

University of Iceland

School of Humanities

English

First Language Acquisition

Are We Born with a Specific Ability to Acquire Language?

B.A. Thesis

Werner Johnsløv

Kt.: 070289-5049

Supervisor: Þórhallur Eypórsson

May 2018

Abstract

The purpose of this essay is to evaluate and assess the established theories behind innateness in first language acquisition. It sets out to explore the basic theories behind Noam Chomsky's Innateness Hypothesis. This includes an overview of the Language Acquisition Device and its relationship to Universal Grammar. The argument for the theory of Poverty of Stimulus, and an overview of the arguments for the likelihood of an existence of a critical period of language acquisition. The paper primarily focuses on how these established theories support Chomsky's views that language is an innate structure that children are born with. Furthermore, the essay sets out to evaluate a more recent book by Noam Chomsky and Robert Berwick on the history of language and its initial emergence, as well as its relationship with how human speech has changed over time. Then finally, I give an overview of Charles Yang's work on computational research based on the formation of productive rules on grammar within first language acquisition. The initial premise of the essay is to evaluate a selection of theories proposed by established linguists that there is an innate structure in human beings that determines how we acquire language the way that we do. The majority of research on this topic seems to agree with nativist theories, although the actual evidence to support these claims is to a certain degree still inadequate. The exception will be the findings by Charles Yang; where others have failed, he has managed to argue persuasively for a computational formula that gives an adequate explanation of how grammar rules are acquired. However, the findings are not indisputable and albeit there is a likelihood of the existence of innate language properties in human beings, the developments at this stage are still in need of much further research.

Table of Contents

1. Introduction	<u>1</u>
2. Noam Chomsky	<u>3</u>
2.1 The Innateness Hypothesis	<u>3</u>
2.1.1 Language Acquisition Device	<u>4</u>
2.1.2 Universal Grammar	<u>4</u>
2.1.3 Poverty of Stimulus	<u>7</u>
2.1.4 Critical Period	<u>10</u>
3. Robert Berwick and Noam Chomsky	<u>12</u>
3.1 Thought Processes and Merge	<u>12</u>
4. Charles Yang	<u>15</u>
4.1 The Tolerance Principle	<u>15</u>
5. Conclusion	<u>19</u>
References	<u>21</u>

1. Introduction

Human beings are unique in the way we communicate compared to other living beings on earth. We are the only creatures who have the ability to express themselves vocally without limitations due to the way our throats are comprised. By pushing air through our larynx, causing our vocal chords to vibrate, we are able to produce sounds and utterances that we use to create words and sentences. In the Cambridge Dictionary, language is defined as “a system of communication consisting of sounds, words, and grammar, or the system of communication used by people in a particular country or type of work” (Cambridge Dictionary, n.d). The way we learn to speak is however a different matter. As adults, learning a new language is for most people a challenging undertaking that takes up a vast amount of dedication and time. But when we are born, it is clear that the way we acquire our primary, or first language comes with ease. Within a relatively short time frame, human beings are able to learn complex language systems including intricate grammar rules without the need to study at all. The question is then, how do we learn to speak? The subject has been a point at issue for decades, and perhaps especially after the linguist Noam Chomsky attained recognition after the theory behind the Innateness Hypothesis was published.

The Innateness Hypothesis as coined by Noam Chomsky is believed to be one of the several arguments for linguistic nativism in humans (Rey, 2014). The idea is that human beings are endowed with a special inborn talent for learning language without any explicit guidance or learning. The much-discussed concept of there being a Language Acquisition Device (LAD) and a Universal Grammar (UG) is argued to be two of the crucial parts behind the justification for the Innateness Hypothesis (Penke & Rosenbach, 2007, p. 28). It is the fact that we are born with a set of rules that allow us to automatically acquire the required knowledge to tackle the notion of language and grammar when learning our first language, regardless of which language we acquire. Secondly, one of the most central arguments that is held by generative grammarians is that of Poverty of Stimulus (PoS) (Otero, 1994, p. 377). In essence it states that the amount of input that a child receives when acquiring language is so little that it would be inconceivable that it would be enough to explain the amount of output that a child communicates. Finally, there is the idea that we have a critical period of language acquisition. It is believed that human beings are only able to fully achieve complete native-like capabilities in a language if we are to begin learning within a certain period

in our youth, as observed by studying feral and deaf children. The work on first language acquisition by Noam Chomsky is well known within the field of linguistics and his views and theories are perhaps the most widely discussed on the subject throughout time. His work is also regarded as the most accepted and established pertaining to how we acquire language from birth (Groot & Kroll, 2010, p. 19). This is something which has been researched for a very long time, and the Innateness Hypothesis is by no means an example of modern exploration within the subject of language acquisition. The majority of research on nativism in linguistics come from the theories that were published by Chomsky, although nativism originally has its roots in philosophy and date back thousands of years (Dobrovolsky & O'Grady, 2007, p. 494).

Next to the established theories of Noam Chomsky's work on the Innateness Hypothesis, there are others who are progressing within this field of research. Computational linguist Robert Berwick has been collaborating with Noam Chomsky in order to find more contemporary arguments as to how we learn language. In one of their more recent works they attempt to explain that language can be seen as a structure of thought and not merely a system of communication. The book also illustrates a theory of how language came to be a tool for communication by way of natural selection. They also contend that the building blocks of language are based on the concept of Merge, a theory that illustrates how words come together in order to create sentences. Finally, there is Charles Yang, a linguist who has been making strides in his research within the field of language acquisition. In his latest book, he has been exploring the concept of innateness within grammar and his work is largely based on computational linguistics. He has formulated a theory that seeks to explain the way in which we learn our grammar. His work is based on the likelihood of children using productive rules when acquiring language and is thereby attempting to explain the inner workings of how we learn grammar. He is quickly becoming one of the names to be considered for prospective investigation in relation to language acquisition as well as cognitive science.

The much-discussed theories behind the Innateness Hypothesis by Noam Chomsky are still to this day behind the leading hypotheses concerning language acquisition. Linguists such as Robert Berwick and Charles Yang are with that also followers of Chomsky, and their research is largely based on his previous theories, albeit with their own outlooks. With that, the likelihood of there being a rational explanation towards language acquisition is indeed very high. That language can be

said to be innate within human beings is potentially something that we will see to be probable, either by way of the established theories within the field, or by more contemporary methods in research to come.

2. Noam Chomsky

Noam Chomsky and his theories on language acquisition is perhaps the most recognized works on the subject of how we learn to speak. Chomsky is the man who looked away from the established theories of behaviorism by Burrhus Skinner, based on operant conditioning. Skinner argued that children are blank slates that go through a trial and error in their ways of learning. Meaning that children acquire language by way of operant conditioning, thereby they learn with re-enforcement and shaping (Skinner, 1957). Chomsky argued against his theories (Chomsky, 1959) and he is the one who managed to take the field of linguistics into what it is today. Chomsky has published a vast array of works on linguistics and within several different genres throughout his career (Noam Chomsky Book List. n.d.). But of all the work that Chomsky has done within linguistics, there is one topic that has caused more deliberation than others, namely his theory of language acquisition and his Innateness Hypothesis.

2.1 The Innateness Hypothesis

The Innateness Hypothesis is in essence the theory of linguistic nativism in human beings. The theory is based on the idea that human beings are endowed with some genetic blueprint for learning language. As famously stated by Chomsky, “the speed and precision of vocabulary acquisition leaves no real alternative to the conclusion that the child somehow has the concepts available before experience with language and is basically learning labels for concepts that are already part of his or her conceptual apparatus” (2001, p. 28). It is easy to comprehend why such a statement would be made given the actual rate at which children do acquire language. It seems inconceivable for an adult to experience the same learning curve that children have when growing up and it is truly a remarkable achievement that we are able to learn to speak in the manner that we do. Chomsky suggests that we have something within us when born that aids us in the learning process, a Language Acquisition Device that makes us automatically acquire the rules of language. Collectively, his theories on innateness within language

learning have been the most recognized hypotheses of language acquisition throughout time and it is still to this day the most established theories on the subject.

2.1.1 Language Acquisition Device

One of the major arguments for linguistic nativism in human beings is the theory that we have a faculty called the Language Acquisition Device, or LAD (Chomsky, 1969). In essence it is the idea that there is a capacity that we are endowed with when born that help explain how children acquire language in the manner that they do. Given that children learn language at the speed and accuracy that they appear to, there is no proven account for how they are able to do this, but the LAD gives a plausible answer to this problem. What Chomsky argues is that LAD explains how children are able to learn any language, describing it as “an innate property of an organism or device capable of language acquisition” (Chomsky, 1969, p 37). There is no part of the brain having been isolated and identified with the LAD, but Chomsky does contend that when children start hearing their primary language, the LAD will likely be activated. “Certain kinds of data and experience may be required in order to set the language-acquisition device into operation” (Chomsky, 1969, p. 33), primarily based on the input they will hear in their surroundings. Another crucial argument for the existence of a LAD is the fact that children will acquire language, irrelevant of which language that may be. There are vast amounts of languages in the world and where some of them have very simple grammar and pronunciation, some have very complex grammar and pronunciation. There has been no specific part of the human brain that have been isolated and identified as the LAD, but given the rate of acquisition, it is plausible that such a device exists within the human brain. Nonetheless, the rate at which children learn to speak still follow the same path, they will acquire their linguistic capability concurrently regardless of which language they learn (Clark, 2016). According to the hypothesis, what children do is learn the basic rules of a language by way of listening and the LAD, then they transform these rules into basic grammar by the help of Universal Grammar.

2.1.2 Universal Grammar

Part of The Innateness Hypothesis is the theory that humans are equipped with what Chomsky calls Universal Grammar. The idea is that given the vast amount languages that exist, and have existed, they are all languages that have some form of grammar. Children seem to acquire the rules of grammar effortlessly and within the same time

frame regardless of which language they are learning, so it would be conceivable that there is an innate ability in us from birth that explains how we learn grammar. As stated by Chomsky and Ronat, “universal grammar is not a grammar, but rather a theory of grammars, a kind of metatheory or schematism for grammar” (1983, p. 183). When children learn to speak, they do so without the knowledge of why they speak as they do. When having learnt our first, or primary language, we speak the language fluently, and with that the grammar rules that we follow and obey are to a certain extent always correct. The thing that is puzzling is then that we are not necessarily able to explain how or why. Even though we speak a language correctly, we are not able to explain the various grammar rules that we use, simply because we never actually learned them as such. Universal Grammar consists of an inhibition that unwittingly we are speaking in the correct way without truly being aware that we are. So it becomes clear that if a native English speaker whose accent is Received Pronunciation (RP) were to hear the phrase *I has read a book*, he or she will automatically know that the correct way to say this sentence would be *I have read a book*. A possible explanation could be simply that it does not sound right, therefore he or she would know that the correct choice of word would be *have*.

The idea that native speakers of a language will have innately acquired the rules of language also seems to be a valid theory considering these two sentences Chomsky fabricated:

(1) Colorless green ideas sleep furiously (2002, p. 15)

(2) Furiously sleep ideas green colorless (2002, p.15)

Both of these sentences are equally absurd and they make no sense to any speaker of English. But whereas they are both nonsensical, only one of them represents a grammatically sound and plausible sentence. While the second sentence fails to represent anything even remotely sound, the first one is an example of a perfectly structured sentence of English. It may be an example of nonsense that has no apparent meaning at all, but it is still clear that any speaker of English will see it as a sentence that will work within the paradigm of English syntax. These sentences were included in the first edition of *Syntactic Structures* in 1957, when Burrhus Skinner’s behaviorist theories were still regarded as the leading cause behind first language acquisition. Chomsky’s “colorless green ideas” (p.15) helped to disprove Skinner’s theory that the human mind was a blank slate upon birth (Baker, 2008). Evidently, as seen when being presented with nonsensical phrases, it is still possible to create absurd but yet

structurally sound sentences. This is something that even children are able to deduce, and the reasoning behind this is that they have a basic understanding of how to grammatically construct sentences. Thereby, it is likely that children are endowed with some form of an innate structure that unconsciously alerts them if a phrase is structurally valid. When even young children can observe that only one of the two sentences are grammatically sound, it becomes apparent that there is more to acquisition of grammar than what was formerly assumed.

Furthermore, to learn languages the way we do, is inherently human, no other creature on earth is able to do so. There are examples of typological claims that argue that all languages do have certain similarities. According to certain linguists, for example Andrew Carnie, he states that all human beings must have some innate properties as “all speakers of human languages share the same basic innate materials for building their language’s grammar” (2002, p.18). Beyond that, regardless of culture, language, or even intelligence, human beings acquire language. Additionally, “the structure of language used by subnormal children resembles that of normal children at the corresponding stage of development” (Martin, 2013, p. 162). Therefore, it is inconsequential if a person is less intelligent, the rate at which he or she will acquire language will still follow the same trajectory as everyone else, and as previously stated, the same applies to grammaticality judgments.

Universal Grammar is one of those aspects of language acquisition that have persevered over time as seen by Chomsky’s continued inquiry (Chomsky, 2006), but future research will still be needed in order to confirm the various theories. When studying acquisition today it can absolutely seem viable that there is such a thing as a Universal Grammar by the way we observe children and their learning curve. It could very well seem as if they learn grammar in a way that would be inconceivable if there is not something to the likeness of Universal Grammar. Furthermore, children acquire grammar at an incomprehensible rate, and the concept of Universal Grammar is one of the major arguments for innateness within language. But it is still difficult to pinpoint exactly what it would entail. Research of that nature we have only just begun to comprehend, and the future of linguistic research should provide further developments.

2.1.3 Poverty of Stimulus

Besides the concept of Universal Grammar, there is the hypothesis of Poverty of Stimulus, one of the strongest cases that help support the theory of linguistic nativism in human beings (Otero, 1994, p. 377). It is in essence also an argument that gives evidence to the theory of Universal Grammar. The hypothesis is based on the idea that when learning your primary language there is not enough input to explain the amount of output a child will produce. The detailed knowledge a child will have of their first language far exceeds the amount of information that they will have heard by their surroundings. Furthermore, there is no explicit teaching present when a child learns his or her first language, the utterances conveyed from a very early age demonstrate that there could be something inborn that explains the linguistic comprehension in children. This idea was originally coined by Chomsky in 1980 and it is one of the strongest advocates for his Innateness Hypothesis as well as Universal Grammar.

The concept behind Poverty of Stimulus is one that has been researched abundantly and there are several examples and experiments that speak for the existence of Chomsky's Poverty of Stimulus. As Morris Halle demonstrates in an article, consider the following words:

(1) ptak thole hlad plast sram mgla vIas flitch dnom rtut (Halle, 2003, p. 95)

These are all words from different languages around the world, and if you were to ask a person whose native language is English, to choose which ones he or she would believe to be words that actually exist in English the answers would likely be: *thole*, *plast*, and *flitch*. Even if you have never heard those words before, which would be unlikely anyway as they are some of the least used words of English, you would still recognize that they could be words in English. This is not knowledge that you would have learned especially when acquiring your first language. "English-speaking parents do not normally draw their children's attention to the fact that consonant sequences [in] English words are subject to certain restrictions" (Halle, 2003, p. 95). But we instantly know that these words could be part of the English language nonetheless, thereby it is knowledge that we have acquired without any apparent teaching.

Another noteworthy example is pertaining to the idea that children do not accept negative evidence. Given the very often quoted dialogue between a mother and child suggests the idea that children accept correction and learn by way of imitation, is false:

CHILD: Nobody don't like me

MOTHER: No, say "Nobody likes me"

CHILD: Nobody don't like me.

MOTHER: No, say, 'nobody likes me'.

CHILD: Nobody don't like me.

MOTHER: No, now listen carefully; say "Nobody likes me"

CHILD: Oh, nobody don't like me

(Loritz, 2002, p.11)

As the mother attempts to correct the child it becomes clear that her correction is not proving to be worthwhile. It is a concise argument for the existence of nativism in children. As the child will not learn by correction, how will he or she ever realise that their discourse is ungrammatical. Furthermore, children have no way of deducing what certain expressions or sentences will *not* mean as they are commonly never corrected, and even if they did, it clearly will not prove to be valuable. Besides the aspect of negative evidence there are issues regarding the general input that children receive. "The input is impoverished because children do not hear only well-formed utterances, but also fragments of utterances" (Foster-Cohen, 2009, p. 97). These utterances will include hesitance in speech, incomplete sentences and slips of the tongue. Therefore, it would stand to reason that children would not be able to acquire a sufficient amount of knowledge of a new language without there being something inherently inborn. If a child is expected to learn his or her mother tongue only by way of listening to their surroundings, it would be inconceivable that they would learn to speak at all. Foster-Cohen also writes in her book (2009) a comparative example of what it would be like to learn your first language based on input alone. It is "like trying to deduce the rules of a new game in which players are cheating, but where this is not marked in any way" (p. 97). It would be an extreme accomplishment, and it appears to be so, as every speaking person on earth will at one point have been subject to this in their primary language acquisition. This is perhaps the strongest advocate for innateness within language in humans, the idea that we are to learn a language without receiving an adequate amount of input is seemingly beyond belief. Especially so when considering that children are apparently not susceptible to correction of their discourse.

When reviewing the Poverty of Stimulus argument, there are certainly theories that go against it. The strongest case that goes against the hypothesis in itself is very

likely to be the article written by Geoffrey Pullum and Barbara Scholz in 2002. Their paper is an extensive analysis of the arguments behind the Poverty of Stimulus, or the APS (Argument for Poverty of Stimulus). In essence what they are arguing is not that there is no Poverty of Stimulus, or linguistic nativism as such, but rather the actual evidence or argument that such a thing does exist is put into question. “We are concerned not with whether empiricist claims about language learning are true, but we are concerned with whether linguists have established their falsity” (Pullum & Scholz, 2002, p. 10). What they are claiming is that while there is a vast amount that have been written on the subject supporting the theory of Poverty of Stimulus, there is not much actual evidence to back it up. The one thing that the article makes the reader understand clearly, is what the conclusion to the argument of Poverty of the Stimulus is presumed to be. It is to show that when born, infants are inherently endowed with an intricate linguistic system that aid in the acquisition of language, or the concept of linguistic nativism. But what is not apparent at all is the interpretation or premise behind how we are to reach said conclusion (Pullum & Scholz, 2002). What is claimed is that there is no actual evidence that truly supports the hypothesis of PoS. There are merely theories that are linked to the idea behind impoverished input in children. Thereby, the paper claims that “instead of clarifying the reasoning, each successive writer on this topic shakes together an idiosyncratic cocktail of claims about children’s learning of language, and concludes that nativism is thereby supported” (Pullum & Scholz, 2002, p. 12). The arguments within the paper are agreeable but what it does not do however, is to deny that there is such a thing as Poverty of the Stimulus or linguistic nativism itself. It claims to actually be in favor of the hypothesis as it would make sense given the fact that children acquire language as fast and effortless as they do. But of all the evidence that have been presented that is supposedly supporting the cause, they claim that none of it is admissible. The hypothesis based on impoverished input in children is one that would likely be believable if it is true that children do not learn by way of listening to their surroundings. However, the claim is difficult to prove beyond a reasonable doubt and so what Pullum and Scholz suggest is further research to be carried out in order to come the conclusion that so many have prematurely taken for granted.

2.1.4 The Critical Period

When children learn their first language there have been many who support the idea that there is a set given time that children are able to acquire their mother tongue. This theory supports the hypothesis of linguistic nativism in that we are only able to become absolutely fluent in a language from a very young age, hence that children are gifted with the power of language. It is originally a theory that was created by Eric Lenneberg and as stated in his book, after the onset of puberty, “languages have to be taught and learned through a conscious and laboured effort” (1967, p. 167). According to the theory, if a language is not acquired before the onset of puberty, the likelihood of becoming well versed is very slight. Another argument for this is also the observation that the vast majority of those who learn a new language after puberty, although adequately so, will never truly become fluent in the language that they attempt to master. This is an issue that has caused much deliberation amongst scholars and more recently the term has been altered so it is believed to be both a critical and a sensitive period (Meredith, 2015). Although it was not Chomsky who originally had the idea, it is very often seen to be in affiliation with him and his theory on Universal Grammar and LAD as it is very supportive of his hypotheses. There are several arguments in favor of the Critical Period Hypothesis, a majority of which are based on situations where children would have to learn speech at later stages in life.

The most notable case that has been studied in regard to researching the possibility of a Critical Period is by far the case of Genie (Newmeyer, 1988). She was a girl that was unjustifiably assumed retarded by her father and consequentially she was chained to a toilet and “kept in a small room with virtually no opportunity to hear human speech from around age two to age thirteen” (Dobrovolsky & O’Grady, 2007, p. 495). Her cognitive functioning was characterized as good at the time of discovery but she was unable to communicate at the level of someone her age. She had signs of serious neglect and her attempts at conversation during her confinement had been penalized by her father who called for silence repeatedly (Curtis & Whitaker, 2014). She was thirteen years old when she was found, and therefore she was believed by psychologists and linguists to have passed the critical stage of language acquisition. She was subjected to extensive research and studying as she was believed to offer insight into the field of language acquisition. Below is a small collection of utterances she made at the age of approximately 20, recorded in August of 1977. As seen the short sentences have clear examples of word order errors:

Utterance:

Dark blue, light blue surprise square and rectangle
 Blue light and square I want a light blue square box
 Think Mama I am thinking about Mama
 Think about Mama love Genie
 No want dark surprise

(Newmeyer, 1988, p. 97)

It slowly became evident that the linguistic ability of Genie would never surpass a certain stage. “Even after more than eight years of linguistic exposure and attempted acquisition, Genie’s utterances remained largely agrammatic – they contained little and inconsistent use of inflectional morphology and other nonlexical grammatical markers” (Newmeyer, 1988, p. 97). The results from the testing of Genie have been disputed but the outcome is still believed that she never truly managed to acquire language to the extent of a native speaker. The complex grammar systems are seemingly not attainable later in life to a native proficiency, thereby it becomes clear that language is not learnable as such after a given time. Consequently, first language acquisition cannot be considered to be an example of learning at all. It is possible to argue that given her situation and extreme upbringing can be to blame, but her case is still very much a strong advocate for the existence of a critical period of language acquisition.

Another argument that supports the existence of a critical period is the rate at which deaf children will acquire sign language. This has been researched and the results are conclusive in that the younger the child will start to learn sign language, the more proficient he or she will become and within a shorter time frame. According to Johnson & Newport (1989), “since 90% of the congenitally deaf have hearing (speaking) parents, only a few deaf individuals are exposed to this language from birth” (p. 62). Thereby it is relatively easy to find subjects who are beginning to learn sign language at a later stage in life. Furthermore, the article examines the rate at which deaf children learn ASL (American Sign Language) and the results are clear, “there is a decline over age in the ability to acquire a first language” (Johnson & Newport, 1989, p. 62). With that there is clearly a correlation between age and language as seen with ASL. The case of the Critical Period is here most relevant, as the opportunities to study children who have not begun learning ASL from birth are more widespread than those that can be

described as feral. Evidently, there is no disputing that a child who begins learning as early as possible will become more conversant in sign language compared to one who would start at a later age.

Within the field of second language acquisition (SLA), there is still deliberation on the matter of the Critical Period. But within first language acquisition, there are clear correlations between those who do not learn language from birth and those who start at a later stage. The evidence seems to point to the theory that there is indeed a given period of time where children are more adept in learning language. Therefore, the possibility of language being innate within human beings is strengthened. If those who did not attain language as children are unable to become fully versed in their language, there is arguably a mechanism or innate ability within us that aid us in becoming proficient at language.

3.0 Robert Berwick and Noam Chomsky

Robert Berwick is a computational linguist who recently collaborated with Noam Chomsky on the book *Why Only Us* (2016). The book is essentially a collection of essays where they attempt to give reason for the evolution of language. The book explains how language and Universal Grammar first emerged, as well as an historical account of previous research into how language will have changed over time.

3.1 Thought Processes and Merge

The basic idea that is put forth by Berwick and Chomsky is namely that language as we see it today was not always meant to be spoken as such. Instead they argue that the premise for language is that it is primarily optimized for thought processes and that it was a system of thought long before homo sapiens began to speak (Berwick & Chomsky, 2016, p. 75). They claim that language will have emerged due to externalization, explicitly so in the same manner as songbirds sing. Then later it will have been developed and in essence built upon the concept of Merge, “the Basic Property of human language” (Berwick & Chomsky, 2016, p. 11). It is being described as the building blocks of our language, claiming that the instance of Merge is what allows us to create an infinite amount of sentences, and hence language itself.

When considering what language is, it is very easy to assume that it is and always has been a medium used for communication between individuals. What Berwick

and Chomsky claim is that language was not necessarily intended for this very purpose. They suggest that language may well have been intended for personal thought processes. Thereby, “language as [an] internal mental tool, does not assume, explicitly or implicitly, that the primary function of language is for external communication” (Berwick & Chomsky, 2016, p. 81). If it is true that language was not intended for communicative purposes, it would mean that it was predestined for human beings as individuals, which is curious given that we are social creatures. However, that human beings are the only apes that are endowed with speech does not mean that we are the only ones gifted with thought. But it is a possibility that language was primarily used for thought and that when it became a tool for communication it did so at a very basic level of speech. Berwick and Chomsky contend that language can be seen to be a representation of Darwin’s Caruso Theory from Darwin’s *Decent of Man* (1871). Accordingly, men would attempt to attract females by way of singing in the very beginning. Men would then be selected according to how well they could sing and consequently those who could sing would procreate and by natural selection, the vocal chords would be perfected over time. The concept of language for communication would then with time evolve and eventually be seen as any other organ of the human body. “Better vocal competence went hand in hand with a general increase in brain size that led, in turn, to language” (Berwick & Chomsky, 2016, p. 3). The arguments for their theory is perhaps far-fetched yet not illogical, the likelihood that humans had language the very minute we originated is unlikely. Thereby, according to Berwick and Chomsky, the building blocks from where language and words originate, are based on thought processes and the motivational state that comes from sexual yearning.

The method for human beings to produce discourse, to combine words into sentences or phrases, is according to Berwick and Chomsky constructed from what is called Merge. “This operation takes any two syntactic elements and combines them into a new, larger hierarchically structured expression” (Berwick & Chomsky, 2016, p. 21). Specifically, this is the operation of taking two words and combining them together in order to create strings of words, or sentences. It must also be mentioned that Merge is not a new theory as it was originally proposed by Chomsky in the early 1990s (see Chomsky, 1993). However, when we learn our primary language, Merge is something that happens automatically, without any explicit teaching as such and it is something which is unique for human beings. When children learn to speak their first words it will be in singular form, often referred to as the one-word stage, a period between 12 and

18 months of age (Dobrovolsky & O’Grady, 2007, p. 476). A few months after the one-word stage, children enter the two-word stage, a period where children begin to produce mini-sentences. As seen by the table below, the utterances do make use of correct word order albeit being incomplete sentences in itself:

<u>Utterance</u>	<u>Intended Meaning</u>
<i>Baby chair</i>	<i>The baby is sitting on the chair</i>
<i>Doggie bark</i>	<i>The dog is barking</i>
<i>Ken Water</i>	<i>Ken is drinking water</i>
<i>Hit doggie</i>	<i>I hit the doggie</i>

(Dobrovolsky & O’Grady, 2007, p. 477)

It is apparent that the way that children acquire language employs a system of Merge from a very early stage in their development and it should therefore be considered the building blocks of language. Human beings are the only animals that have Merge, something that became evident when having attempted to teach chimpanzees ASL. Research found that chimpanzees are able to acquire a certain level of language competence with regards to certain words, or signs, but they are incapable of formulating a structured expression from two different words (Berwick & Chomsky, 2016, p. 147). Chimpanzees have managed to associate signs with the corresponding item as seen by the ape Nim Chimsky. He “stored a list of explicit, mind-independent associations between objects in the external world and the ASL signs for them” (Berwick & Chomsky, 2016, p. 147). With that it is clear that chimpanzees, as one of the animals closest to humans in cognitive ability (Wildman, Uddin, Liu, Grossman, & Goodman, 2003), are not able to comprehend language in the same way as humans do. They would never be able to apply Merge in their use of language, and by that they would not see their use of sign language, as language at all.

The book by Berwick and Chomsky contends to prove that the way language was created from the very beginning was initially by way of thought and not for communication. The theory that language was developed due to natural selection is indeed a conceivable argument, but one that is challenging to truly validate and because of this, the idea has sparked some debate. According to a critical review by Evans, it is

argued that to claim language was never actually developed for communication is a very bold statement, something that is rather difficult to accept (2016). That being mentioned, it is not entirely inconceivable but it is still something which is not altogether self-evident. Chomsky's theory of Merge is also strongly campaigned for, as being something that only humans have. The concept of stringing words together to structured phrases is something that can be seen from a very early age in children. Animals in nature have expressions they use for communication, but only users of human language are able to converse without limitation. Although Merge is not a new concept, it is still a major aspect discussed in this book; and for good reason as it does explain an evolutionary development for language, and it is also an advocate for innateness in human language.

4.0 Charles Yang

Charles Yang is a computational linguist who has worked alongside both Noam Chomsky and Robert Berwick and he is one of the more notable and promising linguists of our time. He has in his most recent book (Yang, 2016) been working on understanding the concept of productive rules within grammar systems. His approach within computational linguistics have yielded encouraging results and he is rapidly becoming a notable scholar within the field of language acquisition.

4.1 The Tolerance Principle

Charles Yang has been working on a method to explain how children acquire grammatical rules when they are learning their first language. Up until this point the basic premise for the argument behind Universal Grammar has primarily been based on the fact that children acquire language and grammar at the incredible rate that they do. But the reasoning behind why and how they acquire their knowledge has not been accurately resolved. It is very easy to assume that children will learn their grammar because of an innate structure in the brain, but there has been little to no actual evidence that support the claim. Collectively, the theories of LAD, UG, PoS, and the Critical Period Hypothesis make up the arguments for innateness in language acquisition. But they are still only theories, without any concrete evidence that proves without a doubt that they are true. However, given the speed and manner of acquisition, there has to be something there, something that would explain how the rules of language is acquired

by children. Yang has been working on a computational solution to this problem, attempting to solve and explain exactly how children would be able to acquire their grammar in the manner that they do. In essence, Yang has written his book about an equation, or about the problems that the equation is meant to solve (2016, p. 1). He attempts to explain how children manage to learn productive rules and by way of giving examples alongside presenting previous work on the subject as well as explaining his solution, or equation. He concludes that there is indeed a way to see exactly how children are able to learn productive rules and generalizing patterns in what he calls the Tolerance Principle.

When children learn to speak, they need to form certain rules as they acquire their first language. Rules of this sort include to add *-s* to a noun in order to make it plural, or by adding *-ed* to a verb in order to make it past tense. This has been studied a great deal, for example by Ivanovska in her study on German nouns and their plural forms (2015), and it is one of the issues the Tolerance Principle seeks to explain. Then there is productivity in child language acquisition: “The acquisition of morphological productivity can be summarized very succinctly: children draw a sharp, essentially categorical, distinction between productive and unproductive processes”. Accordingly, “the simplest assessment of productivity is the celebrated Wug test” (Yang, 2016, p. 31). The wug test was created by psycholinguist Jean Berko and it has been referenced and copied extensively since it was published in 1958 (Yang, 2016). The basic premise for the test is that it examines internalized morphological rules in children. “We know that if the subject can supply the correct plural ending, for instance, to a noun we have made up, he has internalized a working system of the plural allomorphs in English, and is able to generalize to new cases and select the right form” (Berko, 1958, p. 150). Thereby the wug test is as follows:

This is a WUG
 Now there is another one
 There are two of them
 There are two (WUGS)

(Berko, 1958, p. 154)

Yang argues that the test itself was not very reliable, as seen when the test subjects were presented with the verbs *bing* and *gling*. “Only one of the eighty-six children in Berko’s study supplied the analogical form *bang* and *glang*” (Yang, 2016, p. 32). However, due

to the scope of the testing, the results still provided some evidence of how children learn and deal with morphological rules of language. It is also a test that seeks to explain exactly why children manage to apply grammatical rules to words that they not necessarily have heard before. The way in which Yang has created his theory is loosely based on the same principles as the wug test. Although the test is questionable, it is still an example of how innateness can be understood and it is justifiable that Yang would include it in his book.

The wug test itself is one example in the book of how Yang managed to come to a feasible conclusion and in fact an equation that explains how children learn their grammar. The equation, which he calls the Tolerance Principle is an extension to the wug test that is made into a mathematical formula of how children form productive rules when learning grammar. Thereby the Tolerance Principle is as follows:

If R is a productive rule applicable to N candidates, then the following relation holds between N and e , the number of exceptions that could but do not follow R :

$$e \leq \theta_N \text{ where } \theta_N := \frac{N}{\ln N}$$

(Yang, 2016, p. 9)

In essence, the Tolerance Principle is a learning model that evaluates the environment for generalizations when learning a first language. It will consider that when general rules are formed it will be computationally more efficient than storing them individually for each lexical item (Yang, 2016). Basically, what this means is that for example when a learner encounters the rule to add *-ed* in order to change a verb into past tense form, he or she will encounter a large number of lexical items that follow this rule. The learner will also come across a fairly large number of irregular verbs that do not follow this rule. What a child can do in this case is to store each lexical item individually as he or she will hear them in order to learn the grammatical regulation one by one. Alternatively, the child can form a productive rule where only the exceptions to the rule will be stored individually based on frequency of use. Thereby, when a learner hears a new word, he or she would be able to consider which words would have been previously learned to be irregular, and would thereby conclude whether or not to use the regular ending of *-ed*. The Tolerance Principle would then compute in order to see whether or

not it is efficiently viable to form a productive rule or not so as to find out how to best to learn a grammatical ending for each word. “At the present time, it is not clear how the Tolerance Principle is executed as a cognitive mechanism of learning: surely children doesn’t use calculators” (Yang, 2016, p. 64). But children who acquire language will have two ways of learning the words and their respective irregularities, and they would choose whichever one that they would learn faster.

In a survey done by Charles Yang, Kathryn Schuler, and Elissa Newport, they test the Tolerance Principle in order to see if productive rules will be formed when test subjects are presented with an artificial language. The participants were presented with 9 random nonsensical nouns, “children exposed to 5 regular forms and 4 exceptions generalized, applying the regular form to 100% of novel test words. Children exposed to 3 regular forms and 6 exceptions did not extend the rule” (Schuler, Yang, & Newport 2016, p. 2321). The survey shows that even with nonsensical words, productive rules will be formed when presented to the test subjects. When there are more words that are in regular form than irregular ones, learners will generalize. Therefore, when there are more words that have regular endings, it proves that it is computationally efficient to form a productive rule, and that is what children will do. The results from the survey gives a strong indication that the way children comprehend grammar as quickly as they do, will have to be due to generalization. The findings are with that also a signal that human beings are innately bestowed with language.

What Yang has managed to do is to expand on the well-known wug test by constructing a mathematical formula that in essence gives a plausible explanation to how children learn grammar. As Universal Grammar is something that up until this point has been seen as an oracle apparatus with little to no explanation as to what it actually entails, the Tolerance Principle is a likely justification. The mathematical formula that is the Tolerance Principle provides insight into how Universal Grammar possibly works and it is something that can help explain if language is indeed innate in human beings. The equation gives an account of how children acquire their basic rules of language and thereby it can be viewed as a likely advocate for humans being born with an inborn capability for language.

5.0 Conclusion

The widely discussed theories on first language acquisition and the likelihood of the existence of innate capabilities when learning language is a matter that is yet to be proven, but it is something that eventually can be seen to be a probable explanation to how we learn language. The established theories by Noam Chomsky will presumably see further developments and his views are still to this day on the forefront of understanding the concept of language acquisition. Chomsky has laid out extensive groundwork within the field as seen by his work on the Innateness Hypothesis over the course of his long career. The Language Acquisition Device, a faculty that is believed to be an inborn structure in our brains that allows us to acquire language from a very early age is one of the arguments for innateness in humans. To build on his theory of LAD, there is Chomsky's Universal Grammar, perhaps his most widely discussed theory of acquisition as it helps to explain how children are able to comprehend grammar in a way that requires effectively no didactic involvement at all. Furthermore, there is the theory of Poverty of Stimulus, a concept that states that the amount of information a child will receive is less than the amount of output he or she will produce. Which if seen to be true, the feat of learning a language would be a truly extraordinarily achievement given that it is very difficult to grasp how it could be done without effective input. Finally, there is the concept of there being a critical period of language acquisition, as stated, the time of learning will need to be initiated before a certain age in order to reach native-like proficiency in any given language. Collectively these theories argue for the existence of there being an inborn structure in human beings that explains how we are able to learn language without any explicit teaching as such.

Robert Berwick and Noam Chomsky have claimed that at its very basic level, language was not intended for communication but rather for individual thought processes. Then by way of natural selection, language will have emerged and the notion of language as communication will have evolved as any other organ in the body. They also contend that the building blocks of language can be seen to be built upon the theory of Merge, a concept that explains how we are able to create an infinite amount of sentences with a finite amount of resources.

Charles Yang and his work on the formation of productive rules sets out to explain how children are able to learn intricate grammar systems by way of computational modeling. The Tolerance Principle is something that helps to explain the

existence of Universal Grammar and it clarifies how children learn their first language. Yang's formula is unique in that it is the first example of a probable explanation to how the mind works when children acquire their basic grammar. His research is likely the most unveiling of its time and his findings are probably the most outstanding since the Innateness Hypothesis was first presented.

All of the theories presented here are in essence based on the established theories by Noam Chomsky. The leading theories behind the Innateness Hypothesis are still to this day seen as being created by Noam Chomsky and they are by no means examples of modern exploration within the field of language acquisition. But they have stood the test of time and further research is still to this day being executed in order to find evidence that further support his claims. Charles Yang is one of those scientists that seeks to explain and prove his theories, and he has done tremendous strides in his effort behind the Tolerance Principle. Although the basic premise behind the Innateness Hypothesis has seen adversity, it is still very much applicable in today's investigation of first language acquisition. But as stated by Pullum and Schulz on the matter of innately primed learning, until it "is investigated in more detail, linguists will remain ill-equipped to do more than fantasize and speculate on the matter" (2002, p. 47). Even though there have been great strides within research, the simple truth is that there is not at this point any clear evidence that explicitly illustrate how language acquisition occurs. Therefore, the Innateness Hypothesis is still only to be considered that which it is so accurately titled, namely a hypothesis.

References

- Baker, M. C. (2008). The Creative Aspect of Language Use and Nonbiological Nativism. *The Innate Mind*, 3, 233-253.
doi:10.1093/acprof:oso/9780195332834.003.0012
- Berko, J. (1958). The Childs Learning of English Morphology. *WORD*, 14(2-3), 150-177. doi:10.1080/00437956.1958.11659661
- Berwick, R. C., & Chomsky, N. (2017). *Why Only Us: Language and Evolution*. Cambridge, MA: MIT Press.
- Cambridge Dictionary. (n.d.). Language: definition of language in the Cambridge Dictionary. Retrieved 10 February 2018 from <https://dictionary.cambridge.org/dictionary/english/language>
- Carnie, A. (2013). *Syntax: A generative introduction*. Hoboken, NJ: Wiley-Blackwell.
- Chomsky, N., & Ronat, M. (1983). *Language and Responsibility*. Taipei, TW: Wen he.
- Chomsky, N. (1969). *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.
- Chomsky, N. (1980). *Rules and Representations*. Oxford, UK: Basil Blackwell.
- Chomsky, N. (2001). *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: The MIT Press.
- Chomsky, N. (2006). *Language and Mind*. Cambridge, MA: Cambridge University Press.
- Chomsky, N. 1993. A minimalist program for linguistic theory. *MIT Occasional Papers in Linguistics*, Cambridge, MA: The MIT Press
- Chomsky, N., & Lightfoot, D. W. (2002). *Syntactic Structures* (2nd ed.). Berlin, DE: De Gruyter.
- Chomsky, N. (1959). Verbal behavior. *Language*, 35(1), 26-28. doi:10.2307/411334
- Clark, E. V. (2016). *First language acquisition*. Cambridge, MA: Cambridge University Press.
- Curtiss, S., & Whitaker, H. A. (2014). *Genie: A Psycholinguistic Study of a Modern-Day Wild Child*. New York, NY: Academic Press.
- Darwin, Charles. 1871. *The Descent of Man, and Selection in Relation to Sex*. London, UK: John Murray.
- Dobrovolsky, M., & O'Grady, W. (2007). *Contemporary linguistics: An introduction*. Harlow, ND: Longman.

- Evans, V. (2016, February 24). Why Only Us: The language paradox [Review of the book *Why Only Us*, by Berwick, R. C., & Chomsky, N]. *New Scientist*. Retrieved March 1, 2018.
- Foster-Cohen, S. H. (2009). *Language Acquisition*. Basingstoke, UK: Palgrave MacMillan.
- Groot, A. M., & Kroll, J. F. (2010). *Tutorials in bilingualism psycholinguistic perspectives*. New York, NY: Psychology Press.
- Halle, M. (2003). *From memory to speech and back: Papers on phonetics and phonology 1954-2002*. Berlin, DE: Mouton de Gruyter.
- Ivanovska, B. (2015). On the plural forms of nouns in German compared to Macedonian language. *Procedia - Social and Behavioral Sciences*, 180, 66-71. doi:10.1016/j.sbspro.2015.02.087
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60-99. doi:10.1016/0010-0285(89)90003-0
- Loritz, D. (2002). *How the brain evolved language*. New York, NY: Oxford University Press.
- Meredith, R. (2015). Sensitive and critical periods during neurotypical and aberrant neurodevelopment: A framework for neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews*, 50, 180-188. doi:10.1016/j.neubiorev.2014.12.001
- Martin, J. A. (2013). *Voice, Speech, and Language in the Child: Development and Disorder*. Vienna, AT: Springer.
- Newmeyer, F. J. (Ed.). (1988). *Linguistics: The Cambridge Survey*. Cambridge, MA: Cambridge University Press.
- Noam Chomsky Book List. (n.d.). Retrieved February 23, 2018, from <http://www.chomskylis.com/noam-chomsky-book-list.php>
- Otero, C. P. (Ed.). (1994). *Noam Chomsky. critical assessments*. London, UK: Routledge.
- Penke, M., & Rosenbach, A. (2007). *What counts as evidence in linguistics: The case of innateness*. Amsterdam, NL: Benjamins.
- Pullum, G. K., & Scholz, B. C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, 18(1-2). doi:10.1515/tlir.19.1-2.9

- Rey, G. (2014). Innate and Learned: Carey, Mad Dog Nativism, and the Poverty of Stimuli and Analogies (Yet Again). *Mind & Language*, 29(2), 109-132. doi:10.1111/mila.12044
- Schuler, K, Yang, C, Newport, E (2016) Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient to do so. In: The 38th Cognitive Society Annual Meeting. Mahwah, NJ. *Cognitive Science Society*, pp. 2321–26
- Skinner, B. F. (1957). *Verbal behaviour*. New York, NY: Appleton-Century-Crofts.
- Studdert-Kennedy, M., & Terrace, H. (2017). In the beginning. *Journal of Language Evolution*, 2(2), 114-125. doi:10.1093/jole/lzx005
- Wildman, D. E., Uddin, M., Liu, G., Grossman, L. I., & Goodman, M. (2003). Implications of natural selection in shaping 99.4% nonsynonymous DNA identity between humans and chimpanzees: Enlarging genus *Homo*. *Proceedings of the National Academy of Sciences*, 100(12), 7181-7188. doi:10.1073/pnas.1232172100
- Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: The MIT Press.