ,and

ᵛ ᵛ ᵛ ᵛ ᵛ ᵛ

R 9. EXAMPLE TEXTS

.PEN.PIN.TOJ.

N . CAUSE . BECOME . SOLID

n hard.”

✓ ∅ ∇ ∅ ∇

ON TU.PA.SI

OCK . STONE . PLUR VERB . PAST . PREDICATE

rved

✓ ∅ ∇ ∅ ∇

ME FUK.MAJ.JET

PERSON . PLUR AS BUILDING . MATTER . STONE .

tone

✓ ∇ ∅ ∇ ∅ ∇

NA[BI.TU.MUM] TU.PA.SI

SE . AND NAMED[BITUMUM] TU . PAST . PREDICATE

num served

✓ ∅ ∇ ∅ ∇

9.1. THE STORY OF BABEL – JONATHAN ROOSE

AJ.ON ME FUK.MAJ.MOP.

.3TH-PERSON.PLUR AS BUILDING.MATTER.FOAM

n as mortar.

𐤀 𐤁 𐤂 𐤃 𐤄 𐤅

EJ.AJ.ON PA.LAN

-CLAUSE.AND PERSONS.3TH-REMO.PLURAL PAST.SAY

they said

𐤆 𐤇 𐤈 𐤉 𐤊 𐤋

J.MEN.NAP.AM.ON

ERATIVE.VERB.MAKE. BUILDING.1ST.REMOVED.PLURAL

ne, let us build

𐤌 𐤍 𐤎 𐤏 𐤐 𐤑 𐤒 𐤓

PTOS ,AN EK.NAP.LAK.TE

ONE.TOWN AND ACC.BUILDING.LONG.VERTICAL

y ,and a tower

EXERCISE 9. EXAMPLE TEXTS

𐌲𐌹𐌶𐌰 / 𐌲𐌹𐌶𐌰

FON.IN.SOM

S.3TH-REMO.PEAK IN SKY

Peak in the sky

𐌲𐌹𐌶𐌰 𐌲𐌹𐌶𐌰

EK.NA LO.AM.ON

E ACC.NA DAT.1ST-REMO.PLURAL

A name for ourselves

𐌲𐌹𐌶𐌰 𐌲𐌹𐌶𐌰

M.ON

SE.OTHER 1ST-REMO.PLURAL

𐌲𐌹𐌶𐌰 𐌲𐌹𐌶𐌰

MEJ.NE.IP

PROGRESSIVE.PREDICATE VERB.CHANGE.NEGATION.ONE

Scattered

9.1. THE STORY OF BABEL – JONATHAN ROOSE

𐤒 𐤓 𐤌 𐤕 𐤁

N.FA LOK.TEM.

R. AND. FAR ROUND. PLAN-EARTH

ver The earth.

𐤕 𐤁 𐤓𐤕 𐤏 𐤕

L TU.PA.KOM.LIT

ESS. GOD VERB. PAST. COME. DOWN

LORD came down

𐤁 𐤕 𐤓𐤕 𐤌

.TU.SIK EK.TOS

ERE. VERB. SEE ACC. CITY

ok at the city

𐤓𐤕 𐤏 𐤕 𐤓𐤕

EK.NAP.LAK.TE ES

ACC. BUILDING. LONG. VERTICAL DEMONSTRATIVE

tower that

R 9. EXAMPLE TEXTS

✓ ◁◦ ⁄ ✎

ON PO.KEN.NAP

Y.PERSON.PLURAL PERFECTED.MAKE.BUILDING

built

∩ √ ▷◦ ∟

L PA.LAN

SE.AND STRESS.GOD PAST.SAY

ORD said

△ d ∅ ↗

T.IP.SUM

SE.CONDITION POSSIBLE.ALL.ONE.COMMUNITY

e people

↗ ∅ b ∅ d ∅ √

N.SUM TA.AJ.AT.EJ

NE.SAY.COMMUNITY GENA.3TH-REMO.ALL.PERSON

language for all,

9.1. THE STORY OF BABEL – JONATHAN ROOSE

ᵐᵒ ᵐᵒ ᵐᵒ ᵐᵒ ᵐᵒ

ES.SI TU.UT AJ.ON

RELA.DEMONSTR.PRED. VERB.INSTRU. 3TH-REMO.PLURAL

is how they

ᵐᵒ ᵐᵒ -- ᵐᵒ --

A TU.KUK IS.KU

PECTED.BEGINNING. VERB.CUSTOM CONC.CONTR

begun to act then

ᵐᵒ ᵐᵒ ᵐᵒ

L ES AJ.ON

OAC.OR DEMONSTRA. 3TH-REMO.PLURAL

ing that They

ᵐᵒ ᵐᵒ ᵐᵒ ᵐᵒ ᵐᵒ ᵐᵒ

LAN.PI.NIL FI.TU FI.SI

B.SAY.POSSIBL.IMAGI FUTR.VERB FUTR.PREDICATE

R 9. EXAMPLE TEXTS

ose to do will be

o o \ o b 2

ON MU.TO.JAM.

GENI.3TH-REMO.PLURAL META.RANGE.HAND

air reach.

\cap \subset \circ \times \in d

S, O-UF.MEF.LI.OT

.PLURAL CONCLU EXCLA.INTENT.GO.TOWAR.DOWN

en, go down,

○ ↻ /- ↶

LIS.NEF.SIN

U.AND VERB.DISSOCIATION.KNOW.THINK

ound

ON EK.ME.SET.ON

3TH.PLURAL ACC.SAME.WORD.PLURAL

9.1. THE STORY OF BABEL – JONATHAN ROOSE

r speech

𐤁 𐤀 𐤁 𐤁

J.LI UF AJ.ON DEMONSTR.3TH.PLACE INTENT. 3TH.PLUR

e so they

𐤁 𐤁 𐤁 𐤁 𐤁

NE.NEF.SIM SU.EK.OF.ON

URE.NEG.KNOW.THINK ACC.SELF.OTHER.PLURAL

l not understand One another's

𐤁 𐤁 𐤁 𐤁

SAME.WORD.PLURAL

ME.SET.ON

ch.

𐤁 𐤁 𐤁

O.JEL

CLAUSE.CONCLUSION. EXCLA.GOD

9.1. THE STORY OF BABEL – JONATHAN ROOSE

they stopped

𐤀 𐤍 𐤋 𐤁

KEN.NAP TOS.

3. MAKE.BUILD TOWN

ding the city.

𐤓 𐤋 𐤁𐤍𐤏𐤔

TO.SI EK.A.TOS

ONSTR.REASEN.PREDICATE ACC.STRESS.CITY

t is why it

called Babel

PA.NA[PA.PEL]

3. PAST . NAME [BABEL]

𐤓𐤏 𐤁𐤍𐤏𐤔

𐤓𐤏 𐤁𐤍𐤏𐤔 𐤓𐤏 𐤁𐤍𐤏𐤔

FU ES.AJ.LU A.JEL

R 9. EXAMPLE TEXTS

IN-STATE DEMOSTR.3TH.PLACE STRESS.GOD

here the LORD

𐎗 𐎒 𐎗

S.NEF.SIN VERB.PAST.DISSOCIATION.KNOW.THINK

ed

𐎗 𐎗𐎗 𐎗𐎗 𐎗𐎗 𐎗𐎗 𐎗𐎗

N.ME AT.TEM.SUM.

.WORD.PLURAL ALL.PLA-EARTH.COM.

h of the whole earth.

𐎗 𐎗 𐎗 𐎗 𐎗

U.AND DERECTION.ORIGIO DEMONSTR.3TH.PLACE

MI ES.AJ.LU

there

𐎗 𐎗𐎗 𐎗𐎗 𐎗𐎗 𐎗𐎗

U.PA.SI.NE.IP

OD VERB.PAST.PREDICATE.NEGATION.ONE

9.1. THE STORY OF BABEL – JONATHAN ROOSE

LORD scattered

𐤋𐤍𐤏𐤕𐤍𐤏𐤕𐤍𐤏𐤕𐤍𐤏𐤕

AJ.ON KI.AN.FA

3TH.PLUR NEAR.AND.FAR

n over

𐤁𐤏𐤕𐤍𐤏𐤕𐤍𐤏𐤕𐤍𐤏𐤕

AT.SEM.TEM

ALL.SURFACE.PLAN-EARTH

face of the whole earth.

R 9. EXAMPLE TEXTS

The Epic of Gilgamesh – Niek Elsinga

Dream - 𐎧 𐎡𐏁 𐎢𐏁 𐎣𐏁 𐎤𐏁 𐎥𐏁

$\varphi \triangleleft \downarrow$ $\circ \triangleleft \hookrightarrow$ $\hookrightarrow \downarrow$
 $\downarrow \circ$ $\hookrightarrow -$ $\circ \downarrow$ $\downarrow \hookrightarrow$

].SI.PO.MOT PA.TU.PO.LAN.LO.JEL

T.ON.LUM, AN LAN LO.TEN.LUF

-predicate-process-death}{past-verb-process-speak-
instrument-this-word-plural-last}, {and} {speak} {to-
nce-love-of-him}

the last words the dying Enkidu did pray, and say to
his faithful companion:

ㄱ ㅋ ㆁ ㄷ ㅌ ㄴ
 ㄹ ㅎ ㅇ ㆆ ㅡ ㅜ ㅠ

UL.OM SOM.ON AN TEM TU.NI.NOJ

PAK.PIK PO.JA EJ.AM.SI.IP PA.TUF.LI.JUN

9.2. THE EPIC OF GILGAMESH – NIEK ELSINGA

...tion-dream-plural-hallucination}{sky-plural}{and}{earth}
...o-perfective shout}{accusative-sound-bad-big}{progressive
...}{person-first-predicate-one}{past stand direction descruc-
...}
...dreams last night the heavens and the earth poured out
...t groans while I alone stood facing devastation.

𒀭 | 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭
𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭
𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭
𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭

SI.TUT.AN.MEN.TEF TU.PA.FUL.ET.OT LI.EJ.AM I.AM
U.PA.PUS.EK.EJ.AM UT.NEK.TAK.ON.TA.AJ
NAP.MOL.TEF.TA.MOT IN.ES.NA[IL.KAL.LA],
ES.TA.NE.LAS AJ.TU.JON

...ature predicate strong and make fear}{verb past fly co-
...nate under}{direction 1ST}{rel.clause and}{3rd-verb past
...n acc. person 1st}{Instrument nail sharp plural genitive
...{direction}{house full fear genitive death}{in this [name]
...A.LA.}{feminine king genitive not-light}{3SG verb control}

R 9. EXAMPLE TEXTS

and threatening creature flew down at me and pushed
its talons toward the horror-filled house of death wherein
queen of shades, stands in command.

𐌹 𐌸 𐌺 𐌶 𐌵 𐌰 𐌸 𐌶𐌵 𐌰 𐌵 𐌶 𐌰 𐌶

NE.LAS ES MUL EK.NE.SO.EJ TU.SIK LAS.TA.JAN

G}{is}{negation light}{demonstrative}{allow}{acc. not

son}{relative clause verb see}{acc. light genitive day}

darkness which lets no person again see light of day.

𐌹 𐌺 𐌸 𐌶 𐌵 𐌰 𐌸 𐌶𐌵 𐌰 𐌵 𐌶 𐌰 𐌶

LOL NE.LI.FIN.LAS.AN.SUS PU LAS AN SUS FIN

G}{is}{road not-direction life light and joy}

a road leading away from bright and lively life.

𐌰 𐌵 𐌶 𐌶 𐌶𐌵 𐌰 𐌵 𐌰 𐌶 𐌰 𐌶 𐌰 𐌶 𐌰 𐌶
𐌰 𐌶 𐌶 𐌶 𐌶 𐌶 𐌶 𐌶 𐌶 𐌶 𐌶 𐌶 𐌶 𐌶

PUJ.FIN ES.EJ.AJ.ON TU.KOSTOJ EK.TUL.SUK AN

A NET.MEJ.TEP.JE LO.FAN.JIT.PAK.TA.AJ.ON

G}{verb stay-alive}{demonstrative person 3rd plural}{VERB

9.2. THE EPIC OF GILGAMESH – NIEK ELSINGA

sume solid}{acc. dust-dry} {and} {negation verb genitive}
 er change temperature negative} {dative want liquid bad
 tive 3rd-plural}

re dwell those who eat dry dust and have no cooling water
 heir awful thirst.

𒂗 𒍪 𒀭 𒃶 𒂗 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭
 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭
 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭
 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭 𒀭

A EJ.AM.PA.TU.TUF ET.LU.AJ I.EJ.AJ.PA.TU.SIK

AT.ES.EJ.AJ.ON TU.PA.MOT AN A.EK.TES.ON

AJ ES.EJ.AJ PA.TU.MEJ.NE.LAS

TU.TA EK.JEJ.FA.AN.POPTA.AJ.ON

rdinate time} {person 1st do past stand} {coordinate place
 {relative clause person first verb past see} {acc. all demon-
 ive person 3rd plural} {verb past death} {and} {stress acc.
 plural} {king plural} {in-between 3SG} {demonstrative 3SG
 on} {passive verb change not light}{acc. success far and
 ous genitive 3rd plural}

R 9. EXAMPLE TEXTS

and there I saw all those who've died and even kings
whose darkened souls have none of their remote and
glory.

𐤀 𐤁 𐤂 𐤃 𐤄 𐤅 𐤆 𐤇 𐤈 𐤉 𐤊 𐤋 𐤌 𐤍 𐤎 𐤏 𐤐 𐤑 𐤒 𐤓 𐤔 𐤕 𐤖 𐤗 𐤘 𐤙 𐤚 𐤛 𐤜 𐤝 𐤞 𐤟 𐤠 𐤡 𐤢 𐤣 𐤤 𐤥 𐤦 𐤧 𐤨 𐤩 𐤪 𐤫 𐤬 𐤭 𐤮 𐤯 𐤰 𐤱 𐤲 𐤳 𐤴 𐤵 𐤶 𐤷 𐤸 𐤹 𐤺 𐤻 𐤼 𐤽 𐤾 𐤿 𐥀 𐥁 𐥂 𐥃 𐥄 𐥅 𐥆 𐥇 𐥈 𐥉 𐥊 𐥋 𐥌 𐥍 𐥎 𐥏 𐥐 𐥑 𐥒 𐥓 𐥔 𐥕 𐥖 𐥗 𐥘 𐥙 𐥚 𐥛 𐥜 𐥝 𐥞 𐥟 𐥠 𐥡 𐥢 𐥣 𐥤 𐥥 𐥦 𐥧 𐥨 𐥩 𐥪 𐥫 𐥬 𐥭 𐥮 𐥯 𐥰 𐥱 𐥲 𐥳 𐥴 𐥵 𐥶 𐥷 𐥸 𐥹 𐥺 𐥻 𐥼 𐥽 𐥾 𐥿 𐦀 𐦁 𐦂 𐦃 𐦄 𐦅 𐦆 𐦇 𐦈 𐦉 𐦊 𐦋 𐦌 𐦍 𐦎 𐦏 𐦐 𐦑 𐦒 𐦓 𐦔 𐦕 𐦖 𐦗 𐦘 𐦙 𐦚 𐦛 𐦜 𐦝 𐦞 𐦟 𐦠 𐦡 𐦢 𐦣 𐦤 𐦥 𐦦 𐦧 𐦨 𐦩 𐦪 𐦫 𐦬 𐦭 𐦮 𐦯 𐦰 𐦱 𐦲 𐦳 𐦴 𐦵 𐦶 𐦷 𐦸 𐦹 𐦺 𐦻 𐦼 𐦽 𐦾 𐦿 𐧀 𐧁 𐧂 𐧃 𐧄 𐧅 𐧆 𐧇 𐧈 𐧉 𐧊 𐧋 𐧌 𐧍 𐧎 𐧏 𐧐 𐧑 𐧒 𐧓 𐧔 𐧕 𐧖 𐧗 𐧘 𐧙 𐧚 𐧛 𐧜 𐧝 𐧞 𐧟 𐧠 𐧡 𐧢 𐧣 𐧤 𐧥 𐧦 𐧧 𐧨 𐧩 𐧪 𐧫 𐧬 𐧭 𐧮 𐧯 𐧰 𐧱 𐧲 𐧳 𐧴 𐧵 𐧶 𐧷 𐧸 𐧹 𐧺 𐧻 𐧼 𐧽 𐧾 𐧿 𐨀 𐨁 𐨂 𐨃 𐨄 𐨅 𐨆 𐨇 𐨈 𐨉 𐨊 𐨋 𐨌 𐨍 𐨎 𐨏 𐨐 𐨑 𐨒 𐨓 𐨔 𐨕 𐨖 𐨗 𐨘 𐨙 𐨚 𐨛 𐨜 𐨝 𐨞 𐨟 𐨠 𐨡 𐨢 𐨣 𐨤 𐨥 𐨦 𐨧 𐨨 𐨩 𐨪 𐨫 𐨬 𐨭 𐨮 𐨯 𐨰 𐨱 𐨲 𐨳 𐨴 𐨵 𐨶 𐨷 𐨸 𐨹 𐨺 𐨻 𐨼 𐨽 𐨾 𐨿 𐩀 𐩁 𐩂 𐩃 𐩄 𐩅 𐩆 𐩇 𐩈 𐩉 𐩊 𐩋 𐩌 𐩍 𐩎 𐩏 𐩐 𐩑 𐩒 𐩓 𐩔 𐩕 𐩖 𐩗 𐩘 𐩙 𐩚 𐩛 𐩜 𐩝 𐩞 𐩟 𐩠 𐩡 𐩢 𐩣 𐩤 𐩥 𐩦 𐩧 𐩨 𐩩 𐩪 𐩫 𐩬 𐩭 𐩮 𐩯 𐩰 𐩱 𐩲 𐩳 𐩴 𐩵 𐩶 𐩷 𐩸 𐩹 𐩺 𐩻 𐩼 𐩽 𐩾 𐩿 𐪀 𐪁 𐪂 𐪃 𐪄 𐪅 𐪆 𐪇 𐪈 𐪉 𐪊 𐪋 𐪌 𐪍 𐪎 𐪏 𐪐 𐪑 𐪒 𐪓 𐪔 𐪕 𐪖 𐪗 𐪘 𐪙 𐪚 𐪛 𐪜 𐪝 𐪞 𐪟 𐪠 𐪡 𐪢 𐪣 𐪤 𐪥 𐪦 𐪧 𐪨 𐪩 𐪪 𐪫 𐪬 𐪭 𐪮 𐪯 𐪰 𐪱 𐪲 𐪳 𐪴 𐪵 𐪶 𐪷 𐪸 𐪹 𐪺 𐪻 𐪼 𐪽 𐪾 𐪿 𐫀 𐫁 𐫂 𐫃 𐫄 𐫅 𐫆 𐫇 𐫈 𐫉 𐫊 𐫋 𐫌 𐫍 𐫎 𐫏 𐫐 𐫑 𐫒 𐫓 𐫔 𐫕 𐫖 𐫗 𐫘 𐫙 𐫚 𐫛 𐫜 𐫝 𐫞 𐫟 𐫠 𐫡 𐫢 𐫣 𐫤 𐫥 𐫦 𐫧 𐫨 𐫩 𐫪 𐫫 𐫬 𐫭 𐫮 𐫯 𐫰 𐫱 𐫲 𐫳 𐫴 𐫵 𐫶 𐫷 𐫸 𐫹 𐫺 𐫻 𐫼 𐫽 𐫾 𐫿 𐬀 𐬁 𐬂 𐬃 𐬄 𐬅 𐬆 𐬇 𐬈 𐬉 𐬊 𐬋 𐬌 𐬍 𐬎 𐬏 𐬐 𐬑 𐬒 𐬓 𐬔 𐬕 𐬖 𐬗 𐬘 𐬙 𐬚 𐬛 𐬜 𐬝 𐬞 𐬟 𐬠 𐬡 𐬢 𐬣 𐬤 𐬥 𐬦 𐬧 𐬨 𐬩 𐬪 𐬫 𐬬 𐬭 𐬮 𐬯 𐬰 𐬱 𐬲 𐬳 𐬴 𐬵 𐬶 𐬷 𐬸 𐬹 𐬺 𐬻 𐬼 𐬽 𐬾 𐬿 𐭀 𐭁 𐭂 𐭃 𐭄 𐭅 𐭆 𐭇 𐭈 𐭉 𐭊 𐭋 𐭌 𐭍 𐭎 𐭏 𐭐 𐭑 𐭒 𐭓 𐭔 𐭕 𐭖 𐭗 𐭘 𐭙 𐭚 𐭛 𐭜 𐭝 𐭞 𐭟 𐭠 𐭡 𐭢 𐭣 𐭤 𐭥 𐭦 𐭧 𐭨 𐭩 𐭪 𐭫 𐭬 𐭭 𐭮 𐭯 𐭰 𐭱 𐭲 𐭳 𐭴 𐭵 𐭶 𐭷 𐭸 𐭹 𐭺 𐭻 𐭼 𐭽 𐭾 𐭿 𐮀 𐮁 𐮂 𐮃 𐮄 𐮅 𐮆 𐮇 𐮈 𐮉 𐮊 𐮋 𐮌 𐮍 𐮎 𐮏 𐮐 𐮑 𐮒 𐮓 𐮔 𐮕 𐮖 𐮗 𐮘 𐮙 𐮚 𐮛 𐮜 𐮝 𐮞 𐮟 𐮠 𐮡 𐮢 𐮣 𐮤 𐮥 𐮦 𐮧 𐮨 𐮩 𐮪 𐮫 𐮬 𐮭 𐮮 𐮯 𐮰 𐮱 𐮲 𐮳 𐮴 𐮵 𐮶 𐮷 𐮸 𐮹 𐮺 𐮻 𐮼 𐮽 𐮾 𐮿 𐯀 𐯁 𐯂 𐯃 𐯄 𐯅 𐯆 𐯇 𐯈 𐯉 𐯊 𐯋 𐯌 𐯍 𐯎 𐯏 𐯐 𐯑 𐯒 𐯓 𐯔 𐯕 𐯖 𐯗 𐯘 𐯙 𐯚 𐯛 𐯜 𐯝 𐯞 𐯟 𐯠 𐯡 𐯢 𐯣 𐯤 𐯥 𐯦 𐯧 𐯨 𐯩 𐯪 𐯫 𐯬 𐯭 𐯮 𐯯 𐯰 𐯱 𐯲 𐯳 𐯴 𐯵 𐯶 𐯷 𐯸 𐯹 𐯺 𐯻 𐯼 𐯽 𐯾 𐯿 𐰀 𐰁 𐰂 𐰃 𐰄 𐰅 𐰆 𐰇 𐰈 𐰉 𐰊 𐰋 𐰌 𐰍 𐰎 𐰏 𐰐 𐰑 𐰒 𐰓 𐰔 𐰕 𐰖 𐰗 𐰘 𐰙 𐰚 𐰛 𐰜 𐰝 𐰞 𐰟 𐰠 𐰡 𐰢 𐰣 𐰤 𐰥 𐰦 𐰧 𐰨 𐰩 𐰪 𐰫 𐰬 𐰭 𐰮 𐰯 𐰰 𐰱

PIK.TA.TEM TU.PA.SI.PI.EM.FUN I.AN EJ.AM

EF.LI.IN NAP.TA.MOT

big genitive earth} {verb past is passive inverse find}
 clause and} {past verb walk direction inside} {house
 leath}

My greatness was forfeit and I entered then into the death.

འཕྲིན་གྱི་དོན་རྒྱུ་ལྟར་སྤྱད་པའི་དབང་ལྷན་གྱི་མཉམ་སྦྲེལ་གྱི་
 དོན་ལྟར་སྤྱད་པའི་དབང་ལྷན་གྱི་མཉམ་སྦྲེལ་གྱི་

S.EJ.AJ PA.SI.LU.AJ JA.LAK PA.TU.MEF.LI.EF

MAL.POT EK.EJ.AM

ural} {demonstrative 3SG person} {past be place 3rd}

9.2. THE EPIC OF GILGAMESH – NIEK ELSINGA

e long} {past do come to above rel.clause intention verb

t good acc. person 1SG}

ers who've been there long did rise to welcome me.”

Chapter 10

Exercises – Max Geraedts

The purpose of these exercises is to make the reader familiar with the language. To get a feel for the language. They should not be a test of the knowledge of the reader but rather as a way to get your bearings in Atlán.

10.1 – Creating a basic possessive sentence & writing your own name

In the first exercise we will be translating the sentence “My name is Max”. This exercise will demonstrate how to combine words to create new words in a simple and familiar context. We will start by creating the Atlán word for “my”. To help you, I will give you the Atlán translation for “I”; EJ.AM. The word “I” is composed of EJ 𐌺 meaning person, and AM

; 1st removed: speaker. The Atlán word for “my” can easily be made from this by adding the possessive prefix TA ʔ . This leaves us with TA.EJ.AM ʔ ʃ ɔ̃ .

What is the Atlán translation of “name”? And “is”? (Note: “is” is a predicate and comes with a marker). Remember that the subject of an Atlán word always comes first in a sentence, so the word order would be: Name my is [Name].

What comes the easiest part of the sentence, your name. To write your name in Atlán all you must do is transliterate your name into Atlán’s set of 14 sounds and put a cartouche around your name.

) Write your name in Atlán:

Exercise 10.2 – Creating basic active sentences in the present simple tense


Now that we know how to create a basic possessive sentence, we are going to look at how to form basic active sentences. For example, “I am walking” and “He is writing”.

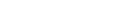
) What is the Atlán translation for “I walk”, “You walk” and “He walks”? (Note: present tense has no need for a marker).

What is the English translation of the following sentences?

R 10. EXERCISES – MAX GERAEDTS

AM TU.LIK EK.POK \circlearrowleft \searrow \circ \subset $\text{---}\circ\text{---}$ $\text{---}\cdot$

AJ TU.JIL EK.FIL 

AM.ON TU.KOS.TOJ 

10.3 – Creating basic sentences in the past simple

move to the past simple tense. Translate the following sentences to Atlán.

orked.

worked.

ey worked.

a played yesterday.

10.4 – Numbers

number system can be used as a ten-base system and a twelve-base system. To see the difference between these two systems you can look at how to spell your name in both ten-base and twelve-base. We will begin with exercises that use the ten-base system as this will probably be more familiar.

the following sentences using a ten-base number

) I have three fish.

) I have eleven fish.

) I have a thousand fish.

Translate the following sentences using a duodecimal (twelve-
) number system:

) I have three fish.

) I have eleven fish.

) I have a thousand fish.

) I have a 1.728 fish.

R 10. EXERCISES – MAX GERAEDTS

solutions

My name is ...” in Atlan would be; NA TA.EJ.AM SI

𐤎 𐤔 𐤓 𐤁 𐤓 NAME.

walk = AM.TU.TOM 𐤁 𐤕 𐤕

u walk = UN.TU.TOM 𐤏 𐤕 𐤕

walks = AJ.TU.TOM 𐤅 𐤕 𐤕

AM PA.TU.KAN 𐤓 𐤁 𐤓 𐤕 𐤏 𐤕

AM.ON PA.TU.KAN 𐤓 𐤁 𐤕 𐤓 𐤕 𐤏 𐤕

LAJ.ON PA.TU.KAN 𐤓 𐤕 𐤕 𐤓 𐤕 𐤏 𐤕

UN.ON PA.TU.LAJ ET.JAN.POP

𐤕 𐤕 𐤕 𐤓 𐤕 𐤕 𐤕 𐤕 𐤕 𐤕 𐤕 𐤕

umbers

AM TU.TA MIS.UP - 𐤓 𐤁 𐤕 𐤔 𐤕 𐤕

AM TU.TA MIS.JI.IP - 𐤓 𐤁 𐤕 𐤔 𐤕 𐤕 𐤕 𐤕

10.1. SOLUTIONS

) EJ.AM TU.TA MIS.NU - 𐎶 𐎠 𐎡 𐎢 𐎣 𐎤

) EJ.AM TU.TA MIS.UP - 𐎶 𐎠 𐎡 𐎢 𐎣 𐎥

) EJ.AM TU.TA MIS.JO - 𐎶 𐎠 𐎡 𐎢 𐎣 𐎦

) EJ.AM TU.TA. MIS.UK.JO.IK

𐎶 𐎠 𐎡 𐎢 𐎣 𐎦 𐎧 𐎩

) EJ.AM TU.TA MIS.NU 𐎶 𐎠 𐎡 𐎢 𐎣 𐎤

ogue

anks to,
ic, for her amazing guidance throughout this whole
a, we wouldn't know where would be without you.
y remember how badly we wanted you to be our su-
Luckily, your enthusiasm for this project was almost
an ours! You always had great ideas and was ready
We cannot thank you enough for all the proofreading
nd your many (!) comments. Your constructive crit-
d thoughtful suggestions brought the quality of our
whole new level. Your insights and expertise added
d credibility to our work, and we couldn't be more

ank you for putting up with our chaotic meetings, late
ons, and last-minute room changes. You never gave
even if there had to be three separate meetings to let
view one documentary. Your commitment and per-

grance made all the difference, and we are sincerely grateful for your support.

, we had a total blast these couple of weeks! Since you can't give this course anymore, we surely hope it was just as much fun for you! Thank you a million times over!

Angela Carpenter, for her support and contribution to this project. Angela, your help and assistance have truly made an impact on this product.

That afternoon when you were ready to lend an ear, despite the crazy time difference, was a true joy for us. How cool it was that we were talking with the professor we saw in a documentary only a few weeks before! Your patience and genuine interest in our non-existing language made our brainstorming session truly delightful. We owe you a debt of gratitude for sharing all that knowledge from your classes. Your enthusiasm, and encouragement meant the world to us!

Humanities Honours Programme, for funding this project, basically making a dream come true. Your help was like finding a hidden treasure in the bustling streets of Hong Kong. Thank you for providing the coffee to keep us going. Your contribution will forever be cherished and remembered.

ography

n, J. (1994). *Language Change: Progress or Decay?*
bridge University Press, pp. 627–637.

S.H. (1970). *Arabic Phonology: An acoustical and
ological investigation*. De Gruyter.

A. (2012). “The usefulness of metrics in the quan-
tion of speech rhythm”. In: *Journal of Phonetics*
pp. 351–373.

.A. and F.A. Navarro (2006). *Medicine: use of En-*
Elsevier EBooks, pp. 752–759.

F.M. (1963 [1759]). *Candide, ou lOptimisme*. Dieter
brandt.

.L. (1975 [1955]). *How to Do Things With Words:*
William James Lectures delivered at Harvard Univer-
1955. Oxford University Press.

i. and adv. (Mar. 2023).

BIBLIOGRAPHY

- er, M. (1998). "The Polysynthesis Parameter". In: *Language* 74.1, p. 129.
- er, R.S. and P.K. Benedict (1997). *Modern Cantonese Phonology*. Walter de Gruyter.
- ney, Michael (2021). "Analysis". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2021. Metaphysics Research Lab, Stanford University.
- lin, B. and P. Kay (1969). *Basic Color Terms: Their Universality and Evolution*. University of California Press.
- nal, B. and N. Altman (2010). "The connectivity of the superior longitudinal fasciculus: a tractography DTI study". In: *Magnetic resonance imaging* 28.2, pp. 217–225.
- swas-Teener, R. and J. Teeny (n.d.). *States of consciousness*. Retrieved: April 9, 2023. URL: %5Curl%7Bhttps://nobaproject.com/modules/states-of-consciousness%7D.
- ek, W.H. (1862). *A Comparative Grammar of South African Languages*. Trubner & Co.
- ins, A.D. (n.d.). *The Analytical Language of John Wilkins*. Retrieved: April 8, 2023. URL: %5Curl%7Bhttps://

RAPHY

ewdblevins.github.io/JohnWilkins/%C2%A0%

(2007). *The Grammar of Words: An Introduction to
istic Morphology*. OUP Oxford.

ky, L. and M. Liberman (2010). *For and Against
istic Relativity*. Retrieved: 4 March 2023. URL: %5Curl%

tps://web.archive.org/web/20120215074211/
:www.economist.com/debate/overview/190%

Jean-Louis (2019). “The Philosophy of Language
the Ontology of Knowledge”. In: p. 18.

R. and E. Lenneberg (1954). “A study in language
ognition”. In: *Journal of Abnormal and Social Psy-
gy* 49.3, pp. 454–462.

(2004). *Rituals, Mantras and Science: An Integral
ective*. Motilal Banarsidass Publishers.

A.R. (november 2018). “The special nature of
ralian phonologies: why auditory constraints on
an language sound systems are not universal”. In:
edings of Meetings on Acoustics 176ASA 35.1.

(2015). *Language Change*. Cambridge University

.

BIBLIOGRAPHY

- mpbell, L. (2010). "Language isolates and their history, or, whats weird, anyway?" In: *Annual Meeting of the Berkeley Linguistics Society* 36.1, pp. 16–31.
- (2018). "How many language families are there in the world?" In: *Anuario del Seminario de Filologia Vasca 'Julio de Urquijo'* 52.1/2, pp. 133–152.
- ron et al., (2017). *Speaking about sounds: a tool for communication on sound features*.
- mbers, J.K., P. Trudgill, and N. Schilling-Estes (2004). *The Handbook of Language Variation and Change*. Blackwell Publishing Ltd eBooks.
- msky, Noam (2002). *Syntactic Structures*. Mouton de Gruyter.
- isholm, H.E. (1911). *Cartouche*. 11th ed. Cambridge University Press.
- en, M. (n.d.). *Definitions and the tree of Porphyry*. Retrieved: April 8 2023. URL: %5Curl%7Bhttp://faculty.washington.edu/smcohen/433/PorphyryTree.html%5C2%A0%7D.
- rie, B. (1989). *Language Universals and Linguistic Typology: Syntax and Morphology*. University of Chicago Press.

RAPHY

- d, N. (1985). “Hark, hark, the lark’: Social motivations for phonological style-shifting”. In: *Language Communication* 5.3, pp. 153–171.
- .W. (1997). *The Complete Lojban Language*. Vol. 15. Logical Language Group.
- len, A. (1997). *Intonation*. Cambridge University Press.
- al (2018). *The Cambridge Encyclopedia of the English Language*. Cambridge University Press.
- ert, J.L. and D. Diderot (2020). *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers*. Open On Demand Ltd.
- , G. (1968 [1661]). *Ars signorum*. Scholar P.
- de, R.G. (1995). *The Development of Cognitive Anthropology*. Cambridge University Press.
- .A. (2016). *Implicature*. Vol. 6. Springer.
- ure, F.[Translated by W. Baskin] (1959). *Course in Formal Linguistics*. Philosophical Library.
- G. and F. Guattari (2019 [1980]). *A Thousand Plateaus: Capitalism and Schizophrenia*. Bloomsbury.
- , K. and A. Lobbeck (2012). *Linguistics for everyone: an introduction*. Cengage Learning.

BIBLIOGRAPHY

- tsch, L. (2019). *The Five Basic Tastes*. Retrieved: April 9, 2023. URL: %5Curl%7Bhttps://www.psychologytoday.com/us/blog/memory-catcher/201903/the-five-basic-tastes%C2%A0%7D.
- , U. Translated by J. Fentress (1994). *The Search for the Perfect Language*. Blackwell Publishers Inc.
- cation, First (2022). *EF EPI: EF English Proficiency Index: A Ranking of 111 Countries and Regions by English Skills*. EF Education First.
- yclopaedia Britannica (n.d.). *Common scale types*. Retrieved: April 9, 2023. URL: %5Curl%7Bhttps://www.britannica.com/art/scale-music/Common-scale-types%C2%A0%7D.
- , Gkonou (2015). *New Directions in Language Learning Psychology*. Springer, pp. 71–96.
- ns, N. and S.C. Levinson (2009). “The myth of language universals: Language diversity and its importance for cognitive science”. In: *Behavioral and brain sciences* 32.5, pp. 429–448.
- hini, A.E. (2020). “Language; An essential component of intercultural communicative competence”. In: *Routledge Handbook of Language and Intercultural Communication*. Routledge, pp. 267–281.

RAPHY

- Mafalda (Jan. 2017). "The Information Visual-
on Thematic, Head to Head with Color Blind-
. In.
- .T., M.D. Hauser, and N. Chomsky (2005). "The
tion of the language faculty: Clarifications and
ications". In: *Cognition* 97.2, pp. 179–210.
- IV and W. Benjamin (2011). *Indo-European Lan-
and Culture: An Introduction*. John Wiley & Sons.
- , M. (2010 [1966]). *The Order of Things: An Archae-
of the Human Sciences*. Routledge.
- , J. (1986). "Aristotle on Species Variation". In:
sophy 61.236, pp. 245–252.
- Jonardon (2021). "Analytic Philosophy in Early
ern India". In: *The Stanford Encyclopedia of Phi-
hy*. Ed. by Edward N. Zalta. Winter 2021. Meta-
cs Research Lab, Stanford University.
- et al., (2012). "The grammatical view of scalar
atures and the relationship between semantics
pragmatics". In: *Handbook of Semantics*. Ed. by
er et al. Mouton de Gruyter, pp. 2297–2332.
- L.P. (2019 [430]). *Plotinus: The Enneads*. Cambridge
ersity Press.

BIBLIOGRAPHY

- Chomsky, N. (1972). "Language and the brain". In: *Scientific American* 226.4, pp. 76–83.
- Cliff, C. and A.E. Wierzbicka (2002). *Meaning and Universal Grammar: Theory and Empirical Findings*. John Benjamins Publishing Company.
- Cliff, C. (2012). "Semantic primes, semantic molecules, semantic templates: Key concepts in the NSM approach to lexical typology". In: *Journal of Multilingual and Multicultural Development* 35.6, pp. 445–467.
- Googletrans (n.d.). *Googletrans 3.0.0. documentation*. Retrieved February 20th, 2023.
- Hakken, C. (2007). "The Contribution of linguistic factors to the intelligibility of closely related languages". In: *Journal of Multilingual and multicultural development* 30.6, pp. 445–467.
- Hawes-Brown, C. (2010). *Dancing for Hathor: Women in Ancient Egypt*. Continuum.
- Greenberg, J.H. (1963). "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements". In: *Universals of Human Language*. Ed. by J.H. Greenberg. MIT Press, pp. 73–113.
- Hakken, L.A. and L.J. Whaley (2005). *Saving languages: An introduction to language revitalization*. Cambridge University Press.

RAPHY

University (n.d.). *Semantic Molecules*. Retrieved: 2023. URL: <https://intranet.curtin.edu.au/schools-departments/linguistics-semantic-metalanguage/what-is-nsm/semantic-molecules%20%7D>.

Levinson, J. and S. Levinson (1996). *Rethinking Linguistic Theory*. Cambridge University Press.

Roberts, C. (1992). "Dutch". In: *Journal of the International Phonetic Association* 22.1-2, pp. 45–47.

Wang, et al., (2019). "The reniform body: An integrative anatomical protocerebral neuropil complex of Eumalacostraca identified in Stomatopoda and Brachyura". In: *Journal of Comparative Neurology* 528.7, pp. 1079–1094.

Chomsky, N. (1987). *The Language Machine*. Cornell University Press.

Frege, G. (1975 [1812]). *Wissenschaft der Logik*. F. Meiner.

Heidegger, M. (2019 [1962]). *Being and Time*. Martino Fine Books.

Greenberg, J. and D. Nurse (2000). *African Languages: An Introduction*. Cambridge University Press.

Wong, D. (June 2015). *How Gay' Came to Mean 'Homosex-*

BIBLIOGRAPHY

- History Computer (2022). *The history of Gottfried Leibniz*. Retrieved: April 8, 2023. URL: <https://history-computer.com/the-history-of-gottfried-leibniz/>
- stadter, D.R. (2021). *Gödel, Escher, Bach: an Eternal Golden Braid*. Basic Books Inc.
- ase, J. (2003). "English as lingua franca: A threat to multilingualism?" In: *Journal of Sociolinguistics* 7.4, pp. 556–578.
- (2010). *Overt and Covert Translation*. John Benjamins Publishing Company, pp. 245–246.
- serl, E. (1993 [1900]). *Logische Untersuchungen*. Niemeyer.
- IUPAC (2021). *Nomenclature - IUPAC: International Union of Pure and Applied Chemistry*. Retrieved: April 9, 2023. URL: <https://iupac.org/what-we-do/nomenclature/>
- son, F. (1982). "Epiphenomenal Qualia". In: *The Philosophical Quarterly* 32.127, pp. 127–136.
- bson, R. (2018). "From Linguistics and Poetics". In: *The Norton Anthology of Theory and criticism*. 3rd ed. W.V. Norton & Company, pp. 1067–1074.
- ge, A. (2012). *Towards the Dynamic Art of Partial Comprehension*. Retrieved: June 4, 2023. URL: <https://www.alexander-ge.com/>

RAPHY

www.laetusinpraesens.org/docs10s/incompb.
7D.

L. (2018). *What is despeciation?* Retrieved: April
23. URL: %5Curl%7Bhttps://www.worldatlas.
articles/what-is-despeciation.html%C2%
D.

D. and J.M. Martin (2023). *Speech and Language*
Processing: 3rd edition. Published online.

1998 [1781]). *Kritik der Reinen Vernunft*. Wissenschaftliche
Gesellschaft Darmstadt.

Gong, and W.S. Wang (2008). “Language change
in social networks”. In: *Communications in Compu-
tational Physics* 3.4, pp. 935–949.

C. (2004). “Tamil”. In: *Journal of the Internationa
Linguistic Association* 34.1, pp. 111–116.

I. (2008). “Dueling context: A dynamic model of
linguistic change”. In: *Journal of Pragmatics* 40.3, pp. 385–406.

er, D. (2012). “The Cross-Linguistic Prevalence
of V and SVO Word Orders Reflects the Sequen-
tial and Hierarchical Representation of Action in Broca’s
Area”. In: *Language and Linguistics Compass* 6.1, pp. 50–

BIBLIOGRAPHY

- Swill, P. (1996). "Children, adolescents, and language change". In: *Language Variation and Change* 8.2, pp. 177–202.
- Gender, R. and M. Kutas (1993). "Subjacency as a processing phenomenon". In: *Language and cognitive processes* 8.4, pp. 573–633.
- pp, K. and C. Meierkord (2002). *Lingua Franca Communication*. Peter Lang EBooks.
- pp, S. and W.B. Michaels (1985). "Against Theory". In: *Against Theory: Literary Studies and the New Pragmatism*. University of Chicago Press.
- msch, C. (2014). "Language and Culture". In: *Aila Review* 27, pp. 30–55.
- ngelbach, M.L. and K.C. Berridge (2009). "Towards a functional neuroanatomy of pleasure and happiness". In: *Trends in Cognitive Sciences* 13.11, pp. 479–487.
- hniarevisch et al., (2015). "Genetic heritage of the Balto-Slavic speaking populations: a synthesis of autosomal, mitochondrial and Y-chromosomal data". In: *PloS one* 10.9, e0135820.
- ov, W. (1990). "The intersection of sex and social class in the course of linguistic change". In: *Language variation and change* 2.2, pp. 205–254.

RAPHY

- 0). *Principles of Linguistic Change, Internal Factors*. Wiley-Blackwell.
- ed, P. (1999). *American English: Handbook of the International Phonetic Association*. Cambridge University Press.
- ed, P. and I. Maddieson (1996). *The Sounds of the World's Languages*. Wiley-Blackwell.
- G. (1980). *Metaphors we live by*. University of Chicago Press.
- 0). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press.
- (2016). *Toki Ponal: la langue du bien*. Tawhid, p. 6.
- J.H. (2010). *Linguistic Relativities: language diversity and modern thought*. Cambridge University Press.
- (1997). *The Whorf Theory Complex: A Critical Reconstruction*. John Benjamins Publishing Company.
- D.A. Vakoch, and L.H. Wurm (1996). "Tone Perception in Cantonese and Mandarin: A cross-linguistic comparison". In: *Journal of Psycholinguistic Research* 25, 103–122.

BIBLIOGRAPHY

- elt, W. (2000). "Producing Spoken Language". In: *The Neurocognition of Language*. Ed. by C.N. Brown and P. Hagoort. Oxford University Press, pp. 83–122.
- inson, S.C. (1998). "Studying Spatial Conceptualization across Cultures: Anthropology and Cognitive Science". In: *Ethos* 26.1, pp. 7–24.
- isen, C. and S.E. Waters (2017). *Cultural Keywords in Discourse*. John Benjamins Publishing Company.
- is, G.L. (2001). *Turkish Grammar*. Oxford University Press.
- erman, A.M. et al. (1967). "Perception of the speech code". In: *Psychological Review* 74.6, p. 431.
- ell, H.G. (1894). *A greek-english lexicon*. Harper.
- geois-Chauvel, C. et al. (1999). "Specialization of left auditory cortex for speech perception in man depends on temporal coding". In: *Cerebral Cortex* 9.5, pp. 484–496.
- y, J.A. (1997). "The Linguistics of Colour". In: *Colour categories in thought and language*. Ed. by C.L. Hardin and L. Maffi. Cambridge University Press, pp. 320–436.

RAPHY

- a). *Grammatical Categories and Cognition: A Case of the Linguistic Relativity Hypothesis*. Cambridge University Press.
- (1968). *Introduction to Theoretical Linguistics*. Cambridge University Press.
- b). *Natural Language and Universal Grammar: Volume 1: Essays in Linguistic Theory*. Cambridge University Press.
- c). (1999). "Linguistic Intergroup Bias: Stereotype Attribution Through Language". In: *Advances in Experimental Social Psychology*. Elsevier B.V., pp. 79–121.
- d). D. and L.I. Bartolome (2014). "Multiculturalism Permitted in English only". In: *International Multilingual Research Journal* 8.1, pp. 24–37.
- e). Y, B. (2007). *Do Primary Colors Exist?* Retrieved: April 9, 2023. URL: %5Curl%7Bhttps://www.handprint.com/HP/WCL/color6.html%7D.
- f). H. (1984). *Patterns of sounds*. Cambridge University Press.
- g). b). *Structural Linguistics*.
- h). (n.d.). *Amplitude, wavelength, and frequency*. Retrieved: April 9, 2023. URL: %5Curl%7Bhttps://djmdme110.

BIBLIOGRAPHY

wordpress.com/2015/11/07/amplitude-wavelength-and-frequency/%C2%A0%7D.

isoff, J.A. (1991). "Sino-Tibetan linguistics: present state and future prospects". In: *Annual review of anthropology* 20.1, pp. 469–504.

thes, P.H. (1991). *Morphology*. Cambridge University Press.

uranen, A. (2003). "The Corpus of English as Lingua Franca in Academic Settings". In: *TESOL Quarterly* 37.3, p. 513.

echam, M. and J. Rees-Miller (2001). "Language in Social Contexts". In: *Contemporary Linguistics*. Ed. by M. Aronoff and J. Rees-Miller. 4th ed. Bedford/St. Martins.

garani, N. et al. (2008). "Phoneme Representation and classification in primary auditory cortex". In: *The Journal of the Acoustical Society of America* 123.2, pp. 899–909.

ge, B. and I. Leglise (2007). "Language and Colonialism". In: *Handbook of Language and Communication: Diversity and Change*. Ed. by M. Hellinger and A. Pauwels. De Gruyter Mouton.

RAPHY

- , C. and M. Hayes (2011). *A Biography of the English Language*. Cengage Learning.
- nn, Friederike (Mar. 2017). *Natural Language On-*. DOI: 10.1093/acrefore/9780199384655.013. URL: <https://oxfordre.com/linguistics/10.1093/acrefore/9780199384655.001.0001/fore-9780199384655-e-330>.
- and D. McCloy (2019). *PHOIBLE 2.0*. Max Planck Institute for the Science of Human History.
- ublishing Group (n.d.). *Nature news - species*. Re- ed: April 8, 2023. URL: %5Curl%7Bhttps://www. ce.com/scitable/definition/species-312/
- en, T. and E.C. Traugott (2012). *The Oxford Hand- of the History of English*. Oxford University Press.
- (2014). *Metaphysics: An Introduction*. Taylor & Fran-
- melin, Martine and Donnchadh O Conaill (2021). *lia: The Knowledge Argument*". In: *The Stanford elopedia of Philosophy*. Ed. by Edward N. Zalta. ner 2021. Metaphysics Research Lab, Stanford iversity.

BIBLIOGRAPHY

- man, J.M. (n.d.). *John Wilkins creates a universal language based on a classification scheme or ontology, and a universal system of measurement*. Retrieved: April 8, 2023. URL: %5Curl%7Bhttps://historyofinformation.com/detail.php?entryid=1874%C2%A0%7D.
- ton, B. (1997). "Language, Identity and the Ownership of English". In: *TESOL Quarterly* 31.3, p. 409.
- MLab (2021). *The Molecules Project: The Building Blocks of Meaning*. Retrieved: March 30, 2023. URL: %5Curl%7Bhttps://www.youtube.com/watch?v=hODJ7mr1rTQ&ab_channel=NSMLab%7D.
- radly, W.D. et al. (2001). *Contemporary Linguistics: An Introduction*. Palgrave MacMillan.
- enheimer, H. (2009). *The anthropology of language: an introduction to linguistic anthropology*. 2nd ed. Wadsworth, pp. 33–34.
- inowska, A. and J.M. Szewczyk (2019). "The more similar the better? Factors in learning cognates, false cognates and non-cognate words". In: *International Journal of Bilingual Education and Bilingualism* 22.8, pp. 974–991.
- lenko, A. (2006). *Bilingual Minds: Emotional Experience, Expression, and Representation*. Multilingual Matters.

RAPHY

- A. (1993). "A language which defies description
ordinary means". In: *The Role of Theory in Language
Description*. De Gruyter Mouton, pp. 87–130.
- A. (2005). *A Tajik-Persian reference grammar*. Brill.
- B. (1994). *The Language Instinct: How the mind cre-
ates language*. Penguin UK.
- . (1989). "Who are the speakers of Esperanto".
*Interlinguistics: Aspects of the science of planned lan-
guages*, pp. 157–172.
- , R. (1980). "Emotion: Theory, research, and ex-
perience". In: *Theories of emotion*. Vol. 1. Academic
Press New York.
- , R. and H. Kellerman (1980). *Theories of Emotion*.
Academic Press New York.
- and B. Seidlhofer (2006). "In and on their own
terms: the "habitat factor" in English as a lingua franca
interactions". In: *International Journal of the Sociology
of Language* 2006.177, pp. 151–176.
- translated by Owen, O.F. (1853 [300]). *Isagoge*.
retrieved: April 9, 2023. URL: [https://
www.tertullian.org/fathers/porphyry_isagogue_
translation.htm%7D](https://www.tertullian.org/fathers/porphyry_isagogue_translation.htm%7D).

BIBLIOGRAPHY

- ner, R. and M. Sala (n.d.). *Romance Languages*. accessed
may 2023.
- st, G. (2011). *An Introduction to Non-Classical Logic:
From if to is*. Cambridge University Press.
- ske, U. and S.C. Gandevia (2012). “The Proprioceptive
Senses: Their Roles in Signaling Body Shape, Body
Position and Movement, and Muscle Force”. In: *Phys-
iological Reviews* 92.4, pp. 1651–1697.
- um, G. (1991). *The Great Eskimo Vocabulary Hoax and
Other Irreverent Essays on the Study of Language*. Chicago
University Press.
- na, O. (2022). *The Sights and Sounds of Life: What are
Sensory Neurons?* Retrieved April 9, 2023. URL: %5Curl%
7Bhttps://www.tempobioscience.com/the-sights-
and-sounds-of-life-what-are-sensory-neurons/
%7D.
- emeyer, J. (n.d.[a]). *Ars Signorum*. Retrieved: April 8,
2023. URL: %5Curl%7Bhttps://www.frathwiki.com/
Ars_signorum%C2%A0%7D.
- n.d.[b]). *Oligosynthetic language*. Retrieved: April 8, 2023.
URL: %5Curl%7Bhttps://www.frathwiki.com/Oligosynthetic
language%C2%A0%7D.

RAPHY

- J. (2004). "British English: Received Pronunciation". In: *Journal of the International Phonetic Association* 34.2, pp. 239–245.
- n, A. (2009). *Lost Languages: the enigma of the worlds deciphered scripts*. Thames & Hudson, p. 61.
- A. (2002). "Critical/Sensitive Periods". In: *Child Development*. MacMillan Reference USA, pp. 101–103.
- M. (2019 [1805]). *Thesaurus*. Retrieved: April 9, 2021. URL: %5Curl%7Bhttps://web.archive.org/20130502153549/http://www.rain.org/~karpeles/transform.html%7D.
- r, Robin and Carlo Ierna (2022). "Christian von Wolff". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2022. Metaphysics Research Lab, Stanford University.
- K. (2001). "Language Death in Central Mexico: The Decline of Nahuatl and the New Bilingual Maintenance Programs". In: *The Bilingual Review* 26.1, p. 3.
- (1929). "The status of linguistics as a science". *Language* 5.4, pp. 207–214.
- roike, M. and Karen Barto (2017). *Introducing second language acquisition*. Cambridge University Press.

BIBLIOGRAPHY

- Chaffman, H. and S. Arokianathan (1986). "Diglossic Variation in Tamil film and fiction". In: *South Asian languages: Structure, convergence, and diglossia*. Vol. 10, pp. 371–381.
- Chomsky, B.A. (1982). "Variation in Esperanto". In: *Studies in the Linguistic Sciences* 12.1, pp. 183–196.
- Chomsky, D. (2012). *Neutralization*. Cambridge University Press.
- Simplifying Theory (n.d.). *Timbre*. Retrieved: April 9, 2023. URL: %5Curl%7Bhttps://www.simplifyingtheory.com/timbre/%C2%A0%7D.
- Chomsky, C. and K.S. Olson (2008). "Indonesian". In: *Journal of the International Phonetic Association* 38.2, pp. 209–213.
- Chomsky, R. et al. (2009). "Long-term crosslinguistic transfer of skills from L1 to L2". In: *Language Learning* 59.1, pp. 203–243.
- Chomsky, J.A. (1983). *The Mesoamerican Indian languages*. Cambridge University Press.
- Chomsky, M. (1952). *Lexico-Statistical Dating of Prehistoric Ethnic Contacts*.
- Chomsky, N. and V. Mann (1995). "The Limits of extended naturalistic exposure on the perceptual mastery of En-

RAPHY

/r/ and /l/ by adult Japanese learners of English”.

Applied Psycholinguistics 16.4, pp. 380–406.

J. (2000). “How relativistic are Humboldts ”Weltan-

en”?” In: *Explorations in Linguistic Relativity*. Ed.

. Pütz and M. Verspoor. John Benjamins Pub-

g Company, pp. 25–44.

. (1995). “Proto-Austronesian and the major Aus-

sian subgroups”. In: *The Austronesians*, pp. 19–

v, Y. (1979). “About Khlebnikov”. In: *The Futur-*

The Formalists, and the Marxist Critique. Ed. by C.

Ink Links, pp. 144–158.

oundation (n.d.[a]). *Relations*. Retrieved: April 8, 2023.

%5Curl%7Bhttp://www.undl.org/unlsys/unl/

005/Relation.htm%7D.

p]). *What is UNL*. Retrieved: April 8, 2023. URL:

rl %7Bhttp : / / www . undl . org / index . php ?

on=com_content&view=article&id=46&Itemid=

ang=en%C2%A0%7D.

rkhoof, R. (2005). “Belgian versus Netherlandic

h: New patterns of divergence? On pronouns of

ess and diminutives”. In: *Multilingua* 24.4, pp. 379–

BIBLIOGRAPHY

- hoeven, J. (2005). "Belgian Standard Dutch". In: *Journal of the International Phonetic Association* 35.2, pp. 243–247.
- steegh, K. (2014). *Arabic Language*. Edinburgh University Press.
- Humboldt, W. (1988 [1836]). *On language: the diversity of human language structure and its influence on the mental development of mankind* (trans. P. Heath). Cambridge University Press.
- son, O. (n.d.). *Ygyde Language - Comparison of philosophical languages*. Retrieved: April 8, 2023. URL: %5Curl%7Bhttp://www.orenwatson.be/saved/ygyde.htm%5C2%A0%7D.
- st, P. (2015). *Duodecimal Finger Counting: Counting to 60 on One Hand*. Retrieved: June 2, 2023. URL: ' 'url%7Bhttps://phillosoph.blogspot.com/2015/11/duodecimal-finger-counting-counting-to.html%7D.
- ite, J.M. (2002). *Everyday Life in Ancient Egypt*. Courier Dover, p. 175.
- itehead, A.N. (2010 [1929]). *Process and Reality*. Free Press Riverside.

RAPHY

- nn et al., (2009). *Where prosody meets pragmatics*. pp. 1–20.
- lia Foundation (2022). *Chromatic Circle*. Retrieved: 9, 2023. URL: %5Curl%7Bhttps://en.wikipedia.wiki/Chromatic_circle%C2%A0%7D.
- J. (1968 [1668]). *Essay Towards a Real Character Philosophical Language*. Scholar Press.
- stein, L. (2010 [1953]). *Philosophical Investigations*. y-Blackwell.
- C. and J. Fletcher (2007). *An Introduction to Phonology and Phonology*. Wiley-Blackwell.
- (2002). *Tone*. Cambridge University Press.
- C. (2009 [1823]). *An Account of Some Recent Discoveries in Hieroglyphical Literature, and Egyptian Antiquities*. John Murray.

Index

- charic, 92, 193, 197
- egocentricity, 147,
151, 153, 175
- erary, 113, 138, 140, 231
- etype letter, 77, 143
- fact, 12, 230
- s, 9–12, 38
- el, 237, 238, 249, 250,
269, 282
- ges, 140, 141
- ouche, 83–87, 211, 231,
292
- nistry, 145
- ate, 182–185, 193, 194,
196
- colour, 24–26, 31, 98, 148,
153, 154, 156–160,
214
- consciousness, 114, 137,
152, 154, 155, 171,
216
- consonants, 45, 47–50,
52–56, 63, 64, 69,
78, 129, 139, 267,
268
- direction of reading, 76
- duodecimal, 80, 81, 94, 294
- dyslexia, 87, 90, 92

erto, 16, 17, 137,
 42, 218, 235, 237
 1, 57, 154, 155,
 64, 169, 170, 176,
 35, 246, 251
 7, 28, 42, 58, 116,
 23–126
 h, 285
 3, 15, 32, 33, 40,
 6, 106, 107, 113,
 14, 117, 173, 182,
 84, 204, 234, 236,
 41, 242, 250, 251,
 53, 254, 262
 g, 117, 118
 onal auxiliary
 language, 7, 13–16,
 8, 117, 256
 onal Phonetic
 alphabet, 50, 63,
 07, 208
 91, 92
 5, 151, 231, 262,

264
 Korean, 52, 90, 103, 197,
 202
 language family, 182, 184,
 189–191, 193,
 195–200, 202
 language variation, 88, 249,
 261–264, 266, 268
 Lexi, 51, 194, 206–209, 263
 linguistic relativity, 18, 19,
 21, 23–31
 loanword, 78, 83, 84, 86,
 211, 230, 261
 Lojban, 29, 30, 106–109,
 112, 113, 194–196,
 205, 206
 mathematics, 11, 82, 160
 morphology, 45, 57, 58,
 108, 123, 263
 Natural Semantic
 Metalanguage,
 173, 174

INDEX

- erals, 78–81, 93, 121,
 - 122, 160, 163
- osynthetic, 128–130,
 - 133, 138, 155, 157,
 - 175, 228
- logy, 127, 132,
 - 135–137, 143, 146,
 - 147, 150, 151, 153,
 - 175
- ography, 7, 61, 86, 128,
 - 159, 173, 175
- imony, 7, 9, 128, 132,
 - 144
- nomenology, 127, 175
- osophical language, 7,
 - 13, 137
- nology, 3, 6, 45, 46, 49,
 - 50, 141, 208, 263,
 - 267
- matics, 234–236, 238,
 - 250, 263
- ody, 45, 56, 131,
 - 239–241
- ctuation, 77, 82
- qualia, 127, 151, 154–156,
 - 160, 171, 172, 175
- scent, 12, 19, 22, 41, 42, 51,
 - 52, 55, 87, 89, 91,
 - 135, 143, 144, 152,
 - 155, 165, 166, 188,
 - 191, 195, 196, 198,
 - 199, 202, 204, 228,
 - 233, 244, 260
- semantic atom, 77, 82, 128,
 - 129, 172, 174, 211,
 - 228–230
- semantic prime, 130, 132,
 - 134, 137, 142, 144,
 - 156, 163–168, 170,
 - 172, 173
- semantics, 127, 128, 132,
 - 172, 175, 236, 238,
 - 263
- sign language, 40, 256, 257
- smell, 154, 166, 215, 220,
 - 228, 258, 259
- sound, 3, 4, 7, 10, 12–14, 19,
 - 21, 24, 25, 27, 34,

2, 47, 51–53, 55,
 7–59, 65, 66, 68,
 9, 102, 104–107,
 09, 120, 122, 127,
 31, 133–135, 138,
 39, 141, 143, 146,
 50, 154, 155, 158,
 60, 161, 163–165,
 67, 168, 174, 175,
 82, 184, 191–193,
 99, 207, 208, 212,
 14, 217, 225, 230,
 32, 236, 243, 244,
 48, 249, 251, 253,
 54, 257–263, 266,
 79, 283, 286, 292
 173, 174
 , 49, 53, 62, 64,
 6, 68–72, 76–78,
 1–83, 90–93, 113,
 15, 128, 129,
 31–134, 142, 157,
 75, 210, 227, 229,
 39
 95–97, 113, 119,

132, 234, 251, 263,
 267

taste, 32, 105, 107, 154, 155,
 165–167, 192, 215,
 220, 258, 262

taxonomy, 135, 138–140,
 142–145

tertium comparationis, 15

TeX, 87, 93, 94

Toki Pona, 30, 86

tone, 44, 56, 57, 60, 62, 85,
 154, 191, 223, 273

translation, 15, 17, 32, 101,
 107, 132, 145, 147,
 182, 185, 196,
 207–209, 228, 230,
 231, 252, 267, 291,
 292

transliteration, 77, 93

typology, 118

vision, 3, 5, 34, 35, 40, 44,
 50, 56, 134, 154,
 155, 160, 190, 212,
 239, 268

INDEX

els, 42, 45, 47, 53–56,
63, 64, 78, 89, 92,
129, 139, 207, 260,

267, 268

Wilkins, 138–143

word order, 114, 263, 292

Atlan

The world being more globally connected than ever, an International Auxiliary Language (IAL) can be of great use for individuals from diverse backgrounds to communicate, collaborate, and understand each other. We, students from the University of Utrecht, embarked on an ambitious mission: to construct such an International Auxiliary Language. In a whirlwind of only a couple of weeks, we created Atlan: a philosophical language based on the principles of neutrality, unambiguity, and simplicity.

This book wants to take the reader on an odyssey with us in constructing this IAL. Analyze with us the fundamental principles of language and the complexities of phonological structures. Learn with us the original writing system, the carefully crafted ontology, and the lexicon that is generated from data from languages worldwide. Follow us as we overcome language barriers, transcend linguistic limitations, and in the process, unite the world in our words.

Step Antonisse

♫ ♫ J ɣ ɓ ʌ ʁ

Niek Elsinga

♫ ♫ ʃ ʁ ɣ ʌ

Max Geraedts

♫ ♫ ʌ ʃ ʌ ɣ

Stijn Janssens

♫ ♫ ʃ ʃ / ʃ ʃ ʃ

Jonathan Roose

♫ ♫ ʃ ʃ ʃ ʃ ʃ