

Samuel Pearce's
PROBE-VA
on the Topic of

UNIVERSAL GRAMMAR
&
LANGUAGE ACQUISITION

2021

Abstract

The goal of this project was twofold: for one half, I intended to expand my knowledge on the topic of Chomsky's Universal Grammar theory by reading an introductory book on the matter and writing a paper summarising my findings and opinion on the concept. For the other half, I conducted an informal experiment with two participants in which I taught them both the language Toki Pona¹ through two of the most common methods; one learned through pure immersion, while the other learned using the Toki Pona official course book. Finally, with the results of both halves of the project complete, I summarised my findings about universal grammar into a separate document and compiled the experiment results here alongside my reflection on the project as a whole.

¹Toki Pona is a minimalist language constructed by Sonja Lang in 2001. It has a very minimal lexicon which lends itself well to such a short experimentation period. I also am already fluent in it, which made it easy to immerse the participant.

Contents

1	Foreword	3
2	Introduction	3
2.1	Experiment Parameters	3
2.2	Hypothesis	3
2.3	Note on Glosses	4
2.4	Toki Pona Pronunciation	4
2.5	Pure Immersion Method	5
2.6	Vocab & Grammar Method	7
3	Results	8
3.1	Conveyance Test	8
3.2	Resultant Data	8
3.3	Discussion	8
4	Summary	9
4.1	Retrospective	9
4.2	Further Research	9
4.3	Conclusion	10
5	Work Schedule	10
5.1	Week 1	10
5.2	Week 2	10
5.3	Week 3	11
5.4	Week 4	11
6	Learning Journal	11
6.1	Time Management	11
6.2	Research Method	12
6.3	Experimentation Method	12
6.4	Writing Method	12
6.5	Future Improvements	12
7	Declaration of Independence	13

1 Foreword

The reason I chose this topic to study is because language is something that truly fascinates me. It is one of the very few important things that separates us from other animals and allows for humanity to not only preserve vital information across generations, but it also grants us many more opportunities to express our uniquely human creativity in the form of pure culture, which is — in my opinion — what makes life worth living.

I would like to thank V.J. Cook and Mark Newson for their book: Chomsky’s Universal Grammar: an Introduction. Without it, I would not have been able to complete half of this whole project. Their accessible but still comprehensive introduction was truly invaluable.

I also want to express my sincerest gratitude towards my two experiment participants, Amin Haidar and Julian Werner. They both took a large amount of time out of their busy schedules to help me with my project. Without them, the project would not have the same impact. For this, I thank them heartily.

2 Introduction

Firstly, I’d like to make a note that this document will primarily deal with the results of the experiment on language acquisition. The summary of universal grammar will be a separate document. This document will either be appended to the end of the current document or in a separate file, if being viewed digitally.

2.1 Experiment Parameters

The experiment consisted of teaching both participants the language Toki Pona. Participant A would learn the language the “traditional” way: using vocabulary and grammar textbooks, while participant B would be forbidden from having any words or concepts explained directly through a third language; everything would have to be learnt through gesture, pictures, and explanations using previously explained words. Once the roughly three weeks were over, I would record the participants and I having a conversation in order to analyse the participants’ fluency. Additionally, a test would be performed where I would give one participant some specific thing to describe and the other would have to explain to me what they were talking about to evaluate their explanation and comprehension skills more directly.

2.2 Hypothesis

My understanding of the language learning and language teaching process at the beginning of the experiment is quite limited. Until now, I have only ever tried teaching a language to my parents, which did not go very well. My expectations for how the process will work is that I will spend fixed periods of time with each participant explaining various concepts and being available for questioning as often as possible.

As for the teaching itself, I guess that conveying the extremely vague Toki Pona words’ meanings through immersion alone without confusing the participant will be quite difficult and could lead to many issues in comprehension.

Though, I assume the amount of practice afforded to the immersion participant would drastically improve their immediate comprehension and speaking ability. My idea for the traditional-learning method would be that the participant will very quickly pick up a larger amount of the lexicon with more concrete understanding than the immersion participant, but will struggle to recall and apply the grammar to them rapidly enough for speech.

As for how the end result would be, my personal hypothesis on the participants' grasp on the language would be that, in general, the immersion participant will be able to recall the most common vocabulary and how to properly implement them into the grammar far quicker than the traditional-learning participant; however, I guess that the traditional-learning participant will have a more correct usage of the language according to its official grammar than the immersion participant, while also being able to convey more complex topics due to their superior faculty over the vocabulary, but they would lack the speed required to utilise this knowledge in speech.

These are my hypotheses regarding how the experiment will go.

2.3 Note on Glosses

A few non-standard gloss abbreviations and elements are used throughout this document, so I have defined their meaning as well as the standard abbreviations here:

- The pictographic elements will be replaced in the glosses with simple descriptions of their contents and meaning in square brackets ([Image ; Intended Meaning]) E.g.: [Tree Image; Plants/Trees/Nature]
- Toki Pona includes context phrases that add extra information about the context of a sentence and build conditionals. The separator will be abbreviated as "CTXT" in glosses.
- All pronouns are abbreviated to just their person and number. E.g.: "You" → 2nd Person singular → "2SG"
- "PM" stands for the predicate marker which separates the subject from the rest of the sentence.
- "ACC" marks the accusative separator which separates the predicate from the direct-object.
- Additionally, some words can act as prepositions some times, but like verbs at other times. When a word is acting as a preposition, it has been marked as such. E.g.: "tawa" → "PREP.to"

2.4 Toki Pona Pronunciation

It's also worth pointing out how Toki Pona pronunciation works so the reader may correctly interpret the examples. If you know how to read the IPA², then you already know how Toki Pona is pronounced; every letter directly corresponds to its IPA equivalent, but given Toki Pona's limited phonology, pronunciation is not as important to get correct as it is in other languages. For those

²The International Phonetic Alphabet: <https://www.internationalphoneticassociation.org>

not familiar with the IPA, most consonants are pronounced as one would expect them to be: ‘t’ like ‘talk’, ‘p’ like ‘pet’, etc. The only special cases are ‘w’, which is always pronounced as in English (like the ‘w’ in ‘way’ or ‘water’), and ‘j’, which is always pronounced as in German (like the ‘y’ in ‘yes’ or ‘you’). The vowels are all consistent unlike English, which means they all have one sound:

- ‘a’ is always pronounced like it is in ‘car’ or ‘palm’
- ‘e’ is always pronounced like it is in ‘dress’ or ‘met’
- ‘i’ is always pronounced like the ‘ee’ in ‘fleece’ or ‘breeze’
- ‘o’ is always pronounced roughly like the ‘ou’ in ‘thought’
- ‘u’ is always pronounced like the ‘oo’ ‘goose’ or ‘choose’

Finally, in Toki Pona, the first syllable is always the stressed one:

- “toki” → “TOW-kee” (/ˈto.ki/)
- “soweli” → “SOW-weh-lee” (/ˈso.we.li/)
- “pimeja” → “PIH-meh-yah” (/ˈpi.me.ja/)

2.5 Pure Immersion Method

For both methods, the expectation that I would be able to spend regular sessions with the participants was quickly disregarded as unrealistic. Instead, due to the Coronavirus pandemic and our individual schedules, it made the most sense to sporadically practice the language through mostly digital interactions.

For the immersion participant, this entailed either discussing topics through a text-channel, which helpfully provided to opportunity to explain basic concepts through the use of pictographic imagery, and meeting — either in-person, or in a video game — the latter of which provided both ample example material and a basic system of gesturing which helped convey the meanings of my utterances.

In total, I would estimate that we spent about 9-10 hours speaking the language. The first few sessions were very difficult, as I had far fewer tools at my disposal to explain the concepts I was trying to convey, but once the most basic grammatical features were understood, I was able to bootstrap the participant’s understanding by progressively building on top of the previously understood material.

The first session took place in a text chat with the ability to send images. I attempted to establish the basic sentence structure and words to confirm their guesses with comprehensible pictographic substitution.

- (1) pona li [Thumbs Up; Good].
good PM [Thumbs Up; Good]
“Pona” (is) “good”.
- (2) ike li [Thumbs Down; Bad].
bad PM [Thumbs Down; Bad]
“Ike” (is) “bad”.

- (3) ala li [Red X-Shape; No/Incorrect].
 no PM [Red X-Shape; No/Incorrect]
 “Ala” (is) “no/wrong”.
- (4) lon li [Green Checkmark; Yes/Correct].
 yes PM [Green Checkmark; Yes/Correct]
 “Lon” is “yes/correct”.

Though there was some confusion at first as to the precise difference between “pona” and “lon”, this distinction, and many similar ones, would be cemented through examples and corrections. After the basic sentence framework was established, I tried conveying how the accusative marker worked through use of the transitive form of “to eat”

- (5) [Burger; Food] li moku.
 [Burger; Food] PM be-food
 “Food” (is) “moku”.
- (6) mi moku e moku.
 1SG eat ACC food
 I eat food.

This particle not having a direct translation became a frequent cause of frustration though, as the baseline we’d established was nowhere near enough to explain a complex grammatical feature. I persevered and assured them that they would gain an intuitive understanding of how it works through practice.

Soon enough, the participant had enough understanding that new concepts were generally easy to explain by relying on previous explained concepts. I was delightfully surprised by how much of the vocabulary they retained just by having to use it to ask questions and explain basic situations. The first moment that stuck out to me as probably the first sentence chain that could actually be considered a conversation was when I explained colours:

I began by asking if they wanted to learn the colours:

- (7) sina wile ala wile kama sona e nimi “kule”?
 you want not want receive knowledge ACC word “kule”?
 Do you want to learn the word “Colour”?

Which they affirmed with “lon”. So I continued with:

- (8) [Image of colour spectrum; Colour] li ‘kule’.
 [Image of colour spectrum; Colour] PM be-colour
 Colour is “kule”.
- (9) [Red Circle; Red] li ‘loje’.
 [Red Circle; Red] PM be-red
 Red is “loje”.
- (10) [Green and Blue Circle; Grue] li ‘laso’.
 [Green and Blue Circle; Grue] PM be-grue
 Grue (Green & Blue) is “laso”.

- (11) [Yellow Circle; Yellow] li ‘jelo’.
 [Yellow Circle; Yellow] PM be-yellow
 Yellow is “jelo”.

The participant’s knowledge was then confirmed by asking the following:

- (12) kasi li jo e kule seme?
 plant PM have ACC colour what?
 What colour do plants have?

To which the participant answered “laso?” which I affirmed with “lon!”.

Eventually, the most important grammar rules were understood and the bulk of the remaining words to be taught were simple verbs which could be explained through a picture accompanied by the phrase “This is X”. Over the course of the whole project, we periodically spoke in Toki Pona to reaffirm existing knowledge and learn new words. Another interaction which stuck out to me was a short remark I made after explaining how the number system worked:

- (13) tenpo kama, la mi sitelen e nimi sin tawa sina.
 time coming CTEXT 1SG write ACC word new PREP.to 2SG
 I will write (text) the new words to you.

Despite the fact that I accompanied the phrase with basic gestures to illustrate my point, I was quite taken aback at the fact that they understood it immediately.

2.6 Vocab & Grammar Method

The traditional-learning participant proved far easier to educate, simply because they were able to educate themselves with very little supervision; I gave them a copy of the official Toki Pona grammar book, which includes lessons about all the grammar as well as a dictionary and some sample texts.

I periodically checked in with the progress they were making, which was somewhat sluggish at first, but soon picked up, they began by learning the entire lexicon in one go. Given Toki Pona’s small size, this was the most practical approach. They stated their ability to understand the grammar through the examples, improved once they knew what all the words meant.

Though it was unintentional that we never spent any time practicing the language, I think

This participant estimates they spent 2-3 hours studying the language independently. It really showed off one of the major benefits of this method of learning: it’s completely possible to do without a teacher and allows the student to proceed at their own pace. The immersion participant lacked this independence, and even though the lessons were tailored to fit their pace of learning, this was only possible, because I only had one student. If you tried to teach any significant quantity of people the language, people would likely be forgotten.

3 Results

3.1 Conveyance Test

In order to gain some numerical data on the participants' comprehension and production abilities, I devised a test in which the participants must "play telephone"; I would give one of the participants a relatively simple concept which has no simple Toki Pona translation, which they would have to explain to the other and then afterwards, the listening participant would tell me what they thought the concept was. After that I could rank how well I thought it was explained on a scale of 1-5 and count the number of guesses it took before the listener correctly guessed the topic. After 4 rounds alternating who's explaining/listening, we can draw some conclusions about each participant's comprehension and explaining ability. As for rules, the participants are forbidden from using any language other than Toki Pona and may not say the the name of the topic.

3.2 Resultant Data

Below are the results from the conveyance test. Participants alternated who explained the concept. In the table below, A is the traditional-learning participant, and B is the immersion participant.

Topic	Explained to	Expl. Quality	Wrong Guesses	Fail/Pass
"video"	A \rightarrow B	4/5	X X X X X (5)	Fail
"train"	A \leftarrow B	5/5	✓(0)	Pass
"factory"	A \rightarrow B	3/5	X X ✓(2)	Pass
"toilet"	A \leftarrow B	4/5	X X ✓(2)	Pass

3.3 Discussion

Due to the differing time availability and studying methods of the two participants, what I gathered from the discussion before the test was that while participant A had a much wider and more complete knowledge of the vocabulary from having learnt them directly through flashcard software, participant B understood and knew how to use the grammar far more than participant A. This lead to the effect that A would typically know the most succinct and logical compound to use to describe the topic, but B often misunderstood it, and once A tried to elaborate on their explanation found it difficult due to a lack of grammar knowledge. On the other hand, participant B always knew how to construct grammatical sentences to describe the topic well and could often get A to understand through example scenarios and the like, A had difficulty understanding the full sentences.

Though the sample size is miniscule, the data still presents us with some interesting findings. For example, contrary to what one might immediately assume, explanation quality is not directly correlated with an easier time guessing. I also noticed the participants working together and using what knowledge they had to cooperatively come to an understanding; in one of the examples, the listener had understood the general premise, but wasn't sure which part of the described scene was the topic, and even though I didn't understand how, because it was ungrammatical, they both found a way to agree on the topic

correctly. In general, one can understand that being able to construct a scene which describes the topic in context is far easier to interpret than a more concise answer. This is somewhat reflected in the data by participant B's generally more easily understood explanations.

4 Summary

4.1 Retrospective

My original understanding of the challenges present with teaching a participant through immersion were generally correct; at one part of the experiment, participant A used the word "sitelen" to mean "image", while participant B had mostly only come into contact with it meaning "writing". If we'd had more time we could have made sure to go through each word and get to know every possible interpretation, but I'm still very happy with how quickly they acquired the language's basics. In the very short few times we met up to speak in toki pona, they were able to construct grammatically correct sentences and understand almost everything I said at normal talking speed. And something which really delighted me to hear near the end of the experiment was when I asked them if they knew how the accusative separator "e" worked and if they knew when it should be used to which he responded: "I just put it wherever it sounds correct". This was exactly what I wanted to achieve through immersion.

As for the traditional-learning participant, I was very surprised by how convenient the method was, I had assumed they would need as much 1-1 tutoring as the immersion participant, but they ended up learning everything they did independently. Though this carried with it the consequence that they often lacked motivation and needed to be reminded to study. It also meant that they got virtually no practice actually speaking, which made it difficult for them to construct sentences for the test. As I hypothesised at the beginning, this participant had a very wide knowledge of the words, but took longer to recall the most common ones, as they hadn't really had the opportunity to use them yet.

4.2 Further Research

Given my resources and time range, I'm quite happy with how the experiment went. If I were to re-do this project, there are a certainly few things that could be more scientifically performed:

- A larger group of participants
- Selecting participants I don't already know
- Testing/measuring more variables
- Longer learning period
- Carefully planned tests
- Periodic testing
- Planned course of learning

4.3 Conclusion

Finally, I'm quite glad to see my hypotheses mostly validated, but I'm even more glad to see that some of them were incorrect and that I was able to learn some valuable things from them. Quite unsurprisingly the main conclusion I've drawn from this project is that, as with most things, a balanced mix of both methods seems to be the best method for acquiring language. From the results of the test data and the numerous testimonials of my wonderful participants, I can also conclude that beginning by learning a large portion (80% grammar through comprehensible input is the most effective method. Once a baseline has been established to make sure you can understand most spoken or read utterances from context and the words you remember, the grammar will be automatically picked up through interaction and practice as well as the rest of the vocabulary. After that point, all that's left is to maintain the language through regular usage and continued vocabulary gain. I'm quite happy to have learnt the things I have from this project and am sure to implement them in future language-learning endeavours.

5 Work Schedule

5.1 Week 1

Planned Tasks

- Read UG literature
- Practice Toki Pona with participants

Performed Tasks

- Practice Toki Pona with participants
- Read UG literature

5.2 Week 2

Planned Tasks

- Write first draft of UG summary
- Practice Toki Pona with participants

Performed Tasks

- Practice Toki Pona with participants
- Read UG literature

5.3 Week 3

Planned Tasks

- Second, third, etc. drafts of UG summary
- Practice Toki Pona with participants

Performed Tasks

- Practice Toki Pona with participants
- First draft of UG summary

5.4 Week 4

Planned Tasks

- Finalise UG summary
- Interview and test experiment participants
- Compile results in final document

Performed Tasks

- Practice Toki Pona with participants
- Finalise UG summary
- Interview and test experiment participants

6 Learning Journal

6.1 Time Management

As with many of my projects, time management was a severe issue when it came to this project. I read the universal grammar textbook at the regular pace I would any other book and mainly focused on organising the experiment. This led to a huge amount of documentation work being pushed into the last week causing me to have to invest nearly all my free time into the project while also not being able to have the paper checked by others for glaring errors. Though I failed to realise the scope of the required documentation for the first half of the project, my time management skills improved significantly out of necessity during the last week-and-a-half; Each day was carefully scheduled to allow me to work on my Probe-VA, other TBZ projects, an IT module, and my regular work simultaneously. I value the experience gained from having to carefully plan my days, and am sure I'll put this newfound knowledge to good use during the VA next year.

6.2 Research Method

Overall, I'm quite happy with my research method, though mainly due to having a good source I was able to spend a good amount of time studying that I could reliably source when writing my summary of it. I must admit it became quite hectic when writing the summary to search for other sources to incorporate into the discussion, which lead to some sources being included without me being able to fully scrutinise them as much as I would have liked to. For the next VA, I will make sure to do more cursory research to gather a wide range of initial sources so I can take the time to analyse each one carefully.

6.3 Experimentation Method

Given the hands-off nature that was enforced by the pandemic and our differing schedules, I don't feel there was much else I could have or should have done to improve the accuracy of my experimental findings; as I mentioned above in 4.2, for a scientifically rigorous study, a much larger sample size would be required and more sterile working conditions, as well as a larger timeframe, none of which were available. Therefore, I'm very pleased with how well the experiment went despite its lackluster conditions.

6.4 Writing Method

As explained above, I ended up having to write the largest part of the documentation in just a few days which lead to a far sloppier writing process than I would normally maintain. Though, I didn't only learn the importance of managing one's time correctly to create a streamlined and highly polished writing pipeline, I also had to take some interesting new approaches when writing the summary, as I had never really had to rigorously cite my sources before which severely hindered any kind of progress at first, but once I got to know what tools I could use to more rapidly find good sources, cite them correctly, and search the physical book I had, my writing sped up by a significant amount.

6.5 Future Improvements

All of the many things I learned about the process of writing carefully cited papers and managing one's time during an important project are very important to me and I plan on implementing and developing these newfound skills more in my future projects, because time management especially has always been an issue of mine.

I guess this is why it's important that we have a *Probe* VA before the real thing to properly prepare our understanding of the task's scale and give us the required experience to tackle the real thing with newly developed skills. Overall, I'm very thankful to have been shown so starkly the consequences of poor time management as well as the important steps to take to avoid it in the future.

7 Declaration of Independence

I respect the intellectual property of other authors and do not claim their work as my own. I therefore clearly mark where I quote verbatim and also point out when I paraphrase or summarise the findings of others. This enables the reader to correctly assess the origin and quality of the information I have used. I make sure that the information I have obtained from others can be clearly distinguished from my own reflections and conclusions. Only then can my own performance be correctly assessed. I make sure that my bibliographical information is accurate enough to enable the reader to find the sources. I also clearly document scientific information obtained from the Internet according to the origin of texts and images with corresponding Internet addresses. I respect the authors' rights of my information sources and comply with the applicable legal regulations. I affirm that I have written my advanced thesis independently, taking into account the above rules.

Place & Date

Signature