INVESTIGATING THE EFFECT OF FREQUENCY ON THE ACQUISITION OF SUPERIORITY EFFECTS IN MULTIPLE INTERROGATIVES USING AN AGENT-BASED MODEL

Eve Whitaker

Linguistics, 2022

Abstract: The overall aim of this dissertation is to investigate the acquisition of superiority effects in multiple interrogatives by using an agent-based model. This would be to investigate what mechanisms a learner might use to learn the rules for superiority and what this might indicate about child language acquisition more broadly, particularly with respect to domain-general learning mechanisms. I investigated the frequency of multiple interrogatives in child directed speech from the Brown corpus in CHILDES and found that children have access to very little evidence of multiple interrogatives in the input (Brown 1973, MacWhinney 2000). This makes them particularly interesting in respect to language acquisition as this leads to the question, how do children acquire superiority effects in multiple interrogatives despite their rarity in the input? However, children have access to much more evidence of single wh-questions than they do of multiple interrogatives. To investigate whether children could acquire superiority effects through the input of single wh-questions I created an agent-based model where learners produce simplified multiple interrogatives based on the frequency of the individual wh-words in the input. I then compare the orders produced by the learners to the orders expected in multiple interrogatives in English, with the hypothesis being that the learners will produce orders expected in English. The results of the model showed that half of the orders learners produce were the same as those expected in English, but the other half were not. Overall, this study finds that frequency-based learning mechanisms which analyse just the input of individual wh-words is not sufficient for the acquisition of superiority effects in multiple interrogatives.

Keywords: agent-based models, superiority effects, multiple interrogatives, child language acquisition, domain-general learning mechanisms, wh-questions

Supervisors: Dr Christine Cuskley and Dr Johannes Heim

Contents

Investigating the effect of frequency on the acquisition of superiority effects in multiple interrogatives using an agent-based model

Eve Whitaker

1. Introduction

This study aims to test whether children could acquire superiority effects in multiple interrogatives, despite their rarity in the input, by using a domain-general learning mechanism rather than language specific one. Specifically, I implement a frequency-based learning mechanism in an agent-based model to test this theory.

Domain-general learning mechanisms are those which are utilised to acquire knowledge broadly, for example, memory is a domain-general ability that facilitates the learning of many different skills. This contrasts with domain-specific learning mechanisms which are utilised to acquire a specific skill, such as language. These mechanisms are language-specific.

Specifically, I will be looking at frequency-based learning mechanisms. Frequency-based learning mechanisms have been shown to play a role broadly in language learning and processing (Juszyk et al. 1994, Ellis 2002, Arnon and Snider 2010). More specifically, frequency has been shown to impact the word order of binomials, with the most frequent item often being placed before the less frequent item (Fenk-Oczlon 1989, Benor and Levy 2006).

The linguistic phenomenon I will be looking at is superiority effects in multiple interrogatives. Multiple interrogatives are questions which contain multiple question words such as, "Who bought what?". In English, superiority effects apply to multiple interrogatives. One of the wh-words must be fronted to the beginning of the sentence and the other stays put, as in (1). The order is also fixed in that it becomes ungrammatical if the opposite wh-word were fronted, as in (2)[1].

[1] The examples in (1) and (2) are from Grebenyova (2006).

(1) Who did John persuade to buy what?

(2) *What did John persuade who to buy?

The issue with multiple interrogatives is that they appear very infrequently in the input (Grebenyova 2006), so there is a question of how children acquire the superiority effects needed to produce them when they have so few chances to witness them. However, single wh-questions appear much more frequently in the input. I aim to test whether children could learn these effects by applying frequency-based learning mechanisms to the input they receive from single wh-questions.

The way I test this is by using an agent-based model. The use of agent-based models exists within the approach to language investigation of using computer models and computational mechanisms to recreate some phenomenon of language. Agent based modelling is often used in the social sciences but has been applied to language research as well. It is particularly useful for testing and formalising theories about populations as many agents can be used and can interact with each other (Baronchelli 2016, Cuskley et al. 2017, Cuskley et al. 2018) however it has also been applied to investigating the mechanisms behind individual language acquisition (Rumelhart & McVlelland 1986).

The model simulates a simplified acquisition of wh-words and of multiple interrogatives. In the first part of the model a learner acquires the mappings between the form of the wh-word, such as *what,* and the meaning, such as *object*. It does this by interacting with an "adult speaker" who has the correct form-meaning mappings and presents the learner with a wh-word. The learner then guesses the meaning and communication either succeeds or fails. If it succeeds, then the weight of that form-meaning mapping is increased. Through successive attempts the learner eventually settles on form-meaning mappings for all the wh-words. In the second part of the model the learner is given two meanings, e.g. *object* and *location,* and assigns two wh-words to them and then orders the words based on frequency, placing the most frequent first. This double production is a simplified multiple interrogative. I can then compare how the learner's double production compares to the orders of wh-words in English multiple interrogatives.

The overall aim of this dissertation is to investigate the acquisition of superiority effects in multiple interrogatives by using an agent-based model. This would be to investigate whether a

frequency-based learning mechanism that acts on the input of wh-words could be used by a learner to acquire the rules for superiority and what this might indicate about child language acquisition more broadly, particularly with respect to domain-general learning mechanisms. I aim to answer the question, "Can a frequency-based mechanism account for the acquisition of superiority effects in multiple interrogatives in English?"

I expect to find that:

1. the learners in the model will acquire the form-meaning mappings in the same order as wh-words are acquired in English.
2. the wh-words in the learners simplified multiple interrogatives will follow the same orders as seen in English multiple interrogatives.

In Chapter 2 I review the literature on domain-general learning mechanisms, particularly looking at frequency-based mechanisms, and find that the monitoring of frequency is used to acquire and process language broadly. I also review studies showing that frequency influences word order in binomials, with the most frequent item often appearing first. In section 2.2 I give a more detailed overview of superiority effects in multiple interrogatives and some of the issues in acquiring them. Through a corpus study I find that multiple interrogatives are very rare in the input a child experiences. In the same study I find that individual wh-words are much more frequent in the input. Finally in section 2.3 I narrow down my research question, state my hypotheses and propose that learners may acquire superiority effects by ordering the wh-words in multiple interrogatives based on the individual frequency of the wh-words.

In Chapter 3 I justify using an agent-based model to test my research question and go over some of the background and standards in using agent-based models. I explain what the ODD protocol is and use it to fully describe my model and its aims.

In Chapter 4 I describe my results finding that the learners successfully acquired all the wh-word meaning mappings and at different times. I also found that the learners produced their simplified multiple interrogatives in orders that matched English order half of the time. In Chapter 5 I discuss my results finding that they partially support the first hypothesis, that the learners will acquire the wh-words in a similar order to how they are acquired in English. The results did not support the second hypothesis, that the orders produced by the learners would match the orders in English. Finally in Chapter 6 I conclude that in terms of the research

question this means that this frequency-based learning mechanism is insufficient for learners to acquire superiority effects in multiple interrogatives.

2. Literature Review

2.1 Child Language Acquisition

Domain-general learning mechanisms refer to those which are utilised to learn and acquire knowledge broadly, whereas domain-specific learning mechanisms are used solely for the acquisition of a specific skill, such as language. An example of a domain-general ability is memory. Humans use memory to acquire knowledge broadly, but it is also essential to acquiring language. What this means is that some of the mechanisms children utilise during the language acquisition process are not built solely for learning language but also play a role in the acquisition of other kinds of knowledge.

This question of domain-general learning mechanisms in language acquisition is situated within the broader question of how do children acquire language? According to the Poverty of the Stimulus Argument the linguistic input a child is exposed to while they are learning is insufficient for the child to acquire a language from it alone. The input is infrequent and often produced with mistakes and interruptions and rarely formed to explicitly teach the child some part of the language. Additionally, as language is a uniquely human trait it is clear there is some biological ability that sets human learners apart from other animals. Children also go well beyond the input, fully acquiring the adult grammar of the language in their environment and being able to comprehend and produce an infinite number of utterances.

For this reason, it is clear children must have some biological ability that allows them to acquire language despite the nature of the input, however, the intricacies of how they do so is a matter of debate. This biological element is described as the language faculty within the theory of Universal Grammar (UG) and principles and parameters framework (P&P) (Chomsky 1981). Here, a child has access to universal grammar, an innate knowledge of all the grammars possible in human language, and sets parameters, syntactic binaries where languages differ, to fit the language in their environment (Laznik and Lohndal 2010).

Domain-general learning mechanisms have been theorised to be involved in the acquisition of language within this framework of UG and P&P. Chomsky (2005) outlines a three-factor model for acquisition where “general cognitive factors” work alongside UG and the linguistic input

during language acquisition. One aspect of these general cognitive factors that he discusses is types of data analysis that may be used for language as well as other forms of knowledge. One example he provides is that statistical methods of identifying words in the speech stream work only if the learner already has the knowledge that words have a single primary stress (Gambell and Yang 2003). He concludes that general methods of data analysis can interact with language-specific knowledge the learner has about the input in the process of acquisition. In this way, domain-general mechanisms work alongside language-specific mechanisms and the linguistic input to allow the learner to acquire a language.

Yang (2002, 2004) posits a variational approach to language acquisition that is anchored within traditional UG and P&P. In this approach, there are a population of grammars in the child's mind that are used to analyse the input. The grammars compete and those that analyse the input correctly are more likely to be selected to analyse the input in the future. At the end of this process, the grammars left are those that fit the language of that environment. Here, grammars are learned probabilistically, gradually, and possibly through domain-general learning mechanisms. This variational approach shows how domain-general learning mechanisms, specifically probability updating, could work alongside language-specific knowledge to acquire the grammar of a language.

Chomsky's three-factor model and Yang's variational approach provide examples of how domain-general mechanisms and language-specific mechanisms are not incompatible. It is not a question of whether the mechanisms behind language acquisition are domain-general or language-specific, but rather which mechanisms can explain which phenomena.

Monitoring frequency is a domain-general mechanism that is utilised by learners in acquiring and processing language in general. Ellis (2002) provides a review of how frequency is utilised to acquire features across language broadly, including acquiring lexical units, phonology, syntax and morphosyntax, and in production, among others. Ellis describes this monitoring of frequency as being more so a feature of synaptic connections, that forms over repeated exposure to linguistic input, rather than conscious counting and so it occurs automatically. They make a point that keeping track of frequencies cannot be the only mechanism that learners use, otherwise learners would only be able to comprehend and produce sentences that they had experienced before.

Regarding this point, I agree that frequency cannot be the only mechanism learners use, for the reason Ellis points to, but I also think that frequency can play a role in some aspects of acquisition and processing as shown by experimental evidence. For example, speakers process more frequent phrases faster than less frequent ones, even when the individual words of the phrases have similar frequencies (Arnon and Snider 2010). Arnon and Snider also showed that learners store the frequencies of compositional phrases, not just those of individual words. Frequency is also utilised by infant learners. In one study 9-month-old infants were found to prefer more frequent phonetic patterns over less frequent ones, and this preference didn't appear so much in younger infants (Juszyk et al. 1994). These experimental studies show that learners specifically process more frequent linguistic information faster.

Frequency has also been shown to be utilised in the acquisition and processing of word order, specifically in binomials. Binomials are three-word phrases where a pair of words are linked by a conjunction, usually *and*. Some examples of binomials are, "rich and famous" or "loud and clear". Often these phrases are spoken in a specific order and are rarely, or never, reversed without changing the meaning.

Fenk-Oczlon (1989) tested multiple constraints on the ordering of the items in freezes (binomials that cannot be reversed) and found that frequency of the individual words was the best predictor of the order of the words in the binomial. Specifically, the more frequent word would be placed first in the binomial and the less frequent placed after *and*. Fenk-Oczlon explained this phenomenon by suggesting that the speakers place the most familiar information first, because it's the quickest for the speaker to recall and produce, and because it avoids "peaks and troughs" of information, placing the newest information at the end of the phrase. It is important to note that this study only looked at freezes and while frequency was the best predictor, it didn't predict the order of every binomial correctly. However, it does suggest that frequency of individual words can have an impact on word order.

Benor and Levy (2006) conducted a similar study, testing multiple constraints on both freezes and reversible binomials. While they found that semantic and metric constraints were more effective than frequency at predicting order, they still found that frequency was a significant constraint, that usually more frequent items were placed first in the binomial. This makes sense since if the most frequent word being placed first is due to ease of recall, then this should also apply to reversible binomials, although it is worth noting, by their nature of being reversible it

is clear that other factors must be in play, not just frequency, and Benor and Levy's (2006) study shows this.

In conclusion, domain-general learning mechanisms and cognitive abilities that are not exclusive to language are needed for language acquisition. This is shown in the theories behind language acquisition and in experimental studies. Frequencies of different linguistic phenomena are used as a basis for processing and acquiring some parts of language. It is used in processing and acquiring language broadly (Ellis 2002) and more specifically frequency of individual items in binomials is a good predictor of the order in which they appear.

To test whether a frequency-based learning mechanism could be utilised by learners to acquire a part of their languages grammar we would need to look at constructions that rarely appear in the input but one where its constituent parts do.

2.2 Superiority Effects in Multiple Interrogatives

One such phenomenon which rarely appears in a child's linguistic input are multiple interrogatives. Multiple interrogatives are questions which contain two, or more, question words such as, "Who bought what?". Compared to single wh-questions, there are unique challenges in acquiring multiple interrogatives, particularly that multiple interrogatives have superiority effects where one wh-word is always fronted to the beginning of the sentence while the other stays put. The question is, how do children acquire this knowledge of superiority effects with so little evidence in the input?

To form accurate multiple interrogatives children must acquire the superiority effects that apply to them. In multiple interrogatives in English there is a fixed order in which the wh-words appear in the sentence,

(1) Who did John persuade to buy what?
(2) *What did John persuade who to buy?

The first question (1) where *who* comes first is much more natural to say than (2). In English, the superior wh-word, or phrase, is fronted while the other stays put (Grebenyova 2006). It is

important to note that multiple interrogatives are formed differently in different languages, for example in Russian all wh-phrases are fronted, not just one as in English (Grebenyova 2006, 2011). In this dissertation I will be focusing on multiple interrogatives in English.

Children encounter little evidence of multiple interrogatives in the linguistic input. Using the Brown corpus (Brown 1973), via CHILDES (MacWhinney 2000), I searched for instances of sentences containing two wh-words[2]. Out of 536 utterances like these 24 were true multiple interrogatives, such as "Who's doing what?" and "What germ are you checking how, Doctor?". This makes up around 0.05% of the input which contains multiple wh-words.

This learning problem is further complicated by the fact that multiple wh-words can appear in a sentence without it being a multiple interrogative. In fact, in the Brown corpus, there were more utterances that contained multiple wh-words that were not multiple interrogatives. Such as (B) where the *when* is not the question but rather referring to a specific time and (C) which isn't a question but contains both *who* and *what*. There were also many instances of repetition or echo questions like in (D) which I decided not to include in my total count of multiple interrogatives.

(B) What did we see when we went to Rhode Island?
(C) Who knows what that is.
(D) What's what?

Some of the utterances that are not multiple interrogatives but still contain multiple wh-words may still follow the rules of superiority effects, for example (C) with different intonation could easily be "Who knows what that is?". However, if (B) were to be rephrased as a multiple interrogative, "When did we see what at Rhode Island?", *when* would be fronted rather than *what*. So, not only are multiple interrogatives extremely rare in the linguistic input, but there are in fact more utterances which contain multiple wh-words which do not follow the rules of superiority effects.

Furthermore, instances of individual wh-words are much more frequent that multiple interrogatives. In the same corpus study, I found 18,453 instances of the words *how*, *what*,

[2] I used the CLANc programme and the command, "combo +s"@whq.txt^*^@whq.txt" +u +f *.cha -t*CHI:" to find all instances of two wh-words and the words between them. This command excludes data from the target child and so it only retrieves data from the linguistic input.

when, *where*, *who* and *why*, compared to 24 instances of multiple interrogatives[3]. So, while multiple interrogatives are rare in the input, individual wh-words are not. This leads to the question of whether children are learning how to construct multiple interrogatives based on the larger input they receive from single wh-questions and individual wh-words.

Grebenyova (2006, 2011) also found that children are exposed to far more single wh-questions than multiple interrogatives. They theorise that learners acquire multiple interrogatives by observing evidence from another part of the language, not from the multiple interrogatives in the input. Grebenyova (2006) investigates how English, Russian, and Malayalam-speaking children learn to form multiple interrogatives, as they are formed differently in each language. They conclude that the evidence the children receive comes from non-wh-constructions, and from these they can deduce whether the language has independent Focus projection or not. This explains how children use evidence outside of the input from multiple interrogatives to acquire the different forms in each language but not how they learn how to order the wh-words in multiple interrogatives.

Another thing to consider is the timings of the acquisition of the different wh-words, as multiple wh- words need to be used in multiple interrogatives. Clark (2003) lists the order of acquisition of wh-question forms as *where*, *what*, *why*, *who* and *when*. Experimental studies have found that children's acquisition of inversion in their why questions takes longer when compared to other wh-words (Labov & Labov 1978, Thornton 2008).

To summarise, wh-words in multiple interrogatives in English have a fixed order, only one of the wh-words is fronted and it must be the correct one otherwise the question is ungrammatical. Children acquire superiority effects in multiple interrogatives despite them rarely appearing in the input. By contrast, wh-words, and single wh-questions, appear much more frequently in the input.

2.3 Research question and Hypothesis

The question of interest is how can children acquire superiority effects in multiple interrogatives when they appear so rarely in the input? I propose that children could be

[3] I used the CLANc programme and the command, "freq +s@whq.txt +o +u +d2 *.cha -t*CHI:", to find all instances of the words *how*, *what*, *what'd*, *what's*, *when*, *where*, *who*, *who'd*, *who's*, *whose* and *why*.

acquiring the order of wh-words in multiple interrogatives based on the frequencies of the individual wh-words in the input.

To produce multiple interrogatives learners must acquire the superiority effects that fronts one of the wh-words in the question while the other stays put. The order of the wh-words is also fixed, the wh-word to be fronted cannot just be chosen at random otherwise the sentence becomes ungrammatical. However, multiple interrogatives appear so rarely in the input that it seems a more likely option that learners are acquiring this knowledge elsewhere. Unlike multiple interrogatives, single wh-questions and individual wh-words appear frequently in the child's linguistic environment. I intend to investigate whether learners could use the input from single wh-questions to acquire superiority effects and produce accurate multiple interrogatives.

Specifically, I propose investigating whether learners could produce accurate multiple interrogatives based on the frequencies of individual wh-words in the input. In section 2.1 I reviewed the literature on domain-general learning mechanisms in language acquisition. Domain-general mechanisms are incorporated alongside UG and the linguistic input in theories on language acquisition (Chomsky 2005, Yang 2002). Experimental studies also showed that domain-general mechanisms, specifically frequency monitoring, are used in the acquisition and processing of language broadly. Frequency also appears to influence word order, specifically in binomials. Fenk-Oczlon (1989) and Benor and Levy (2006) showed that often in binomials the most frequent word is placed first and the less frequent second, this applies to both reversible and non-reversible binomials. I aim to investigate whether learners could produce accurate multiple interrogatives if they used a rule that places the most frequent word first, like we see in binomials.

With this in mind, I narrow down the broader question of how children acquire superiority effects in multiple interrogatives despite their rarity in the linguistic input to the research question:

> Can a frequency-based learning mechanism account for the acquisition of superiority effects in multiple interrogatives in English?

I aim to answer this research question by creating an agent-based model where learners first acquire the wh-words, by learning the form-meaning mappings, and then produce simplified multiple interrogatives by producing two wh-words and ordering them based on their

frequencies in the input. Then I will compare how the learner's simplified multiple interrogatives compare to the orders of wh-words in English multiple interrogatives. I describe this model and its's aims in much greater depth in the next chapter.

I hypothesise that the orders of wh-words produced by the learners in their simplified multiple interrogatives will match those in English multiple interrogatives. The orders in English for the combinations of the wh-words *who*, *what*, *where*, *why* and *how* are:

Table 1. Orders of wh-words in English multiple interrogatives

why first	*how* first	*who* first	*what* first
(why, what)	(how, what)	(who, what)	(what, where)
(why, who)	(how, who)	(who, where)	
(why, where)	(how, where)		
(why, how)			

As well as this hypothesis I will also compare the order of the acquisition of the wh-word form-meaning mappings to the acquisition of wh-words in English as outlined by Clark (2003). This order is *where*, *what*, *why*, *who* and *when.* I expect the learners in the model to acquire the form-meaning mappings in this order.

3. Methodology and ODD protocol

3.1 Introduction to the ODD protocol and Agent-based models

The Overview, Design concepts and Details (ODD) protocol is a standardised format for describing agent-based models. The ODD protocol aims to provide a description of a model that is complete and accurate enough that it can be reproduced based on this description. It also aims to provide rationale for each decision made in the creation of the model (Grimm et al. 2020). The first three sections, Purpose and patterns (3.1), Entities, state variables and scales (3.2) and Process overview and scheduling (3.3) aim to provide an overview of the model. The following sections, Initialisation (3.4), Input data (3.5) and Submodels (3.6) provide full details of the model's processes. I have diverged slightly from the ODD protocol by including this introductory section and there would usually be a section for diagrams at the end of the protocol however I have placed these diagrams in section 3.6 so they can be seen next to their respective submodels. The code for the model is attached at the end of this paper in the appendix.

A key tenet behind the creation of many agent-based models is the minimality procedure (Conte and Paolucci 2014) which is also known as KISS – "keep it simple, stupid" – coined by Axelrod (1997). By prioritising minimality and creating a model with the minimal number of parts needed, very specific hypotheses can be tested that would be difficult to test in experimental studies. I follow this minimality procedure in a few ways in my model, mainly by simplifying the acquisition of wh-words to just form-meaning mappings and by simplifying multiple interrogatives to a double production of just two wh-words. By creating a simplified model, the problem is formalised in code and can be compared to experimental evidence and other theories on the problem.

3.2 Purpose and patterns

3.2.1 Purpose

The purpose of the model is test whether a frequency-based learning mechanism can explain the acquisition of superiority effects in multiple interrogatives.

The broader aim of the model, and the aim of my dissertation, is to investigate the acquisition of superiority effects in multiple interrogatives. In multiple interrogatives in English one of the wh-words is fronted to the beginning of the question, while the other stays put, and this occurs in a fixed order. Specifically, I will be looking at whether a frequency-based mechanism can explain the acquisition of superiority effects in multiple interrogatives. The mechanism I will be implementing is one where the most frequent wh-word is placed first and the less frequent one second.

The purpose of the model is to demonstrate a potential explanation for how a learner could acquire the different wh-words (who, what, where, when, why and how) and acquire superiority effects in multiple interrogatives in English. This also provides a specific phenomenon to test larger ideas about child language acquisition. The mechanism behind this is a frequency-based learning mechanism, where the learner learns the wh-words probabilistically and then applies the frequency with which it has encountered those words to forming simplified multiple interrogatives, which are referred to in this model as double production.

Since this model also builds on previous research and theory, it also serves the purpose of theoretical exposition and hypothesis generation. By creating a model based on existing theory and hypotheses and applying it to this specific phenomenon we can test whether the theory holds up for this phenomenon and if it doesn't, why and what further steps are needed to investigate this.

The model aims to investigate whether children can acquire something complex, like superiority effects in multiple interrogatives, that is not explicitly in the linguistic input and whether a frequency-based learning mechanism can achieve this. Multiple interrogatives rarely appear in the linguistic input, but single wh-questions appear frequently. I propose that a frequency-based learning mechanism can be used to generalise off this indirect input. With this model I aim to use a specific phenomenon to test larger questions about the nature of the mechanisms behind language acquisition and what kind of input is needed.

3.2.2 Patterns

I will evaluate this model by its ability to produce the following patterns:

i. the timings of the acquisition of wh-words

Children exhibit an uneven acquisition of wh-words. So, it would be expected that within this model the wh-words would not be learned simultaneously, they would be learnt at different times, as some words are going to appear more than other words in the input data. This pattern can be measured by looking at the data for all the learners and seeing when each wh-word form-meaning mapping was fully acquired.

ii. likelihood of the order of wh-words in multiple interrogatives

In English multiple interrogatives, the superior wh-phrase is always fronted to the beginning of the question while the other stays put. In this model we would expect to see some orders of wh-words being more likely than others. In fact, in this model, especially once the learners have acquired all the form-meaning mappings, there should only be one order being produced for each pair of wh-words. This leads to the final pattern expected:

iii. likelihood of the order being fixed

There is a fixed order for the wh-words in multiple interrogatives. This should be observed in the model as the learner is going to place more frequent wh-words before less frequent wh-words, so the order should be fixed for each individual learner. However, if the same order appears to be fixed over multiple, or every, learner then the model will have shown this pattern.

If the model produces these patterns, it will demonstrate that this learning mechanism produces the same or similar patterns seen in child language acquisition. This would suggest that a frequency-based mechanism like the one in the model is sufficient for a learner to acquire superiority effects in multiple interrogatives. It would lead to the suggestion that a more complex mechanism is not likely, though the specific mechanisms real learners use would have to be investigated further through other studies, such as experimental studies.

It will also be useful if the model does not produce these patterns. If the model produces patterns that don't resemble those seen in language acquisition, then it means a different

mechanism is being used. It suggests that a more complex mechanism is being utilised and that this one is too simple.

3.3 Entities, State Variables and Scales

3.3.1 Entities

The model has two entities: the learner, which represents a child acquiring their first language, and the speaker, which represents what would be many adult speakers in reality. The rationale for not including many adult speakers, as there would be in a realistic linguistic environment, is that I'm not looking at variation in the input in this model. The main reason for including multiple adult speakers would be to represent variation in the input different speakers provide, since I'm not looking at this one speaker is sufficient. The speaker could also be thought of as the linguistic environment. The data I use is from the Brown corpus (Brown 1973) of child directed speech and so it is taken from a real linguistic environment created by multiple adult speakers.

3.3.2 State Variables

The learner has state variables that represent the learner's hypothesis space. The learner has variables for connection between the wh-words and their meanings. These meanings are person, object, time, reason, location, and manner and these should eventually be mapped to the words who, what, when, why, where and how respectively. However, mappings can be drawn between any of these as it is all dependant on the input. The wh-word mapping variables do not have units.

Variable name	Type	Represents
WHO mapping	Dynamic, probability, range 0-1	The weight between the word *who* and its meaning.
WHAT mapping	Dynamic, probability, range 0-1	The weight between the word *what* and its meaning.
WHY mapping	Dynamic, probability, range 0-1	The weight between the word *why* and its meaning.

WHERE mapping	Dynamic, probability, range 0-1	The weight between the word *where* and its meaning.
HOW mapping	Dynamic, probability, range 0-1	The weight between the word *how* and its meaning.
Age	Dynamic, integer	The age of the learner/ number of interactions.

The speaker only has fixed state variables. It has wh-word mappings, like the learner except these are fixed throughout the programme.

3.3.3 Scales

This model does not represent space, it is not necessary for the model's purpose. The model does represent time as it models a learner acquiring wh-words and multiple interrogatives over time. This will be counted as the number of interactions and is the age variable of the learner.

3.4 Process Overview and Scheduling

1. Speaker selects a random wh-word from the input.
2. Learner executes its "learn" submodel, updating its state variables for the mapping selected and age.
3. The updated state variables, i.e., the weights for the mappings, are recorded.
4. Adult speaker selects two meanings at random, eg.) person, place.
5. Learner executes its "multiple interrogative" submodel, producing a list of two wh-words, e.g.) who, where.
6. The combination and order of the words in this list is recorded.
7. This process repeats until the learner has seen all the input.

3.5 Initialisation

There are two entities at initialisation: the learner and the adult speaker. The adult speaker has state variables for its wh-word form-meaning mappings that are fixed from initialisation. This represents adult speakers in the real world and the state of their I-language. They have already

acquired their language and so the connections between words and meanings are already set. The learner starts with every possible wh-word form-meaning mapping and these will update through the program. These mappings start at 0. This represents a child's I-language before they've been exposed to any input, so they have no prior knowledge of the connections between the words and the meanings. The age of the learner also starts at 0.

The input is based on data collected from the Brown corpus (Brown 1973) from CHILDES (MacWhinney 2000). I describe this data and how I collected it in section 3.6. The content of the input will not differ, so in this way the input it always the same at initialisation. However, the order of the input will differ in that it will be delivered to the learners in a random order on each run.

Initialisation in this model is generic and remains the same on each run, and among scenarios. This is because the differences in results will come from the different order in which the input is delivered to the learner on each run.

3.6 Input Data

The input is based on data collected from the Brown Corpus (Brown 1973) from the CHILDES database (MacWhinney 2000). The Brown corpus is a collection of transcripts recording the linguistic production of three English-speaking children aged 1;6 to 5;1. It also records the input they received during the recording sessions. Using the Brown corpus, I determined the frequency of each wh-word as they appear in child directed speech and then modelled the input the learner will receive based on this. I used the CLANc programme and the command,

```
freq +s@whq.txt +o +u +d2 *.cha  -t*CHI:
```

to find all instances of the words *how*, *what*, *what'd*, *what's*, *when*, *where*, *who*, *who'd*, *who's*, *whose* and *why*. The command also discounts data from the target child so only the child-directed input is searched. I combined the instances of *what*, *what'd* and *what's* into the total for *what*, and I combined *who*, *who'd, who's* and *whose* in the same way. I then had a total number of instances for each of the six wh-words which I divided by three since the corpus

recorded the input received by three children. This way each learner in the model receives the average input one child from the corpus received.

3.7 Submodels

3.7.1 Learn

This submodel is how the learner will acquire the mappings between the wh-words and their meanings. To trigger the learn submodel, the adult speaker selects a random wh-word for learner from the input.

1. Learner picks a mapping to connect a meaning to the wh-word in the input:
    a. If the value for all mappings is the same then a random one is chosen.
    b. Otherwise, learner chooses mapping with highest weight.

2. The communication is “successful” or “unsuccessful”:
    a. If the learners mapping matches the speakers mapping then communication is successful.
    b. Otherwise, communication is unsuccessful.

3. Learner updates probability of that mapping based on whether communication was successful; age also increases by one.
    a. If communication was successful then the probability for the selected mapping is increased.
    b. If communication was unsuccessful then the probability decreases for that mapping.

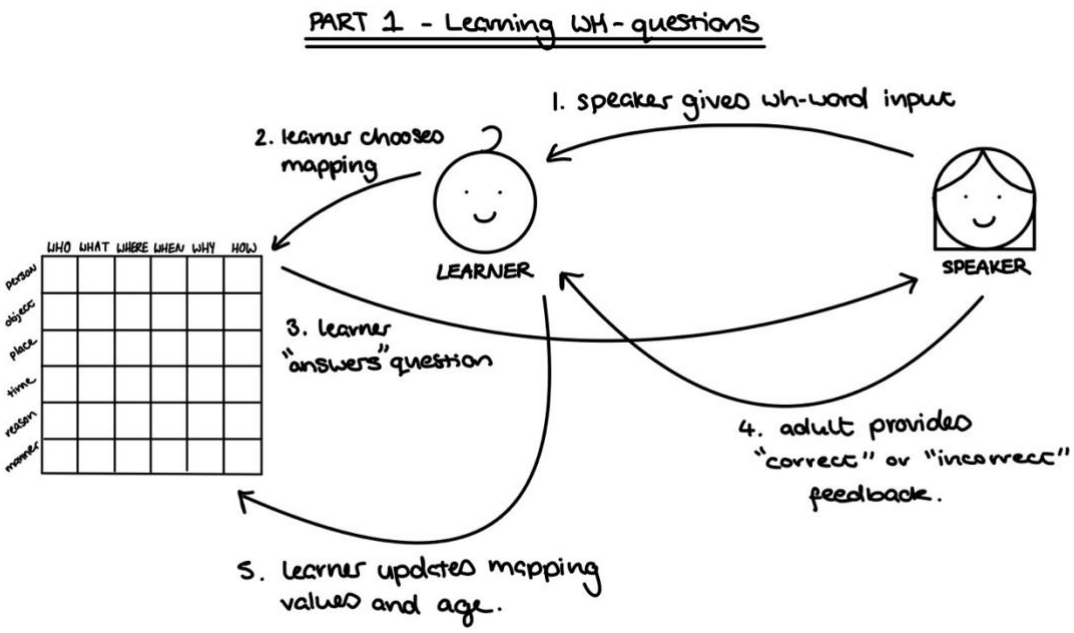

Figure 1. The learn submodel

3.7.2 Double production

This submodel represents a learner producing multiple interrogatives. This submodel immediately follows the learn submodel. The speaker presents a list of two meanings, for example, person and place.

1. Learner selects the two wh-words that have mappings to the meanings presented.

2. Learner creates a list of those two words with the most frequent wh-word encountered placed first and the less frequent second.
   a. If the frequencies are the same then order is random.

3. The combination and order of the list produced by the learner is recorded.

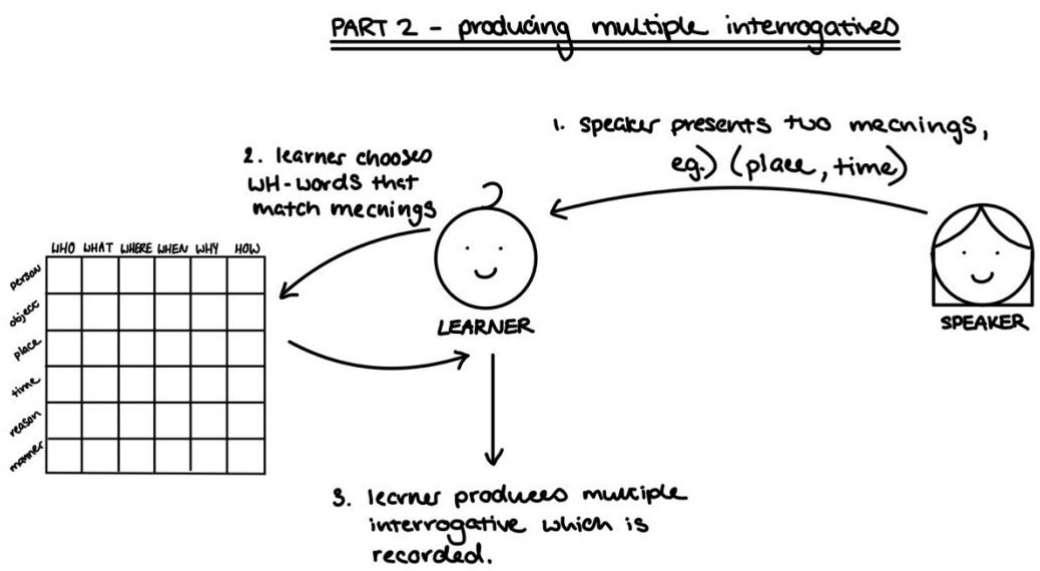

Figure 2. The double production submodel.

4. Results

4.1 Part 1: Learning the form-meaning mappings

In this section I am going to describe the results from the first part of the model, where the learners create mappings between the forms of the wh-words (what, who, why, where, how and when) and their meanings (object, person, reason, location, manner, time).

	WH-words					
	what	who	why	where	how	when
object	100	0	0	0	0	0
person	0	100	0	0	0	0
reason	0	0	100	0	0	0
location	0	0	0	100	0	0
manner	0	0	0	0	100	0
time	0	0	0	0	0	100

Table 2. Number of learners who have acquired each form-meaning mapping at the end of the code

Table 2 shows the number of learners with each form-meaning mapping at the end of the code. There were 100 learners in total and, as the table shows, all learners acquired the correct form-meaning mappings. All learners mapped *what* to *object*, *who* to *person*, *why* to *reason*, *where* to *location*, *how* to *manner* and *when* to *time*.

Next, Table 3 shows the progress of the agents learning over time. It shows the percentage of learners with the correct form-meaning mappings for at every 100 time stamps. Firstly, the form-meaning mappings for all the wh-words are acquired by all learners before time stamp

	WH-words					
Time Stamp	What	Who	Why	Where	How	When
0	0.22	0.13	0.16	0.22	0.11	0.23
100	1	0.75	0.75	0.69	0.78	0.6
200	1	0.95	0.93	0.9	0.97	0.81
300	1	0.99	0.98	0.99	0.99	0.94
400	1	0.99	1	0.99	1	0.96
500	1	1	1	1	1	1
600	1	1	1	1	1	1
700	1	1	1	1	1	1

Table 3. Percentage of learners with correct form-meaning mappings over time for each WH-word

500. Considering there are 6000 time stamps over the course of the code this means that the learners are acquiring the form-meaning mappings very quickly. Secondly, not all the form-meaning mappings are acquired at the same time. Notably, *what* is acquired much faster than the other words with all the learners acquiring this mapping before time stamp 100. Next, *why* and *how* are acquired, both before time stamp 400. At time stamp 400 98% of learners have acquired *why* and 99% have acquired *how*. Next, *who, where* and *when* are acquired at similar times, all before time stamp 500. At time stamps 300 and 400 99% of learners had the correct mappings for *who* and *where*, so they were largely acquired by the learners. *When* took slightly longer to learn, with 96% of learners having the correct mappings at time stamp 400 before it being fulling acquired before time stamp 500. From this the order of acquisition for these learners is *what* first, then *why*, then *how* soon after, then *when* and finally *who* and *where*.

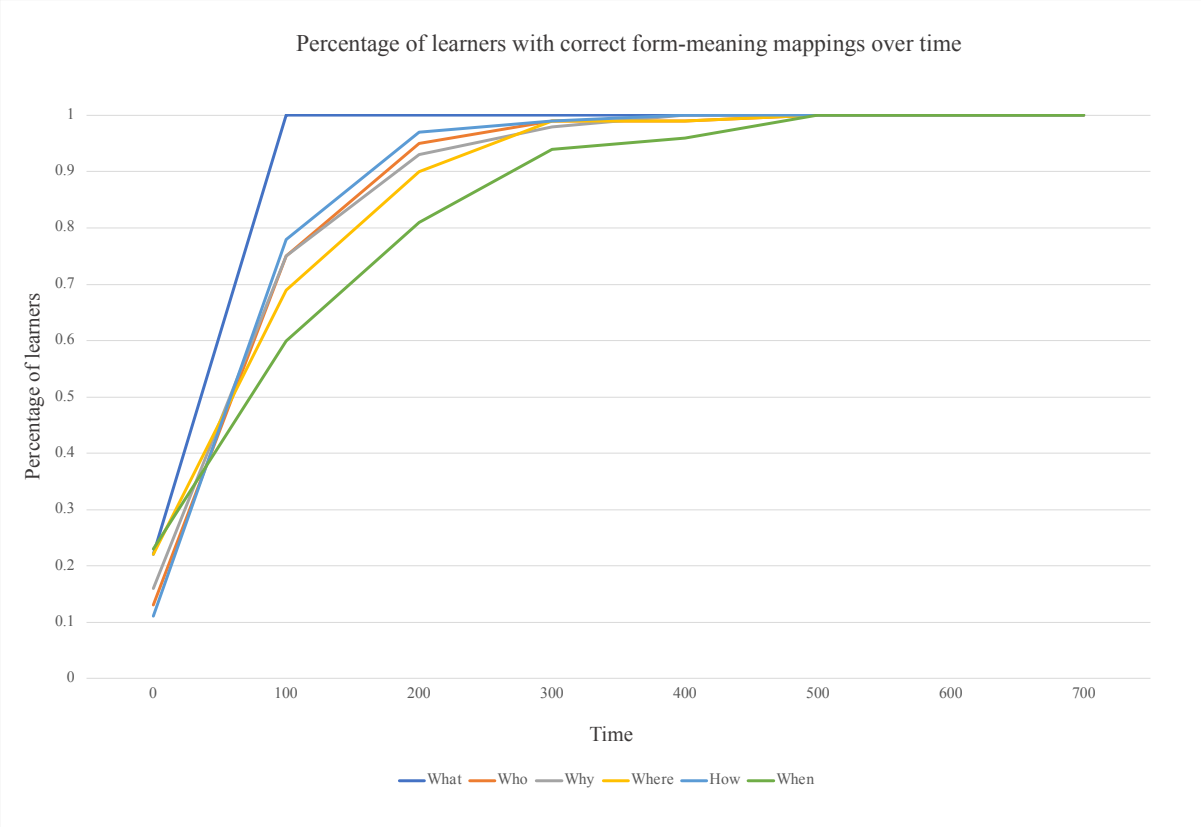

Figure 3. Graph showing the percentage of learners with correct form-meaning mappings over time

Figure 3 shows the data from Table 3 plotted onto a line graph. While there were 6000 time stamps, since all the learners acquired all of the mappings in under 500 time stamps I decided to only include the data until t=700.

4.2 Part 2: Double Production

Next, I am going to present the results from the second part of the model, where the learners produce simplified multiple interrogatives, by producing two wh-words and ordering them based on frequency. In this model the most frequent word will precede the less frequent word.

Pair	Number produced at t=6000
who-what	**0**
what-who	100
what-where	**100**
where-what	0
why-what	**0**
what-why	100
how-what	**0**
what-how	100
who-where	**100**
where-who	0
why-who	**0**
who-why	100
how-who	**0**
who-how	100
why-where	**99**
where-why	1
how-where	**63**
where-how	37
why-how	**99**
how-why	1

Table 4. The number of learners producing each possible pair of WH-words. The pair in bold is that which is expected in English.

Table 4 shows the pairs and orders of wh-words produced by learners during their double production. These results are taken from the double productions produced at the end of the code when the learners have encountered all the input. Firstly, the results show that for each of the pairs the learners overwhelmingly favour one order over the other. For the first five pairs all the learners are using the same order, for example, all are producing *what-who* and none are producing *who-what*. For two of the pairs, 99 of the learners are producing one order, *why-where* and *why-how*, and only one learner is producing the other order (note that this learner may not be the same learner in each instance). The one pair that stands out is *how-where/where-*

how. Here, 63 of the learners are producing *how-where* and 37 are producing *where-how*. One order is still favoured over the other, however, unlike the other pairs, one order hasn't been acquired as a rule by the whole population. Figure 3 shows the data from Table 3 in a bar chart. Again, this shows how for most of the wh-word pairs one order is near universally favoured apart from the *how-where/where-how* pair.

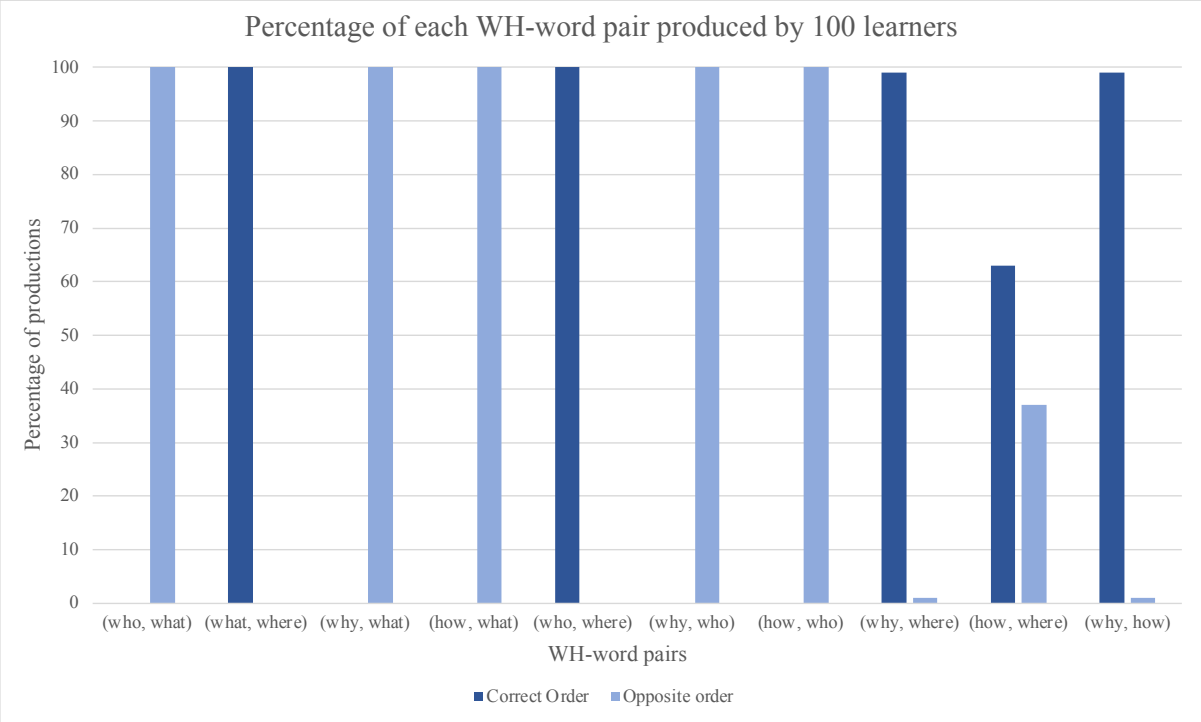

Figure 4. Graph showing the data from Table 3. This shows the number of learning producing each pair of WH-words at the end of the code.

The results in Table 4 and Figure 4 also show that the orders favoured by the learners in the model do not always correspond to the expected order in English. Half of the orders favoured by the learners are also those expected in English. These are *what-where*, *who-where*, *why-where*, *how-where* and *why-how*. The other half of the orders favoured by the learners are the opposite of what would be expected in English. These are *what-who, what-why*, *what-how*, *who-why* and *who-how.*

The first four pairs in the table contain *what* and here the learners always choose the order than places *what* first. In pairs that contain *who*, learners place *who* in first position except when it is paired with *what.* In pairs containing *why* the learners place it before *where* and *how* but after *what* and *who*. With *how* the learners only place it in first position before *where*. With all other pairs *how* is placed second. Finally, when *where* is in the pair the learners never favour orders which place it in first position.

5. Discussion

5.1 Discussion of Results

The research question asked was “can a frequency-based mechanism account for the acquisition of superiority effects in multiple interrogatives in English?”. To investigate this question, I created a model which simulated a simplified acquisition of wh-words and production of multiple interrogatives. The learner gradually acquires the mappings between the form of the wh-word and the meaning. Simultaneously, at each time stamp, the learner is given two meanings, and assigns two wh-words to them based on the form-meaning mappings they have acquired at that point. They then order the words based on frequency, placing the most frequent first and producing a simplified multiple interrogative. My two hypotheses that aim to answer the research question are:

1. the learners in the model will acquire the form-meaning mappings in the same order as wh-words are acquired in English.
2. the wh-words in the learners simplified multiple interrogatives will follow the same orders as seen in English multiple interrogatives.

This model also had three patterns I expected to observe, which were:

i. the timings of the acquisition of wh-words: different wh-words are going to be learned at different times.
ii. likelihood of the order of wh-words in multiple interrogatives: for each pair one order will be more likely than the other.
iii. likelihood of the order being fixed: by the end of the code only one order per pair should be produced by all the learners.

I am going to discuss my results by submodel, so I will discuss the results from the acquisition of the wh-words first and then the results from the double production. Some of my hypotheses and patterns are addressed by the same results so I will discuss them together rather than repeating the same results. Overall, I find that the first two patterns I expected to observe are present in the results, but the third pattern isn’t, the order was not fixed for all the wh-word pairs. In terms of my hypotheses, I found that my first hypothesis was somewhat supported,

the learners largely acquired the form-meaning mappings in the same order as they would be acquired in English. However, I found that my second hypothesis was not supported by the results, the learners did not choose the same orders as would be expected in English for their simplified multiple interrogatives.

First, I am going to discuss the results from the first part of the model, the learn submodel, which I described in section 4.1 in Chapter 4. In this part of the model the learners create mappings between the forms of the wh-words (what, who, why, where, how and when) and their meanings (object, person, reason, location, manner, time). These results aim to answer the first part of my hypothesis, the learners in the model will acquire the form-meaning mappings in the same order as wh-words are acquired in English. The results should also show the first pattern I expected to see in my model, that the different wh-words will be learned at different times.

Firstly, the results show that the learners successfully acquire the correct form-meaning mappings for all the wh-words. The results also support the first pattern I expected to see in the code, that the wh-words would be learned at different times. For example, *what* is learned first, before time stamp 100, whereas *why* and *how* take until time stamp 400 to be acquired.

In terms of my first hypothesis, I expected the learners to acquire the form-meaning mappings in the same order than children acquire wh-words in English. The results somewhat support this hypothesis. The order of acquisition in English is *where*, *what*, *why*, *who* and *when* (Clark 2003). The learners in the model learnt the form-meaning mappings in the order *what*, *why*, (*how*), *who*, *where* and *when.* In both sequences, the words *what*, *why*, *who* and *when* are learned in that order, so in this way the learners in the model are acquiring the form-meaning mappings in the same order as the acquisition of wh-words in English. The big difference is that *where* is learned first in English but in the results it is one of the last words acquired.

These results show that the learners in the model largely follow the same order of acquisition for wh-words in English, with the timing of the acquisition of *where* being significantly different. This indicates that the probability-based mechanism the learners use in the model is largely sufficient for acquiring the wh-words in the same way as in English, though it mustn't be entirely the same.

Next, I am going to discuss the results from the second part of the model, the double production submodel, which I described in section 4.2 in the previous chapter. Here the learners produce simplified multiple interrogatives, by producing two wh-words and ordering them based on frequency, with the more frequent word preceding the less frequent word. These results aim to address the second part of my hypothesis, that the wh-words in the learners simplified multiple interrogatives will follow the same orders as seen in English multiple interrogatives. The results should also show evidence of the second and third patterns I expected to see in my model, that ii.) that for each pair one order will be more likely than the other and iii.) by the end of the code only one order per pair should be produced by all the learners.

The results show that for all pairs one order is much more likely than the other, supporting the second pattern I expected to see. The results also largely support the third pattern I expected to see, that by the end of the code, once the learners had seen all the input, all the learners would be using only one order for each pair. This applies to most pairs but not all. For seven out of the nine pairs all the learners are using just one order. For two of the pairs 99 of the learners are using one order and just one learner is using the opposite order. A significant discrepancy comes with the *how-where/where-how* pair, where 63 of the learners are using *how-where* and 37 are using *where-how*. I believe this result comes from the fact that the frequencies for *how* and *where* are the most similar in the input. The input was collected from data from the Brown corpus (Brown 1973) where how appeared 1338 times and *where* appeared 1336 times. Since the corpus features three children, I divided the instances for all the wh-words by three to find an average for each child which brought the instances for *how* to 446 and *where* to 445. I think this brought the frequencies of *how* and *where* so close together that some learners have *how* as the most frequent and some have *where* as the most frequent. I think this is likely to be the case because the order *how-where* is the most popular among the learners and *how* is more frequent than *where* in the input, just not enough to create a consensus among the learners.

In terms of my second hypothesis, I expected the wh-words in the learners simplified multiple interrogatives to follow the same orders as seen in English multiple interrogatives. The data does not support this hypothesis. For five of the pairs the English order was the most popular but for the other five the opposite order was most popular.

To summarise, I found that the first two patterns I expected to observe in my model are present in the results, but the third pattern isn't, the order was not fixed for all the wh-word pairs. I

found that my first hypothesis was supported somewhat, the learners largely acquired the form-meaning mappings in the same order as they would be acquired in English. In terms of my second hypothesis, I found that it was not supported by the results, the learners did not choose the same orders as would be expected in English for their simplified multiple interrogatives.

The results indicate that the mechanism used by the learners to acquire the form-meaning mappings is somewhat sufficient to acquire these mappings in a way that is similar to how they are acquired in reality. However, the frequency-based mechanism used by the learners to produce the simplified multiple-interrogatives is not sufficient to produce multiple interrogatives that would be accurate in English.

5.2 Significance of results and the limitations of the study

The aim of the study was to answer the research question, "Can a frequency-based mechanism account for the acquisition of superiority effects in multiple interrogatives in English?". The results show that the frequency-based mechanism implemented in the model was not sufficient to produce superiority effects in simplified multiple interrogatives that resembled those seen in English. So, to answer the research question, a frequency-based learning mechanism does not seem likely to be able to account for the acquisition of superiority effects in multiple interrogatives in English, at least for the specific mechanism implemented in this study's model.

Potentially, this means that this learning mechanism is too simple, and that the learner is using more complex learning mechanisms to acquire the superiority effects. One thing that this model removes from the learning problem is any complexities of syntax, the wh-words are all learned in the same way, as though they have the same complexity, and the simplified multiple interrogatives are a simple two-word sequence ordered by frequency. Differences in the syntax of different wh-words and their questions could impact how the learner acquire the wh-words and could impact the acquisition of multiple interrogatives. For example, Thornton (2008) theorises that the reason why children acquiring English fail to use subject-aux inversion in their why-questions, despite acquiring the same rule in other wh-questions, is due to the children accessing a grammar found in Italian. There are specific learnability problems for each wh-word and single wh-questions that are not accounted for in this model, all the wh-words are treated the same. By simplifying down the form of the single wh-questions to individual

wh-words and the multiple interrogatives down to sequences of two wh-words any syntactic complexities a learner could gain information from are unavailable.

Another possibility is that children learning superiority effects in multiple interrogatives are gaining their knowledge from different input. Grebenyova (2006) theorises that children acquire multiple interrogatives by observing evidence from non-wh-constructions. This model looks at whether a learner can acquire superiority effects in multiple interrogatives from the input provided by single wh-questions, not from any constructions other than these.

The study is also limited in its scope as it only investigates multiple interrogatives in English, so it only rules out this frequency-based learning mechanism for English, not for other languages. Multiple interrogatives are formed differently in other languages (Grebenyova 2006, 2011) so further research will have to be done to determine whether this learning mechanism is sufficient for the acquisition of multiple interrogatives and superiority effects in languages other than English.

These results contribute to the field by showing that, with this specific mechanism and using individual wh-words as input, this frequency-based mechanism is not sufficient for the acquisition of superiority effects in English multiple interrogatives. Possible areas for future research would be to investigate languages other than English, to incorporate more of the complexities of the syntax into the input and the production of the multiple interrogatives, and to look at whether learners could acquire superiority effects from input other than single wh-questions.

6. Conclusion

This study aimed to investigate how a learner could acquire superiority effects in multiple interrogatives despite their rarity in the input. Specifically, it aimed to see whether a frequency-based learning mechanism would be sufficient to acquire this. Through a corpus study I found that multiple interrogatives are rare in the input, but individual wh-words are not. I reviewed literature on domain-general learning mechanisms and frequency-based learning mechanisms which showed that learners monitor frequency automatically and that this assists in acquisition and processing. The word order of binomials is also influenced by frequency with the most frequent item appearing before the less frequent one (Fenk-Oczlon 1989, Benor and Levy 2006). I proposed that children could be acquiring the knowledge needed to produce multiple interrogatives from single wh-questions which are much more abundant in the input.

To test this proposal, I planned to create an agent-based model and I outlined two hypotheses to test whether the results it produced supported the research question or not. The two hypotheses were:

1. the learners in the model will acquire the form-meaning mappings in the same order as wh-words are acquired in English.
2. the wh-words in the learners simplified multiple interrogatives will follow the same orders as seen in English multiple interrogatives.

The results of the model partially supported the first hypothesis, all the wh-word from-meaning mappings were acquired, and most were acquired in the same order as they would be by children acquiring them in English, apart from one. However, the results did not support the second hypothesis. The orders of wh-words the learners produced in their simplified multiple interrogatives were not the same as those seen in English. Half of the orders produced by the learners were orders that would appear in English, but the other half were not. With regards to the research question, “Can a frequency-based mechanism account for the acquisition of superiority effects in multiple interrogatives in English?”, this model’s frequency-based mechanism is insufficient for the learners to acquire multiple interrogatives.

This study contributes to the field by showing that, using this specific frequency-based mechanism, and using individual wh-words as input, this frequency-based mechanism is not

sufficient for the acquisition of superiority effects in English multiple interrogatives. Possible areas for future research would be to expand this investigation into languages other than English. Future research could also incorporate more of the complexities of the syntax into the input and the production of the multiple interrogatives, and to look at whether learners could acquire superiority effects from input other than single wh-questions.

References

Arnon, I. and N. Snider. 2010. More than words: frequency effects for multi-word phrases. *Journal of Memory and Language*. 62: 67-82.

Axelrod, R. 1997. *The Complexity of Cooperation: Agent- Based Models of Competition and Collaboration.* Princeton: Princeton University Press.

Baronchelli, A. 2016. A gentle introduction to the minimal Naming Game. *Belgian Journal of Linguistics.* 30: 171-192.

Benor, S.B. and R. Levy. 2006. The chicken or the egg? A probabalistic analysis of english binomials. *Language*. 82: 233-278.

Brown, R. 1973. A first language: the early stages. Cambridge: Harvard University Press.

Chomsky, N. 1981. *Lectures on Government and Binding*. Berlin, New York: Mouton de Gruyter.

Chomsky, N. 2005. Three Factors in Language Design. *Linguistic Inquiry*. 36: 1-22.

Clark, E. 2003. *First Language Acquisition*. Cambridge, New York: Cambridge University Press

Conte, R. and M. Paolucci. 2014. On agent-based modelling and computational social science. *Frontiers in Psycology.* 5: 1-9.

Cuskley, C., C. Castellano, F. Colaiori, V. Loreto, M. Pugliese and F. Tria. 2017. The regularity game: investigating linguistic rule dynamics in a population of interacting agents. *Cognition.* 159: 25-32.

Cuskley, C., V. Loretto, and S. Kirby. 2018. A social approach to rule dynamics using an agent-based model. *Topics in Cognitive Science*. 10: 745-758.

Ellis, N.C. 2002. Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition. *SSLA*. 24: 143-188.

Fenk-Oczlon, G. 1989. Word frequency and word order in freezes. *Linguistics*. 27: 517-556.

Gambell, T. and C. Yang. 2003. Scope and limits of statistical learning in word segmentation. Ms., Yale University, New Haven, Conn.

Grebenyova, L. 2006. Multiple interrogatives: syntax, semantics, and learnability. Doctoral Dissertation. University of Maryland.

Grebenyova, L. 2011. Acquisition of multiple questions in english, russian, and malayalam. *Language Acquisition.* 18: 139-175.

Grimm, V., S.F. Railsback, C.E. Vincenot, U. Berger, C. Gallagher, D.L. DeAngelis, B. Edmonds, J. Ge, J. Giske, J. Groeneveld, A.S.A Johnston, A. Milles, J. Nabe-Nielsen, J.G. Polhill, V. Radchuk, M. Rohwäder, R.A. Stillman, J.C. Thiele, and D. Ayllón. 2020. The ODD protocol for describing agent-based and other simulation models: a second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation.*

Jusczyk, P.W., P.A. Luce and J. Charles-Luce. 1994. Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*. 33: 630-645.
Labov, W. and T. Labov. 1978. 'Learning the Syntax of Questions'. In Campbell R.N., Smith P.T. (eds) *Recent Advances in the Psychology of Language*. Boston: Springer. 1-44.
Lasnick, H. and T. Lohndal. 2010. Government-binding/principles and parameters theory. *WIREs Cognitive Science*. 1: 40-50.
MacWhinney, B. 2000. The CHILDES Project: Tools for analyzing talk. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
Rumelhart, D., G. Hinton and J. Mcclelland. 1986. A General Framework for Parallel Distributed Processing. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. 1.
Thornton, R. 2008. Why continuity. *Natural Language & Linguistic Theory*. 26: 107-146.
Yang, C. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.
Yang, C. 2004. Universal grammar, statistics or both?. *Trends in Cognitive Sciences*. 8: 451-456.

Appendix

```
import numpy as np
np.set_printoptions(suppress=True)
import random
# Form indices: 0-who, 1-what, 2-when, 3-where, 4-why, 5-how
# Meaning indices: 0-person, 1-thing, 2-time, 3-location, 4-reason, 5-manner
# list is [interactions, successes, lastinter, weight,currentWindow]
totalInters=6090#currently made up; will scale with "corpus size"; 10k instances of what roughly in corpus; 10k/161=62*approx250=15000

adultMaps={"what":0,"who":1,"why":2,"where":3,"how":4,"when":5}
conceptMaps={0:"object",1:"person",2:"reason",3:"location",4:"manner",5:"time"}
formMaps={0:"what",1:"who",2:"why",3:"where",4:"how",5:"when"}


useMemory=False
canalizationPoint=5
startWindow=6090
#Memory isn't working here because there is no way to gain forms; there is no explicit feedback,
freqs={"who":499,"what":3959,"where":445,"why":467, "how":446, "when":274}#6090 tokens total
#freqs={"who":0.08,"what":0.65,"where":0.07,"why":0.08, "how":0.07, "when":0.05}
#maps={"who":"person","what":"obj","where":"loc","why":"reason","how":"method"}
corpus=[]
for whword in freqs:
    for i in range(freqs[whword]):
        corpus.append(whword)
numChildren=100
with open ("vocabSnapshot.csv","w") as outfile:
```

```
        outfile.write("ChidNumber,TimeStamp,WhWord,CorrectWeight,CorrectMapping,To
pWeight,TopMapping,IsCorrect\n")

with open("doubleProduction.csv","w") as doutfile:
        doutfile.write("ChildNumber,TimeStamp,meaningA,meaningB,formA,formB,position
1,position2,formACorrect,formBCorrect\n")
class Child:
        def __init__(self):
                #[form][meaning][mapping properties]
                self.matrix = np.zeros(shape=(6,6), dtype = (float, 6))
                #iterate overand add startWindow
                for i in range(len(self.matrix)):
                        for j in range(len(self.matrix[i])):
                                self.matrix[i][j][4]=startWindow
                self.age=0

        def guessMeaning(self,form):
                poss=self.matrix[form]
                #take all the weights for all six meanings, choose the highest, if there are
multiples, choose between them
                myweights=[]
                for i in range(len(poss)):
                        myweights.append(poss[i][3])
                maxweight=max(myweights)
                indices = [i for i, x in enumerate(myweights) if x == maxweight]
                meaningGuess = random.choice(indices)
                return meaningGuess

        def vocabSnapshot(self,interno,childNo):

        keyList=["ChildNumber","TimeStamp","WhWord","correctWeight","correctMapping
","topWeight","topMapping","isCorrect"]
                for whword in adultMaps:
                        whIndex=adultMaps[whword]
                        correctMap=adultMaps[whword]

        row={"ChildNumber":childNo,"TimeStamp":interno,"WhWord":whword,"correctWe
ight":0,"correctMapping":conceptMaps[correctMap],"topWeight":0,"topMapping":"","isCorr
ect":0}
                        topMap=self.guessMeaning(adultMaps[whword])
                        row["topMapping"]=conceptMaps[topMap]
                        row["topWeight"]=self.matrix[whIndex][topMap][3]
                        row["correctWeight"]=self.matrix[whIndex][correctMap][3]
                        row["isCorrect"]=int(row["topMapping"]==row["correctMapping"])
                        with open("vocabSnapshot.csv","a") as outfile:
                                for k in keyList:
                                        if k=="isCorrect":
                                                outfile.write(str(row[k])+"\n")
                                        else:
                                                outfile.write(str(row[k])+",")
```

```
    def doubleProduction(self,interNo,childNo,meaningA, meaningB):
        poss=[]
        for f in range(len(self.matrix)):
            poss.append(self.matrix[f][meaningA][3])
        maxweightA=max(poss)
        indices = [i for i, x in enumerate(poss) if x == maxweightA]
        formGuessA = random.choice(indices)
        successCtA=self.matrix[formGuessA][meaningA][1]
        poss=[]
        for f in range(len(self.matrix)):
            poss.append(self.matrix[f][meaningB][3])
        maxweightB=max(poss)
        indices = [i for i, x in enumerate(poss) if x == maxweightB]
        formGuessB = random.choice(indices)
        successCtB=self.matrix[formGuessB][meaningB][1]
        if successCtB>successCtA:
            pos1=formGuessB
            pos2=formGuessA
        elif successCtB==successCtA:
            dice=random.random()
            if dice>=0.5:
                pos1=formGuessB
                pos2=formGuessA
            else:
                pos1=formGuessA
                pos2=formGuessB
        else:
            pos1=formGuessA
            pos2=formGuessB

    keyList=["ChildNumber","TimeStamp","meaningA","meaningB","formA","formB","
position1","position2","formACorrect","formBCorrect"]#,"orderCorrect"]

    row={"ChildNumber":str(childNo),"TimeStamp":str(interNo),"meaningA":conceptM
aps[meaningA],"meaningB":conceptMaps[meaningB],"formA":formMaps[formGuessA],"for
mB":formMaps[formGuessB],"position1":formMaps[pos1],"position2":formMaps[pos2],"for
mACorrect":str(formGuessA==meaningA),"formBCorrect":str(formGuessB==meaningB)}
        with open("doubleProduction.csv","a") as doutfile:
            for k in keyList:
                if k=="formBCorrect":
                    doutfile.write(row[k]+"\n")
                else:
                    doutfile.write(row[k]+",")


    def updateMe(self,outcome,form,guessedMeaning):
        self.age+=1
```

```
                # list is [interactions, successes, lastinter, weight,currentWindow]
                #update interactions; no matter what, this was one
                self.matrix[form][guessedMeaning][0]+=1
                #update successes, this depends on outcome (0 is failure, 1 is success)
                self.matrix[form][guessedMeaning][1]+=outcome
                #update3 time of last interaction; it's current age
                self.matrix[form][guessedMeaning][2]=self.age
                #update weight; it's successes/interactions

        self.matrix[form][guessedMeaning][3]=self.matrix[form][guessedMeaning][1]/self.m
atrix[form][guessedMeaning][0]
                if useMemory:
                        if outcome:
                                self.matrix[form][guessedMeaning][2]=self.age
                                #print("i jacked my memory!","Form: ",formMaps[form],";
Meaning: ",conceptMaps[guessedMeaning])
                                self.matrix[form][guessedMeaning][4]+=1



totalChildren=100
for c in range(totalChildren):
        interCount=0
        corpusPosition=0
        myChild=Child()
        for j in range(totalInters):
                if j%100==0:
                        myChild.vocabSnapshot(j,c)
                        for x in range(0,5):
                                for y in range(x,5):
                                        if x==y:
                                                pass
                                        else:
                                                myChild.doubleProduction(j,c,x,y)#what who
                if j%len(corpus)==0:
                        random.shuffle(corpus)
                        corpusPosition=0
                thisForm=corpus[corpusPosition]
                childGuess=myChild.guessMeaning(adultMaps[thisForm])
                if childGuess==adultMaps[thisForm]:
                        #reinforce
                        myChild.updateMe(1,adultMaps[thisForm],childGuess)
                else:
                        #failure
                        myChild.updateMe(0,adultMaps[thisForm],childGuess)
                corpusPosition+=1
```