Univerzitet u Sarajevu—Filozofski fakultet

Odsjek za anglistiku

Uniqueness of human language and communication demonstrated through Hockett's design features of language/ Jedinstvenost ljudskog jezika i komunikacije prema Hockettovim obilježjima jezika

- završni magistarski rad -

Student: Arnela Kučinar

Mentor: prof. dr. Nejla Kalajdžisalihović

Sarajevo, 2023.

Contents

Sažetak

Cilj ovog završnog magistarskog rada je predstaviti obilježa ljudskog jezika (engl. *design features of language*) po kojem se ljudski jezik razlikuje od različitih sistema komunikacije u životinjskom svijetu. Metodologija izrade zasnovana je na predstavljanju obilježja ljudskog jezika koje je detaljno predstavio američki lingvista Charles Hockett. Neka od obilježja ljudskog jezika kojima će biti posvećena posebna pažnja u radu i u kontekstu savremenog (digitalnog) doba su: proizvoljnost (engl. *arbitratiness*), dvojnost (engl. *duality*), razmještenost (engl. *displacement*), kulturni prijenos (engl. *cultural transmission*), refleksivnost (npr. jezik kao metajezik), kreativnost (engl. *creativity*) itd. Ova obilježja se kontrastivno analiziraju u odnosu na njihovo prisustvo ili neprisustvo kod nekoliko životinjskih vrsta spomenutih u izvorima koji se detaljno bave ovom temom te izvorima u kojima se detaljnije predstavljaju do danas istraženi slučajevi (ne)uspješnog podučavanja ljudskog jezika na primjeru čimpanzi (Lana, Kanzi, Washoe). U samom radu, propituje se da li je navedenim karakteristikama ljudskog jezika moguće doprinijeti u razvoju teorijskog okvira uzevši u obzir razvoj savremenih tehnologija (npr. *speech-to-text technologies*) i postavljenih istraživačka pitanja.

Abstract

The aim of this master's thesis is to present design features of language that differentiate language of humans from different communication systems used in the animal world. The methodology is based on presenting the design features of language and proposed by an American linguist, Charles Hockett. Some of the features that will be given a special attention in the paper and in the context of the contemporary (digital) world are: arbitrariness, displacement, cultural transmission, reflexiveness (and metalanguage as such) and, last but not least — creativity. These features are contrastively analysed in terms of their presence or absence in several animal species while referring to various sources about cases of (un)successful attempts of teaching human language to animals (e.g., cases of Lana, Kanzi, Washoe).

The paper also attempts to provide responses on the question whether the listed design features of language may be expanded or modified taking into consideration the development of contemporary IT technologies (such as e.g., *speech-to-text technologies*).

1 Introduction

From the moment we are born, our voices are heard, and we are learning to communicate with the world around us in a step-by-step manner. First by cries, later by uttering words — until one day, eventually, we start saying our first strings of words and full sentences. Later, as we get older, we are able to better express ourselves, to tell what we think, how we feel, what we like and dislike.

We could say that humans are storytelling beings. No other species has the capacity for language and ability to use it in the way the humans do, especially not in endlessly creative ways. Moreover, even though some experiments had been conducted on some animals, which will be discussed later in the paper, there is a low chance, if at all, that animals are able to acquire the language in the way humans do. Animals do communicate but in different ways, namely their communication does not include language like ours.

So, what is actually language and how could we define it? According to Chomsky, "language is a component of the human mind, physically represented in the brain and part of the biological endowment of the species" (Chomsky, 2002: 1). In his views, language is something innate, we are born with the ability to acquire language and use it as tool for communication. But what about other species, such as animals?

We know they communicate in many different ways, but would they ever be able to express themselves using language like humans? We will discuss these questions throughout this master's thesis. Moreover, we will also discuss past, present and future influences on language, the way philosophers, linguists and biologists shaped it and discuss what we know today.

Since language has always been evolving, we will also talk about present changes in different languages, what influences these changes and how has AI shaped language in the past few years. With the development of technology, there is also an increase in different ways of communicating among humans, which is an important feature to add to the theoretical framework on the design features of language that will be discussed in further text as a foundation for the research this thesis aims to explore.

2 Theoretical Framework—Biolinguistics and Psycholinguistics

2.1. A biolinguistics perspective

Since the development of human language is not only a question for linguists and linguistics, but also the question to which biologists wanted to provide answers, it is important to mention the emergence of *biolinguistics*. Finch (2010), for instance, claims that most contemporary scholars agree on the explanatory necessity for *protolanguage* explained as follows: language did not spring into being all at once but developed in stages, each one of them serving some function of its own. Proponents of *lexical protolanguage* suggest that language started with isolated, meaningful spoken words: i.e., speech and semantics came first, and syntax last. In contrast, proponents of *gestural protolanguage* suggest that language started in a "manual modality", and that syntax and semantics preceded speech. Finally, proponents of *musical protolanguage* argue that speech initially arose as complex learned vocalizations, more like song than speech, and that semantics was added to this system later (Fitch, 2010, p. 9).

According to Chomsky, the biolinguistics perspective views a person's language in all its aspects—sounds, meaning, structure. Jose-Luis Menidivil, for instance, explains that biolinguistics is a branch of linguistics. From his point of view, biolinguistics is not seen as a combination of biology and linguistics or as an application of biology to the study of language (if such things could be possible at all). Rather, *biolinguistics* is the name given to a branch of *linguistics* that forms a part of natural science. Or, in other words, biolinguistics belongs to the discipline that studies human languages from the viewpoint of natural science. Therefore, him claiming that biolinguistics is a kind of linguistics, his intention is not to deny that biolinguistics is a biological study of language, but only to point out that the expression *biological study of language* should be understood not as *the application of (current) biology to the study of language* but as *the amplification of biology so that it can study language*. This means that biolinguistics should be considered as an abstract layer of the biology of language, a decisive step for the integration of the study of language in natural science (see Menidivil, 2018 and Mendívil-Giró, 2014).

The emergence of the studies on the nervous system, the brain, the human brain, and of the language faculty added huge amounts of additional expertise and caused the subsequent emergence of disciplines that address these areas of complexity (neurobiology, psychology,

linguistics, and other cognitive sciences). In terms of interdisciplinary perspectives, as Gell-Mann suggests, "the enterprise of science involves investigating those laws at all levels, while also working, from the top down and from the bottom up, to build staircases between them" (Gell-Mann, 1994, p. 112).

Biolinguistics contains both investigating the "additional information" of language structure and, at the same time, contributing to the building of staircases in search of unification and principled explanation and "although biology and linguistics have traditionally travelled quite separate paths, there is growing evidence that a rapprochement and synthesis is in the making" (Finch, 2010, p. 4). The two main traits that characterise biolinguistics as a science in relation to other branches and fields of linguistics are related to *methodological naturalism* and *internalism*.

The expression *methodological naturalism* used in this context simply implies that biolinguistics is a kind of linguistics that uses the same methodology as the natural sciences. Of course, this methodological program does not justify in itself that we call the discipline biolinguistics. If anything, it would justify us in calling it natural linguistics or something to that effect. In fact, one might rightly object that the bio- prefix is present because Chomskyan linguistics postulates that the object of inquiry (that is, language) is a natural object (and more accurately a biological one) and not just because it uses the same methodology as the natural sciences. Furthermore, Chomsky has repeatedly stated (see, for example, Chomsky 2000, Chomsky 2002) that language is a mental organ and just another "natural object".

Nevertheless, Chomsky has not been very keen on the term biolinguistics, perhaps because his conception of science does not really grant importance to the distinction between calling something "mental", "neurological", "chemical", "magnetic", " electrical" or "physical"; these matters are purely empirical ones that depend on the historical degree of development of various disciplines, whereas the relevant point is to consider language as a natural object (Mendívil-Giró, 2014). This reasoning explains why Chomsky's intellectual commitment has been always to *methodological naturalism* (i.e., 'the mental' and 'the physical' must be dealt with using natural science) and not to *ontological naturalism* (i.e., 'the mental' is part of 'the physical'). As Chomsky states: "Unless offered some new notion of 'body' or 'material' or 'physical', we have no concept of naturalism apart from methodological naturalism" (Chomsky, 2000, p. 143).

McGilvray (2013, p. 46) suggests the more appropriate: "bio-chemico-physico-compulinguistics". Nevertheless, even in the domain of modern cognitive science, it is not strange to discover that ontological naturalism gives rise to methodological dualism and not, as expected, to methodological naturalism. Furthermore, and according to Mendívil-Giró, "there may be several reasons for this surprising fact, but I find that the main one is the inherent difficulty of the logical path from ontological naturalism to methodological naturalism. If we start from ontological naturalism, we must assume that the natural sciences (that is, physics, chemistry, biology, etc.) should be sufficient to explain the mind and language, for example. However, it is obvious they are not. Neither physics, or chemistry, or biology, can explain or predict the structure and meaning of a passive sentence" (pp. 73-74), for instance.

In addition, also according to Mendívil-Giró, "what is implied in Chomsky's methodological naturalism is that if any theory of language structure is empirically adequate, then that theory is already part of the body of scientific, naturalistic research on language. This stance makes sense if we recognise that we cannot prejudge what kind of physical reality language will have, and if we limit ourselves to studying it like just another natural object. This implies that the discipline at hand, although it does not work with bosons, isotopes or proteins, is a natural science" (p. 74). Biolinguistics is, then, a branch of this 'natural linguistics' and we could, perhaps, define biolinguistics as (an abstract) biology of languages.

According to the same author (Mendívil-Giró 2014), if the notion of languages as abstract organisms is going to make sense at all, we need to turn to the other central feature of biolinguistics: *internalism*. Following Chomsky, internalism is defined as an approach that considers the object of inquiry—that is, the faculty of language (FL)—as an internal property or organ of the mind. What this means is mainly that the primordial source of the structure of FL is not outside the mind and brain but inside it. This internalist conception implies, then, that the mind has its own structure (Ibid.).

Two other notions are discussed by Mendívil-Giró in his paper titled "What are Languages? A Biolinguistic Perspective" (2014). Namely, from a biolinguistic point of view, a language is a person's "language organ". This is what Chomsky (1986) called an internal language (I-language). However, this language organ or I-language is also a historical object and can be discussed across languages:

"Without a doubt, the language organ of a person who speaks Russian is different from the language organ of a person who speaks French. Perhaps both organs share a common unhistorical layout (which we conventionally call Universal Grammar, UG), but they differ because of contingent events that we can only explain historically. Migrations, diverging

changes, borrowings and isolation have produced two different natural objects (in fact, millions of them—as many as there are speakers of both languages). Still, although the I-language of a person who speaks French and that of a person who speaks Russian are historically different, this difference should not lead us to think that they are purely historical objects— in the same way that a rhinoceros and a gazelle are different historical objects but are also natural objects. External languages (E-languages) such as Russian or French (that is, things we can say are spoken in Moscow and Saint Petersburg or in Paris and Québec and that they are written with Cyrillic or Latin alphabets) are also historical objects but are not natural objects. They are, as Saussure (1916) put it, social institutions" (Mendívil-Giró, 2014, pp. 80-81).

2.2. Noam Chomsky on contributors to linguistic studies

In his book *Language and Mind* (2006), Noam Chomsky gives references to scholars and scholarly research that has contributed to linguistics. He mentions the technological advances of the 1940s such as the appearance of the computers, the spectrograph that offered the physical analysis of speech sounds. Moreover, only a few years later machine translation and automatic abstracting were around the corner (Chomsky, 2006, p. 3).

However, Chomsky was also being sceptical of technological advances, because according to him, there was no reason to expect that the available technology can provide a significant insight, understanding or achievements. He also admitted that those judgements are harsh, but with reasons (Chomsky, 2006, p. 4). On the other hand, at the same time, he admits that there had been significant advances in understanding the nature of *linguistic competence*, and as such, they had proceeded from assumptions very different than those that were present in the period he had been discussing. Moreover, he states that at that moment in the development of linguistics and psychology in general, it is alright to turn again to classical questions and see the way they can provide direction for contemporary research and study (Chomsky, 2006, p. 5).

2.3. Chomsky on Juan Huarte and the relevance of Huarte's conceptions for psycholinguistics

Chomsky claims that he tried to find similarities between the intellectual climate of the 17th century and that of today. Therefore, as his starting point, he begins with mentioning the writings of the Spanish physician Juan Huarte, who published a widely translated study of the nature of human intelligence in the late 16th century. In the course of his investigations, Huarte came to wonder at the fact that the word for "intelligence", *ingenio* seems to have the same Latin root as various words meaning "to engender" or "generate". This is what gives a clue to the nature of the mind according to Huarte. Thus, one may discern two generative powers in man, one common with the beasts and plants, and the other participating of spiritual presence. Wit (*ingenio*) is a generative power. The understanding is a generative faculty" (see Chomsky, 2006, p. 8).

Huarte also goes on to distinguish three levels of intelligence. The lowest of these is the "docile wit", meaning that there is nothing in the mind that is not simply transmitted to it by the senses. The next, higher level, the human intelligence, is able "to engender within itself, by its own power, the principles on which the knowledge rests" (Preface, third edition, viii). Normal minds are such that "assisted by the subject alone without the help of anybody, they will produce a thousand conceits they never heard spoke of … inventing and saying such things as they never heard from their master, nor any mouth" (Chomsky, 2006, p. 8). Therefore, normal human intelligence is capable of acquiring knowledge through its own internal resources and by that is capable of generating new thoughts by high they can express their own experiences. The third kind of wit that Huarte mentions is creativity, further he says, "by means of which some, without art or study, speak such subtle and surprising things, yet true, that were never before seen, heard, or writ, nor ever so much as thought of" (Chomsky, 2006, p. 8).

A good example of this is definitely literature, especially genres of fiction and fantasy. Personally, what comes to my mind immediately are Harry Potter books and *Lord of the Rings*. It is really amazing how humans can produce so many words, thoughts, sentences, and how many there are actually that still have not been uttered nor written yet. When we just think about it, about all of the possibilities by which we can express ourselves, our thoughts and emotions, it truly is a gift, many are not aware of.

What Chomsky also mentions regarding Huarte is how language is an index of human intelligence, of what distinguishes humans from animals, and specifically his emphasis on the creative capacity of intelligence. With the rise of romanticism, attention shifted to the third type of wit, to "true creativity" (Chomsky 2006, p. 9).

When it comes to psycholinguistic studies, it is in this part of the paper that we will refer to how Huarte's conceptions are particularly relevant to contemporary psycholinguistics, such as:

(a) the brain as the material site for what we would call cognitive functions
(b) the innateness of cognitive functions
(c) the generative quality of human understanding
(d) qualitative differences between human and animal capacity
(f) creation of words by human convention; not by divine intervention as was frequently believed, and
(g) the defence of the vernacular language use in science.

Chomsky refers to Huarte when he examines the antecedents of transformational linguistics mostly during the seventeenth and eighteenth centuries. Noam Chomsky in *Cartesian Linguistics* (1966) and *Language and Mind* (1968) highlighted the following points inspired by Huarte's work (Chomsky, 1966, see p. 78-80; Chomsky, 1968, see p. 8-9):

(a) Huarte's likely influence on the Cartesian idea which declares that mind is a "cognitive power" with a creative character,
(b) an essential feature of wit (*ingenio*) and understanding (*entendimiento*) is its generative character,
(c) man has two generative powers, one common to the animals, the other common to the spiritual substance and God,
(d) human generative capacity is limited to the generation of internal figures or representations,
(e) Huarte's classification of wits (i.e., creative genius),
(f) the mention of the wit of the eunuchs, one of the disabilities of wits proposed by Huarte,
(g) language as an index of human intelligence (*ingenio*), and
(h) language as a distinctive feature of human intelligence.

As stated earlier, typical of later thought is his reference to use of language as an index of human intelligence, of what distinguishes humans from animals, and specifically, his emphasis on the creative capacity. These concerns dominate both rationalist psychology and linguistics (Chomsky, 1968, p. 9).

2.4. Cartesians and Port Royal Grammar

According to Chomsky, a significant gap separates the system of concept on one hand, and the nature of human intelligence on the other. He further states that a similar realization lies at the base of Cartesian philosophy. Descrates also came to a conclusion in his early investigations that the study of the mind shows a problem of quality of complexity, and not merely degree of complexity (Chomsky, 2006, p. 5). What is interesting is that this argument traces back to the works of the now almost forgotten Cartesian philosophers such as De Cordemoy (Chomsky, 2006, p. 5). Descartes argued that the only sure indication that "another body possesses a human mind", is its ability to use language in the "normal way"; and he argued that this ability cannot be detected in an animal, which shows signs of apparent intelligence exceeding those of a human, even though such an organism might be fully endowed as a human with physiological organs necessary to produce speech (Chomsky, 2006, p. 7).

In Cartesian view, even animal behaviour is potentially infinite in its variety. It means that if animal behaviour is controlled by external stimuli or external states, then as the stimuli vary over an indefinite range, so may the behaviour of the animal. But, the normal use of language is not only innovative and potentially infinite in scope, but also free from the control of detectable stimuli, either external or internal. It is because of this freedom stimulus control that language can serve as an instrument of thought and self-expression, as it does not only for the exceptionally gifted and talented but in fact, for every normal human. According to Chomsky, honesty forces us to admit that we are still far from understanding what enables a human to speak in a way that is innovative, free from stimulus control and also appropriate and coherent. The properties of human thought and human intelligence emphasized by the Cartesians are real enough; they were then, as they are now, beyond the bounds of any well-understood kind of physical explanation. Neither physics, nor biology nor psychology give us any clue to how to deal with these matters (Chomsky, 2006, p. 11).

As the Port-Royal Grammar (1660) was influenced by Descartes, it is not unusual that the central argument of the Port Royal Grammar was that grammatical rules are inborn, universal and mental in origin.

2.5. Ferdinand de Saussure

Chomsky also mentions a Swiss linguist Ferdinand de Saussure who, at the turn of the century, laid the groundwork for modern structural linguistics and claimed that the only proper methods of linguistic analysis are segmentation and classification. In fact, de Saussure occasionally expressed the view that processes of sentence formation do not belong to the system of language at all—that the system of language is restricted to such linguistic units as sounds and words and perhaps a few mixed phrases and a small number of very general patterns; the mechanisms of sentence formation are otherwise free from any constraint imposed by linguistic structure as such (Chomsky, 2006, p. 17).

In his terms, formation is not strictly a matter of *langue*, but is rather assigned to what he called *parole*; it is a process of free creation, unconstrained by linguistic rule except insofar as such rules govern the forms of words and the patterns of sounds. In addition, it needs to be mentioned that not much attention is given to syntax as such in the period of structural linguistics.

Ferdinand de Saussure emphasized that both language and speech are composed of a system of signs (Saussure, 2011, p. 17). The language and speech both precipitated from the social environment. He also insisted that language was not an innate act. Ferdinand de Saussure theorized that the person's language and speech are learned while the person is interacting with other persons and influences. For example, a Japanese person learned to speak the Japanese language while growing up in a Japanese community. Likewise, a British child learned (UK) English from conversing with one's relatives and friends. An Arab child could speak fluent Arabic because he or she had grown up in a community where Arabic is the official language. Or, for instance, a Chinese child will learn the Chinese language from his or her close relatives and friends.

In addition, Ferdinand de Saussure also teaches that people learn to follow instructions through learning. The child learns the word "stop" which may mean that the child should halt actions that are meant to be stopped.

The child also learns that the word “sleep” which may mean that the child should go to his or her room to go to sleep. The child learns the word “eat” which may mean that the child must gobble a morsel of food (Saussure, 2011, p. 20). Chomsky, on the other hand, insists that child’s use of *universal grammar* bridges the gap between linguistic stimuli of the child and the (rich or impoverished) linguistic input that the child receives from his or her parents or relatives or caregivers.

Taking all these well-known approaches into account and expanding them in the given context, in this thesis I will try to answer the following research questions:

RQ1: Is there a theory that provides a checklist for design features of language?

RQ2: If yes, does it distinguish between language of humans and other species, animals in particular?

RQ3: Could this list be expanded by new theories on what human language is when it comes to its production, comprehension, change and usage?

Therefore, this master’s thesis aims to (1) present features that distinguish human language and animal communication systems within the theoretical frameworks of biolinguistics and psycholinguistics, (2) identify functions/features of language specific to humans, (3) provide suggestions that may enhance learning in relation to language functions, (4) discuss issues related to cases and experiments on apes and (5) relate them to contemporary research and potentially new features of human language arising with the advancement of technology.

3 Language and Features of Language

Before we start the introduction on language design features and their author, Charles Hockett, I think that it would be of a great importance to provide definitions(s) of language. So, what is actually language?

According to Pagel, "language is a system of communication uniquely associated with humans and distinguished by its capacity to express complex ideas. Notably, studies analysing the various features of human language have informed our understanding of language as a distinctly human trait. Specifically, language is thought to possess a highly structured system of encoding and representing concepts through either speech sounds or manual gestures, depending on whether they are spoken or signed" (Pagel, 2017). So, we could say that our language, more than anything else, is what makes us human in the way that its features and power allow us to share our thoughts, talk about places and events distant or unexperienced, talk about language using language, creatively express new ideas without attending a language course in our L1, etc. How is this to be proven with regard to other species, for instance?

Currently, there are more views on the nature of the relationship between language and animal communication. Some scholars have proposed a divide between the two phenomena, with language considered as "uniquely human" (Chomsky, 1965). Other scholars have argued for a "linear continuum in line with gradualist evolutionary principles" (Bickerton, 1990). It is also a question to be discussed here whether there is an agreed upon definition of language. Numerous authors have proposed different definitions, trying to provide and propose elements and features that may be considered fundamental foundations to language.

De Saussure (1966), for instance, proposed what may be called "arbitrariness", i.e., that the sound of a word has not direct relation to its meaning. This term will be further explained in the lines that follow.

Dor (2015), for instance, offers a very interesting perspective on language as a social communication technology and argues that language needs to be "invented and constructed" as well as that it "constantly develops as a result of usage".

3.1. Charles Hockett and the design features of language

According to biographical notes on Charles Hockett in *Biographical Memoirs: Volume 89* (2007, pp. 150-179), Charles Hockett was a leading figure in American structuralist linguistics, which flourished particularly in the four decades from the 1930s to the 1960s. Structuralist linguistics was sometimes referred to as Bloomfieldian linguistics from one of its pioneering figures, Leonard Bloomfield. Hockett considered Bloomfield his great influence and was an inheritor of Bloomfield's unfinished work. Hockett referred to his own influential work *A Course in Modern Linguistics* (1958) as "a commentary on *Language*" (see Bloomfield, *Language* (1933)). Hockett was considered by many to be a very influential and important contributor to linguistic theory in the framework of structural linguistics. He believed linguistics is a branch of anthropology, to which he also made serious contributions.

One of his most famous works is *The Origin of Speech* (1960) in which he outlined 13 features of human language. In the study, he listed 13 design features that he deemed to be universal across the world's languages. More importantly, these features distinguished human language from animal communication. While the first 9 features could also match primate communications, the last 4 were solely reserved for human language. Later on, Hockett added 3 features that he saw as unique to human language. Thus, it can be said that human language shares a general set of features that help set it apart from communication among animals. Aitchison (2008), for instance, says that Hockett's longest list contained 16 features and that most authors would agree that 11 features capture the essential nature of language (p. 28). Since the number of design features Hockett considered important changed over the years, in this paper, I will present 13 features:

1. **Vocal-auditory channel**

- except signed languages, natural language is vocally transmitted by speakers as speech sounds and auditorily received by listeners as speech waves

- writing and sign language both utilize the manual-visual channel

- the expression of human language primarily occurs in the vocal-auditory channel

2. **Broadcast transmission and directional reception**

- language signals (i.e., speech sounds) are emitted as waveforms, which are projected in all directions ('broadcasted into auditory space')

- waveforms are perceived by receiving listeners as emanating from a particular direction and point of origin (the vocalising speaker)

3. **Transitoriness**

- language signals are considered temporal as sound waves rapidly fade after they are uttered (rapid fading)

- this temporal nature of language signals requires humans to receive and interpret speech sounds at their time of utterance, since they are not subsequently recoverable

4. **Interchangeability**

- humans can transmit and receive identical linguistic signals, and so are able to reproduce any linguistic message they understand

- 'speaker' and 'listener' may alternate between the conversation's participants via turn taking within the context of linguistic communication

5. **Total feedback**

- humans have an ability to perceive the linguistic signals they transmit i.e., they have understanding of what they are communicating (except in certain conditions caused by health and age-related issues)

- this allows a continuous monitoring of the output to ensure they are relaying what they are trying to express

6. **Specialization**

- language signals are emitted for the sole purpose of communication, and not any other biological functions such as eating

- in normal circumstances, language signals are intentional, and not just a side effect of another behaviour

7. **Semanticity**

- specific language signals represent specific meanings; the associations are 'relatively fixed'

- an example is how a single object is represented by different language signals i.e., words in different languages

- likewise, the crying of babies may, depending on circumstances, convey to parents or caregivers that it requires milk, rest or a change of clothes.

8. **Arbitrariness**

- there is no intrinsic or logical connection between the form of specific language signals and the nature of the specific meanings they represent

- the signal and the meaning are linked by either convention or instinct

It is interesting to mention that "arbitrary symbols are not unique to humans. Gulls, for example, sometimes indicate aggression by turning away from their opponent and uprooting beakfuls of grass" (Aitchison 2008, p. 29).

9. **Discreteness**

- language signals are composed of basic units and are perceived as distinct and individuated

- these units may be further classified into distinct categories

- these basic units can be put in varying order to represent different meanings

10. **Displacement**

- displacement may be used in the context of explaining the human ability to lie or produce utterances which do not (necessarily) correspond with reality ("here and now")

- language signals may be used to convey ideas about things not physically or temporally present at the time of the communicative event such as a topic that is linked to the past or future

11. **Productivity**

- productivity is also called openness or creativity

- it entails reflexiveness, the ability of language to be used to talk about language

- humans can use language to understand and produce an indefinite number of novel utterances

12. Cultural transmission

- although humans are born with the innate ability to learn language, they learn (a) particular linguistic system(s) as their native language(s) from elders in their community

- language is socially transmitted from one generation to the next, and a child reared in isolation does not acquire language

13. Duality of patterning

- the discrete speech sounds of a language combine to form discrete morphological units, which do not have meaning in itself

- these morphemes have to be further combined to form meaningful words and sentences (adapted from NTU Libraries Singapore, on Design Features, 2014).

To conclude, Hockett's list of the design features of language has been cited across disciplines but also severely criticized. For example, social transmission in language, or 'tradition' to use Hockett's terminology, has been argued to have more to do with conceptual content and semantic information being built up across generations of humans, whereas for animal communication the use of the term 'social transmission' relates more to developing communicative modalities, like the vocal-auditory channel (e.g., Garland and McGregor, 2020). Wacewicz and Zywiczynski (2015) in particular state that Hockett's language design features focus too much on communicative means and structure, rather than the underlying cognitive aspects of language, from theory of mind to executive function. Other scholars as well have presented briefer attempts than Hockett on comparing language with animal communication but are rarely cited (see Amphaeris, 2022).

3.2. Human language vs animal communication systems

Human language, according to Pagel (2017) and numerous other scholars, is unique among all forms of communication. Sign language that will be mentioned in the context of the great apes is nothing like human language and it lacks structure dependency that is discussed in great detail in *The Articulate Mammal* (Aitchieson, 2008).

Another term that is introduced by Pagel (2017) is "compositional" in the sense that human language allows us to express our thoughts in sentences subjects, verbs and objects—such as 'I kicked the ball'—and recognizing past, present and future tenses. Compositionality gives human language an endless capacity for generating new sentences as speakers combine and recombine sets of words into their subject, verb and object roles. For instance, with just 25 different words for each role, it is already possible to generate over 15,000 distinct sentences. Moreover, human language is also *referential*, meaning speakers use it to exchange specific information with each other about people or objects and their locations or actions (Pagel, 2017).

We as humans can always ask what the other person meant if the message we heard was not understood properly. However, when it comes to animals, we can never be completely sure that what they signalled was completely intentional or not. The way animals communicate has always been a mind-boggling topic to linguists and according to Adret (2001) there will always be more to research about it.

Although, we cannot give the best possible definition of a language that will completely describe the way it is, I think that currently the best one which Lawrence (1998) mentioned in his book would be Chomsky's definition of language stating that human language is as he said the "*last bastion of human uniqueness*". According to Pagel (2017), animal communication is different than human language in the sense that animals use: sound signalling for object or an action, they are not able to use their communication in creative ways or to exchange information. This assumption could be discussed and questioned in the context of displacement in bees and turn-taking in birds (see Aitchieson, 2008).

3.3. The influence of the home-raising experiments

During the twentieth century, some scientists that will be mentioned later came up with the idea to create an artificial environment so that the animals may start learning and speaking human language. In this case, it was about chimpanzees, the closest relatives to humans. They thought that if the animals were surrounded by humans and human language since a young age, that they would be able to pick up the words and even start talking. To some, this idea may seem brilliant, to some it may raise questions as how far would humans go just to prove something what they believe in.

According to Kellog (1968) it became a usual practice to say so, to keep nonhuman primates as animals and even conduct experiments on them. Ever since, it was a usual practice in some countries (see Great ape research ban). What poses a problem is that the majority of those who were doing the experiments were not trained, did not have any knowledge of animal training and also were ill-equipped so they were not able to record animals' reactions properly. As Kellog (1968) points out, it would be a much different story if the experiments were done by trained and qualified psychobiologists in controlled experimental surrounding.

The experiments had been conducted on many chimpanzees, but in this thesis, we will take a look at some of them which include: Viki, Nim, Kanzi, Lana and Washoe.

3.4.1. Viki

The first case that will be discussed is Viki. Viki also remains as the most successful attempt to make a chimpanzee speak, conducted at the Yerkes Laboratories of Primate Biology in Florida. The experiment was led by Keith and Catherine Heyes. They adopted Viki only a few days after her birth and tried to raise her in an environment suitable for a human child. Viki lived there for six years and according to Hayes (Hayes and Hayes, 1952) during this time she learnt four words that could be recognised as "mama", "papa", "cup" and "up".

Although she managed to utter these words, it was obviously difficult to pronounce them. Even though the training lasted for several more years, not much advancement could have been made. She could only pronounce or, better say, try to do so with the words that she first had learnt ("mama" and "papa"), but the difficulty in doing so was inevitable. The words such as "cup" and "up" were easier to pronounce because they resemble the sounds that chimpanzees naturally produce (Fouts & Rigby, 1980).

3.4.2. Nim and ASL

Nim Chimpsky, unlike Viki, Nim was trained to learn the American Sign Language (ASL). The so called "Project Nim" was led by a psychologist Herbert Terrace who wanted to do was prove that Chomsky was wrong. As Cowie (2008) put it, Chomsky has claimed that for human children to learn a language they must know the rules of the Universal Grammar which, according to him, are something innate, something we are born with.

Nim was brought to New York where he lived with Terrace's family for few years. As a result of the experiment, Nim learnt around 125 signs. But, what did that show? According to (Hartsfield, 2022), Terrace was not satisfied with the results and had admitted that Chomsky was right because, although the chimp learnt the signs, it was not able to use the language the way humans do, to create sentences and express ideas.

In an interview for Columbia News (Project NIM Revisited, 2023), Terrace said that the uniqueness of the human language lies in the ability to name things and create new meanings by combining words, and such things are beyond a chimp's ability.

3.4.3. Kanzi

The third experiment was conducted on Kanzi, a bonobo[1] that learnt to use symbols and lexigrams (symbols representing words). The experiment was first attempted on his mother, but instead Kanzi learnt almost 400 symbols (Wayman, 2011). The chimpanzee was raised by Duane and Sue Savage-Rumbaugh. It underwent rigorous control procedures such as hiding facial expressions or eye movement under the mask or having Kanzi wear earphones so that he can hear the words being read from another room (Donald, 2001).

With such strict procedures, they wanted to make sure that the results were clear and that the chimp was not being hinted with the possible correct answers. However, despite all the restrictions, the results still did not seem fruitful enough to say that Kanzi passed the experiment. Even though Kanzi had learnt several hundred symbols and could identify photos, point at objects, and understand sentences such as: "Take the vacuum cleaner" or "Give Pinky some water" (Donald (2001)), it still was not able to express its mental state, feelings or thoughts.

As Donald (2001) stated, Kanzi could comprehend grammatical relationships between some words in a sentence, but that was the limit, since Kanzi did not have that human ability to say so, to use it for anything deeper and that would demand him to have the innate capacity that always somehow seems to get in the way during these experiments and also shows us once again that human language is much more complex than it has been thought.

[1] also called the *pygmy chimpanzee*, an endangered great ape.

3.4.4. Lana

Another chimpanzee that was used for the experiment was Lana. Lana was used as a subject to question her language capacity by a computer-controlled training led by Rumbaugh, Gill and von Glasersfeld (1973). After six months, she was able to read projected word characters, complete an incomplete sentence based on its meaning and serial order and also reject incomplete sentence that were grammatically incorrect (Fouts & Rigby, 1980).

The computer in questions was PDP-8 controlled by two separate consoles each containing twenty five keys. Each key contained a lexigram developed by the experimenters. The symbols were white geometric symbols that could be used separately or in combinations. The keys on which the symbols were displayed consisted of three colours that could be used separately or together. She also could ask for food, music, movies, toys and so on, when she pressed the key correctly (Fouts & Rigby, 1980). When the key is available for use by Lana, it is softly "backlit". When Lana presses a key, it becomes brightly lit. When a key is not available for use, it has no backlighting. The console also could be used between the experimenters and Lana so that the computer behaved as a kind of mediator between them and Lana (Sebeok & Umiker-Sebeok, 1980).

According to Fouts & Rigby (1980) Lana's training would begin with her pushing a single key on the console. She was requested to start with a "please" and end with a "period". The computer would analyse the sentence and if it was correct the tone was played and Lana was rewarded with what she had requested; if not, the computer would erase the projection display and the keys would be reset on the console. What was interesting was that Lana learnt to attend to the lexigrams on the projectors without training. During her spontaneous learning, it was possible to observe and examine her ability to read sentence beginnings, to discriminate between valid and invalid beginnings, and to complete sentences. Lana ranged from 70% to 100% correct on the various beginnings. They concluded that Lana accurately read and perceived the serial order in Yerkish, the sign language of great apes, and was able to discriminate between valid and invalid beginnings of incomplete sentences in order to receive a reward.

3.4.5. Washoe

The last chimpanzee that will be discussed is Washoe. Washoe was an infant female chimpanzee between 8-14 months old in June 1966 when she was adopted by Allen and Beatrice Gardner in the United States. She grew up in their backyard and lived in a house trailer which had a toilet, a kitchen and a bed. During this project, the Gardeners only used Ameslan (American Sign Language) to communicate with Washoe and also with one another in her presence (see Gardner and Gardner, 1975, Fouts and Rigby, 1980).

The project began in June 1966 and ended in October 1970. During those years, Washoe accumulated a vocabulary of over 130 signs. Despite such a rather small number of signs, Washoe was able to use them very well. She was able to produce spontaneous combinations of signs that demonstrated her ability and also correctness. According to Schoenemann (2002), she signed correctly with such ease using different signs and combinations when the new situations occurred. The researchers were with Washoe throughout the day in shifts. They were trying to immerse her in Ameslan as much as possible. During activities such as cooking meals, cleaning, brushing her teeth and correcting her lapses in toilet training (Fouts and Rigby, 1980).

Washoe was recorded daily. The researchers wanted to see if the signing was spontaneous or prompted. Washoe would sign correctly when spontaneously answering a question or making her own request such as “open”. For prompted signs, she would need the researchers’ assistance. Infomation about correctness was being recorded as well. The criterion was based on fifteen simultaneous days during which she should show spontaneous and correct signing. The Gardeners kept a diary of various signs, when and how she used them, in which situations and contexts. After thirty-six months, she used 85 signs and by June 1974 the number went up to 160 (see Fouts & Rigby, 1980).

During the experiment with Washoe, several methods were used. Fouts & Rigby (1980) included the following:

(1) Manual babbling

The first method used for the experiment with Washoe was *manual babbling*. It was considered to be an equivalent to the infant babbling. After the end of the second year and Washoe’s progress in acquiring vocabulary, the babbling significantly decreased. They came to a

conclusion that, with the acceleration of signing, the babbling decreased. It could be said that happens in humans when normal speech starts developing.

(2) Shaping

This method includes awarding Washoe every time she acquires a new sign. Fouts & Rigby (1980) mentioned that the sign for *open* consists of placing the two open palms against the object to open it and then moving them up and apart. Washoe quickly learnt the sign and used it spontaneously for doors, books, boxes and drawers. At first, it seemed as a great procedure used to introduce new signs but soon it turned out to not be as efficient as other methods.

(3) Guidance

According to the Gardners, this method was considered the most effective for Washoe. Fouts & Rigby (1980) described it as placing hands and arms in the appropriate positions for the sign, usually in the presence of an object or action representing the sign. The Garderners gave an example of the sixth sign that Washoe acquired, and that was *tickle*. The sign was made by holding one hand open with fingers together, palms down, drawing the extended finger of the other hand across the back of the first hand (Fouts & Rigby, 1980).

(4) Moulding

According to Fouts (1972), other methods used for sign acquisition are used: moulding, which involved physically guiding Washoe's hands into the correct position and movement for the sign, imitation, during which the experimenter would show the sign which Washoe would imitate, and free style, which can be considered a combination of moulding and imitation. Moulding was proved to be the procedure that produced the most rapid acquisition of signs followed by free style and, lastly, by imitation.

(5) Observational learning

The last method that was used during the experiment with Washoe was *observational learning.* It consisted of practicing Ameslan in Washoe's presence, and was an equivalent to Fout's (1972) use of imitation. Washoe learnt to sign *sweet* and *flower*. He signed sweet by touching the lower lip or the tongue with the extended index and second finger of one hand, while

remaining fingers were pressed into palm. They stated that she acquired the signs effortlessly after months of exposure (Gardner & Gardner, 1971).

To quantify the accumulated data correctly, new ways of testing had to be designed. After several attempts, they finally found the right method; the test satisfied both, the researchers and the chimpanzee (see Fouts & Rigby, 1980, Chapter 37):

(1) Tests using flash cards

Washoe was shown pictures on large cards and then questioned. However, the test was paced and required a great amount of discipline to do it correctly and on time.

(2) Tests using a box

Washoe was supposed to identify three-dimensional objects placed in a box. Although the test was better than the one with the flash-cards, it was logistically difficult to achieve.

(3) The slide test

The slide test was the last and also the easiest one. It consisted of 35mm colour transparencies used as examplers. It was efficient and doable, and most importantly paced by Washoe (Fouts & Rigby, 1980, Chapter 37). According to Gardners' report (1971) Washoe has correctly identified 53 items out of possible 99 in the slide test. Although Washoe's performance was more than satisfying and also encouraging to the researchers, what was more interesting were also the errors that fit into conceptual categories such as food, animals or grooming items. The examples of such errors include Washoe signing *dog* instead of a *cat*, *brush* instead of a *comb* and food instead of *meat*.

When Washoe was shown pictures of three-dimensional replicas of objects (e.g., a toy cat, a toy dog) used in the box test, the baby sign occurred frequently among her errors. When the slides were real items, that did not happen Gardner & Gardner, 1971). Washoe made four out of ten errors in the examples such as signing a doll instead of a cat, and all four errors were baby signs.

During the tenth month of the project, Gardeners (1971) claimed that Washoe signed the phrase *gimme sweet* and *come open*. What is even more interesting is the resemblance

between chimpanzee and children in the use of language at the same age between 18 and 24 months (see Aitchison, 2008).

The combinations of signs that Washoe used was done in more or less the same manner the way human signers do. The Gardners (1971) described Washoe raising her hands in the signing space while completing the combination. She would end the combination by touching some object or a surface. That is similar to human singers' hands in repose.

The context in which Washoe would use signing was very good. She would correctly use phrases such as *go in, go out* or in *down bed*. Washoe also signed at a locked door thirteen times in the right context correctly: *gimme key, more key, gimme key more, open key, key open, open more, more open, key in, open key please, open gimme key, in open help, help key in, and open key help hurry* (Gardner & Gardner, 1971).

Between April 1967 and June 1969, the researchers recorded 245 different combinations of three or more signs in their diary. During the analysis of Washoe's recorded combinations, it turned out, according to the Gardners (1971), that Washoe had a preference for word order. The combinations in question included *you me* in 90 percent of time of the taken samples and *me you* in the remaining 10 percent. The Gardners were hesitant to accept this order as an indicator of syntax in Washoe's manual language. They stated that it may only be a mere imitation of researchers' preferred order. This is also interesting in relation to the two-word stage in humans and pivot grammar (Aitchison, 2008, p. 121). Project Washoe had ended in October 1970, after the members of the research team obtained their degrees and left the project. Washoe was brought to the Institute of Primate Studies in Norman Oklahoma. She passed away after a short illness in 2007 at the age of 42 (Carey, 2007).

3.4. Conclusions on home-raising experiments

As Yerkes and Learned (1925) claimed, the chimpanzees resemble the most to humans regarding physiology, intelligence and the ability to imitate. Some other home-raising attempts rather more unsuccessful than not, were reported to have been conducted, but without any considerable differences in vocal language capability such as in chimpanzees. Since the results with the home-experiments were not satisfying, a number of researchers concluded that there are actually differences between humans and apes.

Noam Chomsky (1968) suggests that the uniqueness of the humans lies in the ability to use language in such a way that enables him creating thoughts, which is in line with Hockett's design features of language pertaining not only to creativity but also structure-dependency and understanding deep structures, i.e., that one sentence may have the same deep but different surface structures and vice versa and that order matters. Using and repositioning linguistic symbols, humans have the power to produce new thoughts, to use tenses and with modifying concepts, humans can create new combinations of symbols.

The home-raising experiments resulted in two possibilities. According to Hayes & Hayes, (1951) the only animals capable of complex vocal communication are humans, and that is impossible with lower primates (e.g., speech planning, errors and pauses (filled or breathing pauses) are all subject of various studies on language production).

The second possibility asserts that speech used only by humans is not a suitable medium of communication for chimpanzees. The Gardners (1971) also brought the attention to the fact that chimpanzees' inability to speak is not only behavioural, but also anatomical. Some parts of animal behaviour are highly resistant to modification, which also leads us further to conclude that teaching chimpanzees vocal language was a failure because of the aforementioned resistance to modification.

4 Human Language and Artificial Intelligence

4.1. Natural Language Processing

With the technological advancements over the years and the appearance of the Artificial Intelligence (AI) it was just a matter of time to start implementing it in our daily lives during our most usual tasks. It could be said that we are not fully aware of how AI has changed our lives for worse or for better, but one thing is sure: we are using it more than we are actually aware. Sending voice messages, hearing recordings of our voice that makes us feel uneasy, and the iPhone users can command their virtual assistant Siri to check the weather or play their favourite song just using one channel – *voice*. How and why AI affects human language and what changes has it brought over the last few years will be discussed as well. In order to understand that, we first need to know what Natural Language Processing or NLP is.

NLP is a branch of computer science, more precisely of artificial intelligence or AI. It combines computational linguistics and its rules are modelled on human language with statistical, machine and deep learning models. As such, these technologies enable computers to process *human language* in the form of a *text* or *voice data*, striving to convey speaker's message as accurately as possible. It encompasses computer operations such as translating text from one language to another (e.g., Google Translate, Deep L), responding to spoken commands and summarising large volume of text rapidly in real time (What Is Natural Language Processing? | IBM, n.d.). It also involves voice-operated GPS systems, speech-to-text dictation software or customer service chatbots and many other options.

Since human language is complex and can be ambiguous as mentioned earlier, the main task for the programmers is to apply language rules on applications and make them distinguish the differences and ambiguities that human language has. They would also have to make the AI "understand" the different parts of language: idioms, metaphors, grammar etc. (What Is Natural Language Processing? | IBM, n.d.). When comparing teaching human language to animals, we could also say that humans are conducting the same experiments with artificial intelligence. Therefore, Hockett's design features of language need to be updated and revised in this context as well. It is quite certain that only the language of humans may contain several design features of language simultaneously.

Therefore, despite many advances that the AI has made, I think that no matter how much time and effort is put in the future in improving AI's ability, it will never reach the level of

understanding and using language the way humans do. It can be practical when we do not have time but want to get something done, since simple voice commands or translations are done in no time. But, in the long run, it will never reach or surpass the function that the human language possesses. We cannot integrate the human brain into AI. What I would also like to point out is that sarcasm, metaphor and also idioms cannot be comprehended by AI on a deeper level. Many of such attempts can result and mostly do, in funny and silly mistranslations. Those language layers and its depth is specific for humans' understanding only.

4.2. NLP Tasks

What artificial intelligence programmes can be "taught" is speech recognition or speech -to- text conversions. It can be used for spoken commands and questions. However, it is sensitive in the sense that it cannot process the action if the commands uttered are quick, slurred or mispronounced. The process is even more complicated when it comes to different accents and dialects.

Part of speech tagging is also known as grammatical tagging and it represents the process of determining the part of speech based on its use and context. Word sense disambiguation, for instance, selects the word multiple meanings that make the most sense in given context. It can help distinguish the meaning of verb *make* in "*make the grade*" (achieve) vs "*make a bet*" (place). Named entity recognition or NEM identifies words or phrases as useful entities. For instance, NEM identifies *Kentucky* as location and *Fred* as a name.

Co-reference resolution has the task to identify if and when two words relate to the same entity. The most common example would be determining a person or an object to which a noun refers, e.g., she –Mary. Finally, sentiment analysis has the role to extract subjective qualities including: attitudes, emotions, sarcasm, confusion from the text. "Natural language generation" could be the opposite of "speech recognition", as it has the task of putting structured information into human language.

4.3. NLP usage

NLP is the main force behind machine intelligence. The article (*What Is Natural Language Processing? | IBM*, n.d.) mentions the following cases of its use:

(1) Spam detection

Although it may not be obvious at first, NLP is a part of spam detection. It scans emails searching for words that indicate spam or phishing.

(2) Machine translation or MT

The most well-known and widely available NLP is *Google Translate*. Useful MT involves more than replacing words from one language into another. It has the role of capturing accurately the meaning and tone of the input language and transmitting it the same into output language.

(3) Virtual agents and chatbots

Two most famous known virtual agents include Apple's *Siri* and Amazon's *Alexa*. Both use speech recognition for voice commands and respond in natural language generation with appropriate action or command. Chatbots perform the same in response to typed text entries. They also learn contextual cues about human requests in order to provide them with better responses over time. There is also question-answering, where a Chatbot answers our questions. Moreover, they have a set of readily made questions and answers.

(4) Social media sentiment analysis

NLP has become an important role for revealing hidden data insights from social media platforms. Sentiment analysis has the ability to gather information based on our reactions on social media posts, responses, reviews and more. Collecting the data is useful for companies that can attract the audience and expand the need for their services and products.

(5) Text summarization

In this case, NLP makes huge volumes of digital texts smaller, creating summaries and synopses for indexes and research database. The best applications use semantic reasoning and natural language generation (NLG) to provide the service with useful context and summary conclusions.

5 Computational Linguistics and the Design Features of Language Revisited

According to Schubert's (2020) definition, computational linguistics is scientific and engineering discipline which has the task to understand written and spoken language from a computational point of view and building artefacts that process and produce language in bulk or dialogue setting. A computational language gives us insight into thinking and intelligence. The reason computational linguistics is mentioned at the end of this paper is to reflect on an important design feature of language of humans and that is the ability of humans to *think about language using language* and to play with language data, tag parts of speech, feed electronic texts into various pieces of software, use crowdsourcing, getting information through voice or reading from different parts of the world from different internet users, etc. This feature, or these features, is/are emerging with the emergence of technology and the fact that various channels can now be combined to use language and store data in new and creative ways for future generations.

The main objectives of computational linguistics involve: formulating grammatical and semantic framework of languages allowing computational implementations of syntactic and semantic analysis, the discovery of processing techniques that use properties of language and the development of cognitive and neuroscientific computational models used for understanding language processing and learning in the brain (Schubert, 2020). The aims of the field are broad and varied and as Schubert (2020) mentions, we will list the following:

- efficient text retrieval on a topic
- machine translation (MT)
- question answering (QA)
- text summarization
- analysis of text or spoken language
- sentiment analysis
- creation of computational systems based on human ability for dialogue and acquiring knowledge from text

5.1. Manual knowledge coding, knowledge extraction, crowdsourcing

Language as we already know is complex, and in order to make machines linguistically competent, one needs to possess a deep knowledge and understanding of language. Ideally, the pre-programmed knowledge of a machine would be restricted to kinds of human language that are considered to be innate such as object persistence, basic models of animacy and mind, means of organizing events in time and other know-how (Computational Linguistics, Stanford Encyclopaedia of Philosophy, 2014). Most current methods of knowledge acquisition, according to the same source, include *manual knowledge coding*, *knowledge extraction from text* and *crowdsourcing*. (Computational Linguistics (Stanford Encyclopaedia of Philosophy, 2014).

5.1.1 Manual knowledge coding

A well-known manual creation of commonsense knowledge is the Cyc Knowledge Base (KB). It consists of a few hundred thousand concepts and several million facts and rules. Cyc base has been used in various sectors including business, education and military. It is heavily relied upon specific predicates and high order operations, rather than language understanding. Component Library (CLib) is another example of hand-coded knowledge base. It provides an upper ontology of several hundred concepts and axioms about basic actions and changes. However, the coverage of English lexicon is spare because with the use of frame-based Kleo, knowledge representation is not close to language (Schubert, 2020).

Manually coded lexical knowledge is limited, just as in the cases mentioned above, the reason being a lack of language understanding and vocabulary incompleteness due to always expanding and shifting vocabulary, jargons and styles of expression in all living languages. Besides lexical knowledge, there are also sources if world knowledge such as *tabulations* and *gazettes*, info-boxes in online sources like *Wikipedia* (Schubert, 2020).

5.1.2 Knowledge extraction from texts

According to Schubert (2020), these methods collect information from reliable sources (WordNet) or narrative texts (Wikipedia) mapping it into syntax for exploring generic

knowledge. Indirect methods collect generic knowledge found in miscellaneous reports, stories and essays.

5.1.3 Crowdsourcing

Using the method of crowdsourcing, general knowledge is acquired by soliciting information from large numbers of web users, often with small financial rewards or simple games. Crowding is proved to be reliable in terms of annotation and clarification tasks. *The Open Mind Common Sense* project has a network of informal commonsense knowledge which is based on simple English statements collected from contributors around the world. It has been useful in improving interpretation in speech recognition and some other sectors (Schubert, 2020).

6 Conclusion

The aim of this master's thesis was to introduce design features of language that make the language used by humans different from different communication systems used in the animal kingdom, i.e., by other species. The methodological framework was based on presenting the design features of language in the context of different experiments conducted so far in the attempt to teach apes human language. In the discussion, some of the design features of language were given a special attention such as: arbitrariness, creativity, cultural transmission, metalanguage and displacement. In the discussion, a link was created between different disciplines such as psycholinguistics in relation to biolinguistics and computational linguistics whereas the greatest part of the theoretical framework, aside the design features presented by Hockett, relies on the theoretical works of Noam Chomsky.

With great certainty, I think that we can agree upon one thing: that human language is an amazing gift that we as humans have been given, gifted or born with. During the writing of this thesis, I came to a realisation that we have taken this gift of language for granted. We should be taking more care of it, learning about it, being proud that we have such an amazing ability to express ourselves, to be creative, witty, funny, to say what is happening in and around us and the fact that there are so many languages that we have the absolute ability to learn or at least learn about them.

We have shown through many examples of the long-ago conducted animal experiments (in this case chimpanzees) that language in all its design features is species-specific to humans (especially if we take into consideration what may stand for "learnability"), even though some of them showed some advances during the experiments. I think that this does not indicate a lack of intelligence of the chimpanzees, not at all, because like we already know they are clever in many other ways, but it rather tells us that language is something innate like Noam Chomsky has claimed. All of these experiments were, in my opinion, unnecessary because they were violating animal rights.

When it comes to the AI and its advances, we must admit that they sound really incredible, but not as incredible as the ways in which the human language functions. As we could see, no matter how advanced the AI has been, there will always be a lack of human language and the copy will never be as good as the original. We can say that artificial intelligence may be beneficial to humans in some ways, but it will never outgrow the way humans use and shape language. Still, the knowledges obtained in the field about the design features of language may assist teachers, translators, interpreters and IT specialists in the way that the notion of "learnability" may be expanded for non-humans.

In this thesis, and in terms of research questions, I have presented a checklist for the design features of language using Hockett's classification primarily, but also through consulting other materials and sources. The checklist may be expanded in relation to new approaches to human language in relation to artificial intelligence, for instance. I have also shown, using illustrative examples, the design features of language that are present only in the language of humans. Numerous other examples are available (e.g., examples of dolphins and bar-pressing, imitation as used by parrots, turn-taking in birds, etc.) but are beyond the scope of this paper in which I also attempted to join the different approaches and relate them to the field of translation studies as well.

Finally, this master's thesis has only scratched the surface of such a broad and interesting topic, and I am sure that many other advances will be made in the future when it comes to artificial intelligence as a testing ground for new discoveries about how human language works.

REFERENCES

Adler, M. (2008, May 28). The chimp that learned sign language. *NPR*. https://www.npr.org/2008/05/28/90516132/the-chimp-that-learned-sign-language

Adret, P. (2001). Songs, Roars, and Rituals: Communication in birds, mammals, and other animals. Lesley J. Rogers , Gisela Kaplan. *The Quarterly Review of Biology*, *76*(3), 394–395. https://doi.org/10.1086/394097

Aitchison, J. (2008). *The articulate mammal: an introduction to psycholinguistics.* 5th ed. London; New York, Routledge.

Amphaeris, J., Shannon, G., & Tenbrink, T. (2022). Overlap not gap: Understanding the relationship between animal communication and language with Prototype Theory. *Lingua*, *272*, 103332. https://doi.org/10.1016/j.lingua.2022.103332

Allen J., W. de Beaumont, L. Galescu, J. Orfan, M. Swift, and C.M. Teng, 2013, "Automatically deriving event ontologies for a commonsense knowledge base," in *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)*, University of Potsdam, Germany, March 19–22. Stroudsburg, PA: Association for Computational Linguistics (ACL).

Bickerton, D. (1990). *Language and species*. The University of Chicago Press.

Carey, B. (2007, November 1). *Washoe, a chimp of many words, dies at 42*. The New York Times. https://www.nytimes.com/2007/11/01/science/01chimp.html

Chomsky, N. (2006). *Language and mind*. https://doi.org/10.1017/cbo9780511791222

Computational Linguistics (Stanford Encyclopaedia of Philosophy). (2014, February 26). https://plato.stanford.edu/entries/computational-linguistics/#StaNLP

Donald, M. (2001). *A Mind so rare: the evolution of human consciousness*. http://ci.nii.ac.jp/ncid/BA55483155

Encyclopaedia Britannica. (1998, July 20). *Submissive behaviour | Dominance, Aggression, Hierarchy*. Encyclopaedia Britannica. https://www.britannica.com/science/submissive-behaviour

Fitch, W. T. (2010). *Language*. Cambridge University Press eBooks, 73–142. https://doi.org/10.1017/cbo9780511817779.004

Fouts, R. S., & Rigby, R. L. (1980). Man-Chimpanzee communication. Springer eBooks, 261–285. https://doi.org/10.1007/978-1-4613-3012-7_15

Gardner, B. T., & Gardner, R. (1971). Two-Way Communication with an Infant Chimpanzee. *Behavior of nonhuman primates*, 117–184. https://doi.org/10.1016/b978-0-12-629104-9.50010-8

Gardner, B. T., & Gardner, R. A. (1975). Evidence for sentence constituents in the early utterances of child and chimpanzee. *Journal of Experimental Psychology: General*, *104*(3), 244–267. https://doi.org/10.1037/0096-3445.104.3.244

Gardner, R., & Gardner, B. T. (1988). The role of cross-fostering in sign language studies of chimpanzees. *Human Evolution*, *3*(1–2), 65–79. https://doi.org/10.1007/bf02436591

Garland, E. C., & McGregor, P. K. (2020). Cultural Transmission, Evolution, and Revolution in Vocal Displays: Insights from Bird and Whale Song. *Frontiers in psychology*, *11*, 544929. https://doi.org/10.3389/fpsyg.2020.544929

Great_ape_research_ban.(n.d.). https://www.bionity.com/en/encyclopedia/Great_ape_research_ban.html.

Hartsfield, T. (2022). Koko the Impostor: Ape sign language was a bunch of babbling nonsense. *Big Think*. https://bigthink.com/life/ape-sign-language/

Hayes, K. J., & Hayes, C. (1951). Imitation in a home-raised chimpanzee [Dataset]. In *PsycEXTRA Dataset*. https://doi.org/10.1037/e533882004-001

Hayes, K. J., & Hayes, C. (1952). Imitation in a home-raised chimpanzee. *Journal of Comparative and Physiological Psychology*, *45*(5), 450–459. https://doi.org/10.1037/h0053609

Kellogg, W. N. (1968). Chimpanzees in experimental homes. *Psychological Record*, *18*(4), 489–498. https://doi.org/10.1007/bf03393798

Mendívil, J. (2018, December 11). *What Kind of Linguistics is Biolinguistics?* Philosophy of Linguistics. https://philosophyoflinguistics618680050.wordpress.com/2018/12/11/what-kind-of-linguistics-is-biolinguistics/

Mendívil-Giró, J. (2014). What are Languages? A Biolinguistic Perspective. *Open Linguistics*, *1*(1). https://doi.org/10.2478/opli-2014-0005

Open Indiana | Indiana University Press. (1977, August 1). https://publish.iupress.indiana.edu/read/how-animals-communicate/section/79ea1292-481b-4aa0-a750-1a5d2e8f3bfa#ch37

Pagel, M. (2017). What is human language, when did it evolve and why should we care? *BMC Biology*, *15*(1). https://doi.org/10.1186/s12915-017-0405-3

Project NIM revisited. (2023, September 21). Columbia News. https://news.columbia.edu/news/chimpanzee-language-project-nim-herbert-terrace

Pearce, K. L., Port-Royal, 2015, doi:10.4324/9780415249126-DA063-2. Routledge Encyclopaedia of Philosophy, Taylor and Francis. https://www.rep.routledge.com/articles/thematic/port-royal/v-2

Rumbaugh, D. M., Gill, T., & von Glasersfeld, E. C. (1973). Reading and sentence completion by a chimpanzee (Pan). *Science*, *182*(4113), 731–733. https://doi.org/10.1126/science.182.4113.731

Schubert, Lenhart, "Computational Linguistics". *The Stanford Encyclopaedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.). https://plato.stanford.edu/archives/spr2020/entries/computational-linguistics/

Sebeok, T. A. & Umiker-Sebeok, J. (1980). *Speaking of apes*. Plenum Press.

Wayman, E. (2011, August 11). Six talking apes. *Smithsonian Magazine*. https://www.smithsonianmag.com/science-nature/six-talking-apes-48085302/

What is Natural Language Processing? | *IBM*. (n.d.). https://www.ibm.com/topics/natural-language-processing

Wynne, C. (2008). Aping language: A skeptical analysis of the evidence for nonhuman primate language. The Skeptics Society and Clive Wynne.

Yerkes, R. M., & Learned, B. W. (1925). *Chimpanzee intelligence and its vocal expressions*. http://ci.nii.ac.jp/ncid/BA17920721