European Funds Knowledge Education Development
Republic of Poland
NAWA POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE
European Union European Social Fund

Katarzyna Postrzednik-Lotko

ANALYSIS AND SYNTHESIS OF NATURAL LANGUAGES

Lublin 2021

Analysis and synthesis of natural languages

Katarzyna Postrzednik-Lotko

Analysis and synthesis of natural languages

Katarzyna Postrzednik-Lotko

Lublin 2021

Reviewers:

- Prof. SUT Aleksandra Kuzior, PhD., DSc.
- Dr. Ida Skubis

Funding: Research, preparation of materials and preparation of the textbook were carried out under the project – grant no. PPI/KAT/2019/1/00015/U/00001 "Cognitive technologies – second-cycle studies in English" and were carried under the KATAMARAN program Polish National Agency for Academic Exchange (NAWA). The program is co-financed by the European Social Fund under the Knowledge Education Development Operational Program, a non-competition project entitled "Supporting the institutional capacity of Polish universities through the creation and implementation of international study programs" implemented under Measure 3.3. Internationalization of Polish higher education, specified in the application for project funding no. POWR.03.03.00-00-PN 16/18.

The project was carried out in cooperation with the Silesian University of Technology (project leader – Poland) and the Kiev National University of Construction and Architecture (project partner – Ukraine).

Project manager: prof. SUT. Aleksandra Kuzior, PhD., Dsc.
Department of Applied Social Sciences
Faculty of Organization and Management
Silesian University of Technology

Typesetting:
Monika Maciąg

Cover designer:
Marcin Szklarczyk

ISBN 978-83-66489-54-7

Publisher:
Wydawnictwo Naukowe TYGIEL sp. z o.o.
ul. Głowackiego 35/341,
20-060 Lublin
www.wydawnictwo-tygiel.pl

Table of contents

Introduction

The purpose of this book is to provide a definition of natural language, to provide a historical overview of language synthesis and to identify the main tasks and challenges of natural language processing (NLP) based on the available literature, and to present issues related to obtaining the desired information from available databases. In addition, attention was paid to selected problems of the natural language understanding system. Then, the methods of visualization and data mining techniques were indicated, and the summary also indicated application areas (including machine translation).

Chapter 1. Definition of natural language

The term natural language is understood in two ways. First, the concept of natural language is used as a term to describe a language developed through the historical development of certain ethnic or national groups, as opposed to artificial languages. Secondly, it is a language used by humans for interpersonal communication, and the term is also used in both mathematical and technical disciplines, as opposed to formal languages such as programming languages.

Natural language processing (NLP) is an interdisciplinary field that combines issues of both artificial intelligence and linguistics. It mainly deals with the automation of analysis, understanding, translation and natural language generation by the computer. Natural language processing consists mainly in transforming the information stored in a computer database into a language that is easy to read and understand by humans. In contrast, a system that understands natural language transforms natural language samples into more formal symbols that are easier for computer programs to process.

Many NLP problems relate to both language generation and comprehension. For example, the morphological model of a sentence (the structure of words) that a computer should build is both necessary for the sentence to be understandable and at the same time grammatically correct.

Also, NLP largely overlaps with the field of computational linguistics and is often considered a sub-division of artificial intelligence. On the other hand, the term natural language is used to distinguish human languages (such as English, German or Polish) from the formal language, i.e. computer language (such as C ++, Java and others). While natural language analysis can deal with both text and pronunciation, work on speech synthesis has developed as a separate section.

Chapter 2. Synthesis of natural language[1]

[1] Chapters 3-15 are based on: Köster J.-P., *Historische Entwicklung von Syntheseapparaten mit Erzeugung statischer und vokalartiger Signale nebst Untersiuchungen zur Synthese deutscher Vokale*, [in:] von Essen O., Gutknecht Ch., Köster J.-P. (all Ed.), *Hamburger Phonetosche Beiträge – Untersuchungen zur Phonetik und Linguistik*, Vol. 4, Hamburg, 1973.

The origins of the synthesis of natural language go back to the beginnings of human history. The fact of a close relationship between song and speech and the shared use of the same carrier elements facilitates access to the theory that a primitive man who saw a singing/speech instrument in a speech apparatus must have been able to use it, an artificial instrument as a pseudo-lingual producer, as a synthesizer. Otherwise, it is not surprising that even today in primitive cultures you can find whistling, drumming and balloon "tongues". It seems that the melody and rhythm, and sometimes also the recorded changes (as evidenced by the Navaho percussion language in North America) of the language, have been transferred to elements of the instrumental "language". This stage in the development of the instrumental language, which is still visible today in the seclusion of primitive cultures (Melanesia, West Africa, some parts of South Africa and other isolated areas), can generally be used to depict the original conditions. Therefore, the earliest synthesizers should be considered drums, animal horns and Pan flutes. In fact, the natives of Liberia call these instruments "the language of the game" (the Liberian Jabo can distinguish four tones in their language). Of course, these primitive synthesizers can only be considered true synthesis machines in the broadest sense. They do not directly perform the function of generating a speech signal in its complex form, but merely take over the carrier elements with which they can convey less specific, more general information content. The synthesis signal is therefore not synthesized speech, but a substitution of speech. After all, in the communication chain they take over the function of speech organs and develop speech signals thanks to the use of elements from natural speech. In this sense, they are the true precursors of modern speech synthesizers, even though such an application cannot be demonstrated in terms of structure and design. Primitive synthesizers are of particular value to the synthesis system and its traditions through the fact that they are musical instruments. The equation of the speech apparatus and musical instruments of ancient times comes from the apparatus for the production of animal voices, which were increasingly used in antiquity and developed through the 17/18 century through the construction of organs. In the register, animal and human voices were to imitate, until today, electronic synthesis of speech. In the publications of the research group, which under the supervision of E. Leipp at the Sorbonne, developed the synthesizer of the Ikophone and is now successfully confirmed in the field of rule synthesis, the close relationship between the human speech apparatus and musical instruments is repeatedly indicated. Besides primitive cultures, primitive synthesizers are still used today as hunting bait. Bird evoking synthesis is easily accomplished with a whistle and has been eagerly practiced since ancient times. Other animal voices can also be synthesized with simple devices. Finally, the toy market is another area where primitive synths play a role today (nose whistle, bird calls, doll voices, etc.). To define a group of primitive synthesizers, we note that all the instruments of primitive man, once made with the intention of imitating speech and using language carriers, are primitive synthesizers designed to imitate animal sounds. Almost all basic forms of primitive synthesizers, whether as musical instruments, hunting baits, or toys, have survived to this day. The drum is one of the most important primitive synthesizers. It is still used today in almost all jungle areas with primitive settlements as a means of long-distance communication. Although Bantu Congolese drumming uses only two different tones, it is nevertheless suitable for printing messages such as the holding of festivals and their description, dangers, formation of enemy groups, battles, peace treaties, etc. The highly developed drum language has been proven for some tribes of Papua and people in

Africa. Pitch and rhythm play a decisive role as information carriers. The Monumbs in New Guinea have developed their percussion language to such an extent that they can convey extremely diverse information. What the drums primarily attribute to a group of primitive synthesizers is the fact that in percussion languages, the rhythm and pitch are related to the accent and refraction of the language represented. Occasionally, even in percussion languages, there remain linguistic subtleties of earlier linguistic states that have long since changed in the languages themselves. Sometimes the phonetic subtlety of speech persists in signals, although it has been lost in language. Horns, shells, and pipes, as primitive synthesizers, are almost no less important than the drum. They can be easily used to mimic a speech melody, and tuning can also be based on the accents and articulations of the language. In this particular context, reference can be made in particular to the Malaysian Pan flute. Contemporary analyzes of the whistling languages of Turkey, Africa, etc. have shown that their codes reflect a number of physiological features of a normal language. The whistle or flute has also found use in imitating the calls of birds, as simple apparatuses that mimic the calls of birds with a degree of perfection can already be found in the Byzantine Philo (late 3rd century BC) and the Alexandrian Heron (around 100 BC). Then, in the following centuries, they underwent continuous improvement and enjoyed vitality until the 18th century. Horn and shell signals, which have the same carrier properties as tubular signals, have also been used as primitive synthesizers. The horn signals of Liberian apple trees find their source in the folding of watchmakers' slogans. The Jewish shofar signals (horn signals) reflect in their global form the sung biblical text. In the recent past, short appeals for animals were stimulated by the horn tunes of pig and sheep farmers in Great Britain, Hungary, Romania and the bordering steppes. Although neither the drum, nor the horn, nor the flute produce pure tones, their energy spectra essentially have one central region that corresponds to the respective tuning of the instrument. Instruments operating on the basis of individual tones represent the simplest type of primitive synthesizers. In addition to this simple type of primitive synthesizers, primitive cultures also encountered instruments capable of producing vocal sounds, thus approaching the nature of real speech synthesizers. The special sonic effect of these instruments is based on the principle of resonance, which has been used since early scientific speech synthesis in the late 18th century to explain the process of speech formation and manage its simulation with increasing success. This theory is based on the assumption that the spectrum of the sound energy produced by the generator can be constructed by resonators such that any speech sound can be composed. The African marimba (balafon) and Javanese Génder are xylophone-like instruments, with bamboo tubes serving as resonators under their wooden pieces that act as springs. The sound-generating mechanism is based on the fact that the pieces of wood are made to vibrate by hitting the hammers, which are amplified by the resonators placed under the generators, thanks to which the vocal impression is created by a single element or by hitting several elements simultaneously. It can be assumed that these effects were originally interpreted linguistically because they are better than the melody and rhythm of individual tones as linguistic signs. Even if there is no direct evidence for such an assumption, the fact that the principle of solid marimba production is one of the oldest in the world and that it was common not only in Africa and the Javanese, but also in Asia, indirectly increases the importance of marimba (it seems the same principle was applied, also in other forms – than Javanese Génder – when building several Asian and African instruments) and the probability that the marimba simulated singing and thus the language already in primitive times, at the same time falling under primitive synthesizers.

Of the primitive synthesizers mentioned so far, the jew's harp is the most used instrument because it is an excellent source simulator – jewish harp, also known as guimbarde – requires the use of metal; its distribution area covers the whole world.

The jew's harp consists of a flexible steel tongue that is attached at the end to a tin or iron frame. The free end of the tongue is bent at a right angle and can be easily hit with a finger. To play, the instrument is placed in the mouth and held by the pressure of the teeth against the parallel ends of the frame.

The language itself produces a very deep sound (from 50 Hz to 100 Hz depending on the model). It takes on the function of a pseudo-laryngus and produces a vocoder-like language with excited sound.

In addition to producing vocal sounds, the first formant, which is highly visible on the spectrum, can be manipulated according to the melody by changing the shape of the mouth, through the movements of the tongue, jaw and mouth, so that the fluctuations are perceived by the listener as specific pitch variations that overlap with basic (similar to the language of parrots). It takes a few Jewish harps to play complete melodies because only a multiple of the basic tone of the same steel tongue can be produced. Wheatstone mentions a certain Mr. Eulenstein who used 16 Jewish harps and was therefore able to play all the keys.

A rule with the Jewish harp is a flute or a nose whistle that has even been marketed as a toy. By changing the mouth opening and buccal articulators, the mounting tube acting as a resonator is changed, and with it the emitted stroke. This is the same phenomenon as in the case of the so-called cubic flute tubes, the pitch of which depends on their volume and opening area (ocarina). The fact that this does not produce vowels – as is the case with the Jewish harp – is solely due to the fact that the sound generated by the lip whistle is quasi-sinusoidal, i.e. it cannot in any way be compared to the Jewish harp generator that produces up to 30/40 tons.

Chapter 3. The place of primitive synthesizers in the field of speech synthesis

Primitive synthesizers, as already mentioned, are not real synthesizing machines. They are not able to comprehensively recreate the final speech signal or simulate its generation mechanism. Therefore, primitive speech synthesizers have only historical significance.

The impulse to transfer language to instruments in its simplest form was most likely due to magical-religious reasons, and therefore must be understood from the cult tradition of primitive peoples. If despite the great interest in the language generation mechanism (starting with the ancient Egyptian papyri between the 22nd and 14th centuries BC), the language is still used for magical and religious reasons, a solution still cannot be found, and the 14th century BC shows that the Egyptians had one. There were already some ideas about blood, breath, and voice circulation, there was no better artificial speech language production than that achieved by primitive synthesizers, the reason must have been an imperfect objective insight into the anatomy and physiology of speech and acoustics. Of course, this should only be successful if it allows some freedom from magic and religious embellishment (unless objective observation itself becomes subject to religious demands); but this was achieved only with the beginning of the natural-philosophical direction, of which Aristotle was its principal representative.

The magical-religious direction of speech synthesis covers prehistory and antiquity, including Assyro-Babylonian, Egyptian and ancient Indian cult circles. It finds its temporary limitation in the natural-philosophical direction, starting with the Greeks, but without a complete ending, because both the Greeks and Romans, and beyond medieval attempts at synthesis to the linguistic machines of the 18th century, magic, which is very often associated with religion, was one from the essential carriers of speech synthesis.

Chapter 4. Ancient language machines

Language was of great importance as a means of influencing people as early as antiquity, when religious habits became decisive rules in great cultures. The code of primitive synthesizers was probably no longer available to the common man then, and was probably primarily a secret language for the privileged classes (priesthood, group of political leaders). Since the conditions for a true synthesis of language did not yet exist, attempts were made to establish direct contact with people by other means.

The history of speech synthesis in ancient Egypt must be seen in connection with religion. The Egyptians did not make important decisions without first consulting with the gods. The way of contacting the supernatural was different. Divine counsel was obtained through:

1. Movements and signs of gods statues.
2. Mysterious voices.
3. Noises.
4. Prophetic dreams.

It should be assumed that the communication between the believer and the statues of the gods was initially limited to gestures and the statue answered the questioner "yes"; or "not"; through moving parts. It is extremely difficult to leave the terrain of speculation here, as the insight into the practices of the rituals is hampered by the fact that the initiate was obliged to remain silent and was threatened with death if he broke this silence.

In connection with the monuments of the gods who communicated their will to believers through gestures, Macrobius tells (beginning of the 5th century CE) about the statue of Heliopolis in Lower Egypt, to which Lucian also alluded. This statue answered questions by moving forward (yes) or stepping back (no). It is known that in Etruria and Rome, sacred art also referred to movable monuments. Macrobius also notes the "talking" statues of certain goddesses of fortune that, like the figures in Heliopolis, moved forward and backward to answer questions put to them.

In the second stage, sounds and mysterious voices could take over communication functions of gestures. The voices could not be an artificially generated speech, but with a high probability a system of pipes through which (in a plastic hole) could be heard, in a modified form, what the priest said at the other end of the connection, or voices which come from the priests hidden in monuments. In this category of distorted speech synthesizers seem to match those mentioned by Chapuis and Gélis, recognized as oracles and speaking monuments to the Chaldeans, which were described by Maimonides (1135-1204), the personal physician of Sultan Saladin (1137-1193).

In addition to monuments whose speech ability derives from a distorted way, antiquity also knows monuments or other objects that may have been able to generate simple sounds. This message represented such simple signal generation, often exaggerated, and made it the ability to speak fluently.

Famous in ancient times was the ability to speak the Colossi of Memnon near Thebes. They were after Strabo (66 BC to AD 24; see "Geographica") partially destroyed by Cambyses (he took the throne in 529 BC as king of Persia, in 525 he was in the possession of Egypt and died in 522). As soon as the first rays of sunlight fell on the statue's mouth, it is said that he was able to read the Greek inscription on her left leg in plain language and even make prophecies seven verses long.

Contrary to these claims, Pausanias (who lived in the time of Hadrian, 76 to 138) reports that the Egyptian sun monument at sunrise only made a sound that would be similar to a broken string in a lyre. This opinion was shared by Strabo (66 BC to AD 24) and Juvenal (47 to 130), who spoke of the emission of one or more sounds.

There should be no doubt that stories with longer, consistent statements are legend, which over time turned into a "talking" statue. Therefore, one should rather follow Brewster, who claimed: *The simple sounds, which issued from the statue were, in the progress of time, magnified into intelligible words, and even into an oracle of seven verses* .... This notion is supported by a younger account that Sir A. Smith wrote after his travels to the Great Colossi in 1821. Smith also heard the sound of a broken lyre string and interpreted it as isolated sounds through the foot bases. Among other explanations, this as the expansion of the material under the influence of solar radiation seems to be most understandable that the dew entering the cracks at night is evaporated by the sun in the morning and escapes through one or more holes from the center of the rock to the outside hissing.

It is interesting that, as reported by Brewster, Jomard, Jollois and Devilliers, who stopped in Egypt in the 19^{th} century, they wanted to hear the same sound of a string breaking in the temple of Amun in Kanaku. Also Humbold – according to Brewster – reports that during one of his trips to the Orinoco, he heard a similar sound near the rocks.

Since one can infer a whole line of the prophecy from simple sounds – a breaking string – it is not surprising. This simple Egyptian seemed to be interpreting the priest's interpretation of how the divine message was translated.

Also outside Egypt, deceptive systems were used to transmit the sound of the images of the gods. Bruner derived in his "exhaustive description of sound generators or speaking figures" speech systems from an old discovery of the Brahmins and Chinese. With his help, priests could eavesdrop on confessions of guilt bringing sacrifices to believers and relay their messages to them. Through the legend, the message of a speaking head from the Germanic area is also conveyed. Odin, a mighty sorcerer from the north, had a speaking head set in gold[2].

In connection with the assignment of mysterious voices to the statues, it is mentioned as an instrument that cannot be assigned to primitive synthesizers. Nor was it used to receive sounds or noises, but to change the natural spoken language to give it a different, foreign tone.

It is a mirliton, which existed in almost all old cultures for voice customization. It consists of a thin diaphragm which, when put to the mouth, vibrates through speech and gives the voice a nasal, metallic sound. The example of this instrument shows with great accuracy the simplicity with which an ancient man could generate an "artificial" voice.

With the flourishing of Greek and Roman culture, the natural philosophy current – is becoming part of the history of phonetics. It is characterized by vast innovations in the area of understanding nature's processes. Apart from the works in the field of sound

---

[2] Cf. ibidem.

systematics (such as the works of Plato or Aristotle), to which the Indians showed a great approach (not to mention the legendary Theyth, an Egyptian who would be the first to systematize speech sounds and name individual elements), they are characteristic of the new current important observations on the anatomy and physiology of the speech mechanism, acoustics and philosophy of speech.

Pythagoras (582 to 500 B.C.E.) designated the discussion of whether the voice would be something corporeal or non-corporeal. This problem was then picked up by Plato and Aristotle and further observed by the Stoics.

Hippocrates (460 to 377 BC) looked at the human larynx like a flute – the strings to which the voice assigned no meaning. In his view, the sound is generated when air is forced outwards and thus made vocal strings to vibrate. The head amplifies the sound that is produced, and the tongue forms it by hitting the palate and teeth. Hippocrates also describes the lungs and trachea, although his arguments were less detailed on the basis of barely deeper anatomical and physiological knowledge. This lack is particularly related to the great problems with the acquisition of experimental material. Doctors of Hippocrates' day, who acquired a lot of anatomical skills, had only fetuses and newborns at their disposal, who had been abandoned as being perceived as offspring of little value. Since the larynx, in particular, has yet another very specific structure in early infancy, precise knowledge of anatomy and physiology in adults has developed very slowly.

Comprehensive information on the structure, shape and function of the larynx was found only in the Gauls on the basis of animal vivisection.

Plato (429 to 348 BC) mentions the diaphragm and lungs and comments on their function in the breathing process. Plato is also the first to reflect on the essence of pitch. He interprets pitch differences as a change in the speed with which the sound travels from the emitting instrument to the ear. Aristotle (384 to 322 BC) also explained the basic phenomena of anatomy and linguistic tools, even though it hardly relates to details and, like Plato, makes mistakes. Thus, he erroneously concluded, from evaluating language as the major organ in articulation, that deaf people have a stiff tongue. It was Afrodisias who first explained that the deaf are silent and cannot hear. A wealth of information on treating problems with linguistic pathology can be found in the Greeks, especially Aristotle. However, these problems are mainly solved theoretically. The work of Greek and Roman grammarians on language synthesis, on the other hand, addresses problems in the written language. Aristotle considered the physiology of breathing extensively. Our knowledge of breathing today does not begin to develop until the mid-nineteenth century. However, in the middle of the fifteenth century, Leonardo did an excellent job of the anatomy of the organs involved in breathing (see also Vesal, who in his publication of 1543 for the first time gave a description of the respiratory device corresponding to nature); however, his results went unnoticed for some 450 years. The vocal cords do not yet play any role in Aristotle's anatomy. They are only partially indicated. Only Bebricius of Aquapedente (1537-1619) recognizes the beginning of the vocal fold function. Until Aristotle, the design of the extension tube was roughly known. Aristotle's comments on pronunciation include the timbre, pitch, strength and duration of speech sounds. What is known from Pythagoras, Hippocrates, Plato and Aristotle in terms of suggestions about the structure and function of the linguistic organs was eventually systematized and deepened by Galenus (129-199). In his case, the participation of the larynx in the production of the voice is considered to be secured. Aristotle continued to use the same name for Pharynx and Larynx and assigned them the essential voice generation. Although Galen provided a very good overview of the respiratory organs (diaphragm, lungs, trachea) and larynx and interpreted them in detail in the texts, he incorrectly assigned the role of the trachea in tuning. This view was only corrected by Dodard, who rigorously eliminated the involvement of the trachea in voice production and limited its function to the supply of air from the lungs to the larynx. What distinguishes Galen above all from Aristotle is the fact that he no longer limits himself to pure observation, but instead experiments, systematically and intensively using vivisection in which he possessed astonishing virtuosity. In this way, he was able to develop previously unimaginable knowledge in the field of anatomy and physiology.

The use of acoustic vases in ancient times indicates that builders of amphitheaters and large conference rooms have already used the effects of resonators to improve room acoustics. In the fifth book "De Architectura" (written around 13 BC) Vitruvius describes the use of air vases, the volume of which was adjusted to the size of the respective system. According to Tarnócza, the vases embedded in the wall are resonators that dampen in closed rooms, but strengthen in open structures. The assumption that significant new knowledge in the field of medicine and the natural sciences, as well as philosophy, would also give speech synthesis new impulses by which it could break out of the field of primitive synthesizers or false speech synthesis is not correct. In natural philosophy and visionary epochs there is no approach to the true solution of the problem of synthesis. As in the Assyrian-Babylonian, Indian, and Egyptian regions, the predictions of the Greek and Roman oracles are imitated by false voices. Since at least Greco-Roman times, people have dealt with pronunciation through statues and strange voices. Gods and oracles often expressed their statements verbally. In order to surprise believers and influence them, the priests of the Greeks and Romans do not refrain from contacting the gods with the help of tongues. They usually did this by speaking statues, wonderful voices and oracles. Such statues were found in the ruins of Pompeii, which had pipe systems through which they could speak using a pipe system with a hidden speaker. One of the most famous statues is the talking head of Lesbos. She was known in Greece and Persia and reportedly foretold the death of Cyrus in the land of the Scythians. Through Lucius, we know that Alexander, by transmitting his voice to the statue - this method was general - could himself hide behind it and make his voice. The fact that the Greeks and Romans donated statues was due, among other things, to the fact that art here was able to perfectly recreate the human body. It is obvious that when faced with such situations, all the so-called "talking" statues of antiquity must be met with extraordinary skepticism. This is especially true for reports of the legendary times of the Greeks. They talk about characters that can move and talk, but must be seen as metaphors right from the start. Evaluation of the operation of Roman and Greek antiquities in the field of speech synthesis may be limited to the conclusion that, despite the decisive advances in medicine, philosophy and the natural sciences and the resulting insight into the mechanism of language production, no verifiable real attempts have been made. As in the millennia before this era, primitive synthesizers still exist, but they probably have one problem, and that is that directly perceived linguistic information has lost its meaning. The use of false speech systems has been improved. The unconditional rejection of ancient language systems as true speech synthesis machines and their evaluation as a false tool for priests and lay rulers, if not fictional, is shared by all antiquity speech synthesis authors. In most cases, this assessment also covers the entire Middle Ages. Despite the negative balance of speech synthesis, which results from the evaluation of the Greco-Roman era, some positive development in the field of animal voice synthesis has been noted. This is especially true of the artificial production of bird sounds, which can be seen in a multitude of models of Philo of Byzantium and Heron of Alexandria. Philo of Byzantium lived at the end of the third century BC and was known as a Greek tactician and engineer. He is counted among the representatives of the Alexandrian school. His traditional writings include that he reported an automaton that could be placed in temples along with a suitable spring. This is the form of an animal that drank water with sucking and splashing sounds. The first construction of a bird powered by an artificial mechanism is also attributed to Philo. The apparatus used air pressure or steam as a driving force.

In summary, primitive synthesizers existed thousands of years before this era, but have probably lost their importance as carriers of immediately perceived linguistic information. The use of false speech systems has been improved. The unconditional rejection of ancient language systems as real-language machines and their evaluation as a deceptive tool for priests and lay rulers, if not fictional, is shared by all authors who deal with the synthesis of ancient speech. In most cases, this assessment also extends to the entire Middle Ages.

Despite the negative speech synthesis balance that results from the evaluation of the Greco-Roman era, there has been some positive development in the field of animal voice synthesis. This is especially true of the artificial production of bird sounds.

Heron lived about 100 years before Christ was born and was considered the perfect mathematician and mechanic of his time. He has preserved excerpts from some writings on the production of various machines. As the inventor of sucks and fountains, he also used compressed air, squeezed by whistles, to synthesize bird voices. He also managed to use pipes in combination with water. These early bird voice synthesizer models made by Philo and Heron survived the 17th and 18th centuries.

Chapter 5. Middle Ages – experimental requirements for the construction of language machines

With the fall of the Western Roman Empire, antiquity came to an end. The beginning of the Middle Ages marks the time of the decline of phonetics. In terms of the perception of vocal and sound processes in language production, the new era is characterized by a relapse to a level that remains well below the level achieved by Galen. Galen's teachings are forgotten. None of the medieval anatomical works produces results in a phonetic relationship that can be supported by research.

Due to his critical scientific attitude, Kempelen generally uses it as the most outstanding representative of earlier speech synthesis. But Kempelen could not completely ignore his time's tendency to mysticism. He took this into account when he delivered his machine, the description of which was only published seven years after its completion.

To demonstrate the fascination with the artificial language of amazed people in the 18th century, a brief mention is made of an anonymous journalist who, in a letter to the "Teutsche Merkur" of 1784, showed a Kempelen machine that had not yet been completed, described as follows: "You cannot believe, dear friend, because of the strange feeling that the first hearing of a human voice and a human tongue, which was clearly not from a human mouth, made us all amazed. We looked at each other in silence and concern, then he confessed frankly that at first a little secret shower would have flown over us. For now, I can tell you about this strange invention, which maybe a hundred years ago could have brought Kempeln to the stake, dear friend, he himself, when he is finished, give us your thoughts on these linguistic organs, the discoveries and inventions of this extremely important machine in a separate work that he wants to publish about it"[3].

Finally, Kempelen has definitely adapted to the sensation of its peers thanks to the construction of the machine. But it was precisely because of this "scientific error" that the machine was received with great skepticism (for example, when the language machine was released in 1783, an anonymous pamphlet shows: "The talking machine [...] was intended to provide evidence that the machine was an invalid automatic machine".

It was only in the Treaty of 1791 that Kempelen abandoned any secrecy and described the development of his synthesis apparatus with scientific objectivity for an already perfect and clearly defined sale. "I can flatter myself so much without exaggerating self-love that although it is imperfect, it can at least be a good reason for a perfect speech machine"[4].

At the beginning of his description of a speech machine, Kempelen confessed: "... it never occurred to me to work on such a machine. When I started trying, I apparently wanted to imitate some sounds of the human voice with some instrument; […] Which I found very difficult and I thought completely that it was impossible to combine them with pronunciation, in fact I even had the most important sounds or letters in detail. For many years, I have only been able to put them together from a distance, and thus be able to create syllables and words. How I came up with this idea very gradually and very late

[3] Anon., *Űber Herrn Kempelens Schachspieler und Sprechmaschin,* [in:] "Der Teutsche Merkur", November 1784, pp. 180-185.

[4] Kempelen W. v., *Wolfgangs von Kempelen k. k. wirklichen Hofraths Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine/ Le mécanisme de la parole, suivi de la description d' une machine parlante et enichi de XXVII planches*, Vienne 1791, preamble.

will turn out: it is possible to create a versatile machine"[5]. Kempelen therefore took a long time to develop the most perfect language machine model and is characterized by several distinct stages:

- 1st stage of experimental preparatory work;
- 2nd stage of the first machine (isolated sounds);
- 3rd stage of the second machine (isolated sounds combined with mechanical control);
- 4th stage of the third and final talking machine.

This treatise reflects Kempelen's knowledge not only of philosophy and the study of language, but also of the physiology and anatomy of the speech organs, which were developed in the 18th century. However, Kempelen was not familiar with the observations described by Leonardo da Vinci in his Quaderni d 'anatomia and consequently did not use them in his deliberations. On the other hand, it is clear from Kempelen's writing that "Memoire sur les causes de la voix de l'homme et de ses diffrents tonnes "by French physiologist Denis Dodard had an effect (from which Kempelen recognized that the trachea had nothing to do with the patient's voice , except for the delivery of air). Kempelen clearly recognized the importance of the anatomy and physiology of the speech organs for vocal articulation: "... therefore the most necessary information should be provided here about the structure, mood and function of each tool that contributes to speech. "Among the main tools used in the production of speech, Kempelen lists: 1. glottis 2. nose with soft palate 3. mouth 4. tongue 5 teeth 6. lip among the tools that produce voice: 1. lung 2. trachea 3. larynx 4. glottis.

They are basically descriptive elements that are also used today in articulation phonetics. Since Kempelen's time, the basic functions of the Extension Tube have been recognized. He overcomes the theory that the larynx has the primary function of producing speech and the secondary function of the extension tube. E. Brücke's appreciation of the physiology of the speech apparatus shows what aspects of Kempelen's thoughts on the physiology of the speech apparatus he represents. Brücke writes that Kempelen's physiological phonology "[…] was so well established that it has given and will continue to be the surest foundation for all further research. His work on the mechanism of human language is one of the best physiological books I have ever read [...]"[6]. Kempelen's achievements in the field of the physiology and anatomy of the speech organs are characterized by systematic experiments aimed at elucidating the mechanism of speech tools. Descriptions are comprehensive, theoretical conclusions conclusive. Kempelen never stops at the level of pure anatomical-physiological observation, but subordinates his observations to the goal of constructing mechanical analogues. He actually understands the lungs as the respiratory organ that guarantees the supply of air necessary for the phonation process. The mechanical model he used for this purpose is a bellows, such as that used by Kratzenstein and Mical. In the wind tracheal assessment, Kempelen clearly stands out from the wrong opinion that he has a special and necessary voice share and agrees with Dodart's opinion that he only has the function of delivering air from the lungs to the tuning organ, the larynx. Kempelen identifies the larynx as a voice socket.

---

[5] Ibidem, pp. 388-389.
[6] Brücke E., *Grundzüge der Physiologie und Systematik der Sprachlaute*, Vienna 1856, p. 7.

As a voice-generating mechanism, he exactly mentions the glottis. As early as 1769, Kempelen was working with various musical instruments to overcome the difficulties of producing the human voice. This explains Kempelen's view that vocal folds acted in layers like membrane reeds. There are also other reasons why the mechanism of vocal folding has not been properly identified. Since the vocal folds could not yet be observed from the outside, it depended on examining the heads of the dead. However, these studies only allow the observation of the vocal folds in a top view, so that only horizontal movements can be recorded. This led to the erroneous conclusion that the vocal folds are horizontally stretched membranes. In addition to the erroneous model as the target of consideration, it is incorrect to assume that these phenomena can be recorded alive in the larynx of the deceased, because the larynx of the deceased is a completely passive model that never allows simulation of muscle contractions of the living vocal folds, and a few hours after death due to loss of moisture changes its reaction to blowing. Based on a synthesis of the Ferryin and Dodard theory, the larynx is a stringed instrument. Kempelen adheres to the model of the membranous reed (Kempelen finds both theories compatible and developed an attempt at synthesis: "Since there cannot be any changes in the glottis, that is, it cannot become wider or narrower without stretching or stretching its edges, and vice versa, never it must not be stretched more or less unless the glottis also becomes narrower or wider at the same time"[7]). Kempelen's phonation theory is based on this model: "this opening is not so large that the air pressed against the tongue can escape completely freely, but it must be forced through with some force, rubs against both edges of the membranes and causes them to tremble. The tremor is so remarkable that the jolts, when air is blown through the skin back and forth, almost merge into one and are no longer distinguishable by their ears, just as the spokes of a speed wheel seem to converge into a dial. In this way, the repeated gusts of air become sound to the ear, and this sound is called a voice"[8]. If Kempelen, despite his theoretical attachment to this model, uses other sources in his linguistic machines, it is only a sign that the artificial glottis, which has been theoretically misused, has not proved successful in practice. Von Kempelen clearly talks about the relationship between the voice generator and the articulation apparatus: "[…] the mouth is two vocal strings; they indicate the sound, the rest of the structure of the instrument serves, like the mouth in the voice, to sharpen the sound, round it and, if I may express myself that way, articulate it"[9]. Von Kempelen also noted the whispering tongue in the side note: "I can speak quite audibly with the naked wind on my speech device if I put a small piece of wood into the tuning tube, thus preventing vibration".

Kempelen's approach to solving the problem of artificial language production is clearly genetic. With this, then, he is faced with a theoretical concept of solution, standing alongside his contemporaries Kratzenstein and Mical. He is able to distinguish from them by a practical result which, in his last machine, brings him closer to the genetic solution than his opponents. In his last machine, Kempelen is able to generate all the vowels from one resonator for the first time. It achieves this goal through a mainly empirical working method. Both theoretically and technically, Kempelen represented

[7] Kempelen E., 1791, p. 83.
[8] Ibidem, p. 84.
[9] Ibidem.

a new type of researcher who was able to defend a general bias towards the natural mechanism of speech and reject without concessions all conventional considerations that did not correspond to this concept. It follows that Kempelen could not have a direct practical and technical relationship with any of its predecessors.

In order to create a language of a general kind, the development of its machine was not only in the fact that the ultimate task was to articulate a specific language or to speak only the European language, but it should be realized to such an extent that it would be able to master all the languages of the world.

On the one hand, the goals set in this way contrasted with the critical remarks made in the scientific field about the limitations of the results achieved and the admission that the research was not exhaustive: "What is the theory or the mechanism of human language, which is the largest part of this book, I am far from believing that I have exhausted everything in this knowledge. I only show the discoveries that I have made in my experiments, I will arrange them in a specific order, I extract consequences and rules from them, I try to report on the errors I have discovered in various writers"[10].

The relationship between theory and model so clearly evident in Kempelen's work still does not seem to be the ultimate goal of his deliberations, but has rather stabilized intuitively or can be explained as an inevitable product of his way of working. Indeed, the purpose of Kempelen's efforts was to technically create a talking machine; Kempelen needed a theory to make his technical ideas come true faster. His theory in turn, which led him from the scientific observation of the previous speech act in clamping application to – if only prompted – analogous mechanical functions, guarantees his machine an undeniable first place in a series of linguistic machines that simulate a clamping tube, although it ultimately does not represent it. In this regard, Kempelen clearly distinguishes itself from Kratzenstein and Mical, who, in lengthy experimental work, constructed five complex containers that were not analogous to the articulation organs, but only because of their resonant behavior with the sound of the five vowels they were supposed to represent.

Kempelen undoubtedly contributes to a better understanding of the function of the articulation organs and, indirectly, the acoustic relations more clearly than his predecessors and their principles in systematically following the development of the sophisticated mechanics of his machine. In this way, it recognizes the three-fold division of the speech generation mechanism (air generator, voice generator, articulator) and takes them into account when designing the machine. He also understood the fundamental importance of the connection of the larynx with the cavities and distinguished between the primary and secondary functions of the organs involved in articulation: "[...] everything agrees with the proof that apart from the glottis, there is no proper tool for the tongue and that our nose, mouth, tongue , teeth and lips were originally little used for speaking like fingers for playing the flute and eyes for reading, although they were indispensable tools [...]"[11].

[10] Ibidem.
[11] Ibidem, p. 177.

Nevertheless, Kempelen's views are not without flaws. The use of the membranous tube as an image of the human glottis clearly shows that there is a fundamental misinterpretation of the natural voice generator. The example of a voice generator also shows how much technical practice has placed limitations on the application of its theoretical insights to the designer.

It is based on the experience of the first attempts with a wind instrument and selects muffled props that are sufficiently justified. This inconsistency can only be explained by a restriction of the technical possibilities. None of the models derived in the physiological part from articulation factors can find application in synthetic machines.

Nonetheless, Panconcelli-Calzia and Scripture consider Kempelen to be a researcher who has succeeded in producing synthetic sounds in a purely genetic manner. However, this view should not be completely accepted as the Kempelen machine as an articulation analog does not withstand rigorous testing. Ultimately, it is a model that, as a mechanism, only roughly mimics the physiological processes involved in speaking, and is also a terminal analog based on the final acoustic score.

Brekle shows that this is not necessarily a disadvantage: Kempelen's main advancement is actually in this dual modeling, since pure articulation modeling gives almost no scientific information, it merely imitates the observed articulation process. Simultaneous acoustic modeling, however, highlights the relevant properties of the extension tube and thus provides scientific information on the speaking process, it is also more generic and thus more explainable, as it can describe articulation processes in terms of general acoustics[12].

The final abstraction of the acoustic phenomena that occur in speaking in this articulation process has not yet been done by Kempelen.

Here he uses the experience of the first preliminary experiments with wind instruments and decides to use a muffled reed. This is a clear and insufficiently justified inconsistency. This inconsistency can only be explained by the compulsion of technical possibilities. None of the physiological parts derived from the articulation conditions are used in the synthesis machine.

Even so, Kempelen sees himself as an investigator who has managed to produce synthetic sounds by purely genetic means. This view cannot be fully supported, however, as the Kempelen machine does not stand up to rigorous testing as an articulation analog. Ultimately, it is a model which, as a mechanism only roughly imitates the physiological processes of a speech mechanism, is also a terminal analog based on a final acoustic residue.

However, this does not necessarily have to be a disadvantage: Kemplen's significant advance is actually in this dual modeling, since pure articulation modeling gives almost no scientific information, it is merely an imitation of the observed articulation process, but simultaneous acoustic modeling has important properties and thus informative scientific about the speaking process, has greater generalities and at the same time greater explanatory capacity.

12 Cf. ibidem, p. 36.

The final abstraction of the acoustic phenomena occurring in the speech act precisely from this process of articulation has not yet been completed in Kepelen's work (he could have used a cylindrical tube closed on one side to acoustically explain the vocal articulation). However, it is on the way to achieving this goal because it simulates articulation with materials whose nature and function are not comparable to the corresponding parts of the articulation apparatus. Access to the theory of acoustics of speech sounds in Kempelen's work can be found in the explanation of how speech tools work.

Even though Kempelen could not yet make the transition from articulation phonetics to acoustic phonetics, he nevertheless gave it an initial impetus. Willis, who laid the foundations for the acoustic theory of vocal articulation by measuring the acoustic properties of circular cylindrical tubes of different lengths and diameters, refers to Kempelen for his decisive contribution to the development of synthesis. He further develops Kempelen's idea of the relationship between sound and tone quality and abstraction - from the mouthpiece of a speech machine – to physiological conditions in humans[13]. The final step towards unconditional synthesis is therefore reserved for Willis.

Thanks to Kempelen's speech machines, sounds, syllables and words can be produced by the demonstrator. This means that it is not a talking automat as it would also need to be mechanized from the control part. However, since the Kempelen machine was used to generate synthetic speech via feedback by interrogating a demonstrator, coupled with the dexterity of making changes via direct control, it cannot be a question of wanting to deal with the machine in the sense of automatically managing the entire communication process. The functions that the device cannot perform are part of the neurophysiological transmitter, part of the auditory part, and part of the neurophysiological receiver. This means that it is not possible to save the rules of the synthesis process, which can be stored and ensure reproducible synthesis at any time. This is a significant drawback as it also blurs the problem of accurately generating sound transitions that has no chance of being brought up to the scientific question level. It remains veiled in the user's intuition. The copy of Liénard's machine and his experiments with it show very precisely the qualitative weaknesses of the apparatus: "p", "l", "m", "f", "r" and "h" cannot be produced as independent sounds, "m" almost it does not differ from "n", and "v" hardly differs from "f";

A certain number of sounds must be compensated by "p" ("d", "t", "k", "b", "g") and "s", "z", "ç" and "j" are not produced through the mouth and therefore do not combine well with the vowels. Liénard's research on generating vowels with a copy of the Kempelen synthesizer shows that only 'a' could be generated clearly, while the remaining vowels are only imperfectly produced with difficulty. Spectrographic analyzes show that synthetic vowels were characterized by a region of stronger intensity around the root note, which served as the first format (the root note was at the height of the child's voice in the case of Kampelen) and a more or less significant energy range between 2000 Hz and the root note, which can be considered as the second format. This imperfect acoustic structure is apparently sufficient for contrasting identification, as Kempelen himself admitted that his vowels became properly understood only in contrast. Kempelen was, of

[13] Cf. Willis R., *On the vowel sounds, and on reed organ pipes*, Cambridge, 1830, p. 231.

course, also aware of the fact that this limited achievement consisted largely of promoting the perception of artificially articulated language through prosodic means. Therefore, during the demonstration, he paid attention to the exact imitation of the accent of the words and sentences produced.

Moreover, Kempelen apparently used the means of self-suggestion consciously. Many authors reporting on the machine and related demonstrations emphasize that the machine often repeated the same words or imitating words that the listener had previously requested (i.e. spoken). The fact that the power of the machine may be minimal under such conditions seems understandable to those familiar with the role of suggestion in acoustic perception. Liénard also emphasizes this fact: "On est surtout aisément induit en erreur quand on sait d'avance le mot que la machine doit dire, et lorsqu'elle a prononcé on s'imagine d'avoir entendu ce que l'on a voulu entendre"[14].

Another trick to pretend an acceptable quality of artificial speech was the high level of the reed whistle, which gave the effect of a child's voice. On the one hand, the energy spectrum of a child's voice is imprecisely defined, even with natural sound production, and on the other hand, the listener does not intuitively place great demands on the quality of the baby's gibberish and is pleased to roughly hear what is expected.

The decisive advantages of the machine in relation to the previously constructed models is taking into consideration the importance of connecting people with each other and its value as the basis for the development of the speech articulation theory, which, however, would only develop in later years. For the first time, she clearly brought the role of contrast and the influence of suggestion into the discussion.

Another advantage was the spontaneity of creating an artificial language, which was lacking in many electronic machines (OVE II, PAT, Pattern Playback). However, the quality of these machines, with an average intelligibility of 98%, clearly stands out from the capabilities of Kempelen machines. Nevertheless, the distance between the Kempelen synthesizer and the first electronic model for generating continuous artificial language, Voder von Dudley (1939), is not misleading. Both machines have specific advantages and disadvantages:

1. Voder uses electrical simulation, thanks to which the acoustic properties are well defined and predictable.
2. The Voder simulates acoustics only, the Kempelen machine simulates acoustically and articulatingly.
3. The Voder makes better sounds, the Kempelen machine has allowed for better expressions on linguistic production.
4. They both need a trained user, he only disappears with the invention of the spectrograph during World War II, which led to the development of an automatically controlled spectrogram reader[15].

It is worth noting that experimental phonetics has its roots from Kempelen, i.e. before the development of the comparative language method, and that not only the first works on language physiology, but also phonic tables were the subject of research and questions

[14] Liénard J.S., *La machineparlante de Kempelen*, 1968, p. 11.
[15] Koenig W. et al., *The sound spectrograph*, JASA 17, 1946, pp. 19-49.

throughout the 19th century. Enormous also unreservedly recognizes Kempelen's merits as a pioneer of phonetics and speech synthesis. Without being mathematically and physically pre-trained, Kempelen developed his famous speech machine on this basis after experiments and careful observation, in which articulation processes reveal themselves as purely acoustic processes. This successful vocal synthesis attracted the attention of some 19th-century scientists. The acoustic function of the tube was investigated in the second half of the 19th century, which allowed for physiological considerations on the activity of the uvula, soft palate, vocal folds, etc. in the background. The path was clear to a deeper look at the phenomenon of vowel articulation.

Willis, Weatstone, Helmholtz, Bell and others use the work of Kempelen in the 19th century. The entire physiological and genetic synthesis of the nineteenth century radiates on Posch and Faber, who developed new language machines on this basis. This is true up to the Riesz model from 1932.

With the triumvirate of Kratzenstein, Mical and Kempelen, speech synthesis developed in the 18th century. It was the first real breakthrough in the field of mechanical speech synthesis in the 19th century. From then on, one can almost continuously follow the stages of development of scientific efforts in the field of artificial language production. Speech synthesis was no longer a field of alchemical hoax, but rather the sphere of interest of talented researchers. All three designers started from the idea of mimicking a speech mechanism by mechanical means and developed theoretical concepts of varying value for this purpose. However, none of them succeeded in putting the theoretical, basic idea into practice.

Hermann von Helmholtz (1821-1894), a student of Johannes Müller, is a member of the Virchow, Brücke and du Bois-Reymond group. It should therefore be expected that genetic-based synthesis would be further developed by him. This assumption is not correct, however, as Helmholtz synthesizers are to – under certain limitations – be treated as a simulation of acoustic vocal articulation processes and thus be unconditionally subject to the genetic concept. The field of genetic synthesis found its last representative in Faber and was to lie fallow until the times of Riesz. The career of Helmholtz, who was initially a professor of anatomy and physiology and only later (1871-1894) took up the chair of physics at the University of Berlin, shows strong similarities to the Kratzensteins.

The preparatory work for the theory of sound experiences took place in 1855-1863, when Helmholtz developed elements of his vocal theory. They are based on the Wheatstone resonance theory and Fourier's findings that any periodic sound process can be decomposed into the sum of sinusoidal and cosine harmonic oscillations.

Helmholtz defines a simple tone as a periodic movement of air that repeats in the same way at even and sufficiently small intervals. With this definition, Helmholtz followed Ohm, who might again be referring to Fourier. At the same time, he was guided by Ohm's principle of defining sound.

Sound quality was determined in 1843 by Ohm using the following laws:

1. all musical sounds are periodic;
2. the sound quality depends on the right combination and number of simple tones;
3. any change in air pressure corresponding to a complex musical sound can be decomposed into the sum of its sinusoidal waves.

According to Ohm, the sound consists of:

1. basic tone;
2. partial tones (the fundamental tone is the first partial tone).

Helmholtz extended Ohm's law by the relative strength of partial tones and the exclusion of any influence of this phase:

1. Music sound quality depends on:

   a) numbers,
   b) the relative strength of the partial tones,

2. Component phase is not essential.

Following Ohm's clean-tone ideas, Helmholtz also adopted the views of Ohms' opponent Seebeck, which contradicted Ohm's or Helmholtz's definition that simple sound could also be caused by air movements that deviated significantly from simple swing movements. Helmholtz exposed the controversy as a fictional issue when Seebeck referred to sounds, but Ohm wanted to define simple sinusoidal tones.

According to Helmholtz's hearing theory, sound, which is made up of individual periodic vibrations of air, is still processed separately for a period in physico-acoustic perception, while the psychological result of physico-acoustic perception leads to a fusion of periodic vibrations into sound.

Helmholtz, making a distinction between the fundamental tones (= the lowest tone contained in the sound) and the subtext (a series of harmonic multiples of the fundamental tones), turned to the old knowledge of the laws between the numbers of oscillations (already Pythagoras, 582-479 BC determined the conditions for oscillations on the strings of a monochord; Galileo, 1638; Newton; Euler, 1729 and D. Bernouilli, 1771, have come to the abstraction of these laws) generally applied to any sound, and agreed with the precursors of the theory of speech and sound associated with subtext (Bartoli / Rameau / Seiler / Garcia; Helmholtz himself). He also noted that the vowel sung for piano strings is preserved due to the different strings and after their subsequent vibrations), while for researchers such as Wallis (1677) and Sauveur (1701), who noticed the existence of supersonic in musical instruments, they can be considered perfectionists in in the sense that they generalized their laws and consciously linked them to the sounds of speech. Helmholtz managed to prove the general validity of the law both theoretically and empirically.

Unlike sound, which is defined as the sum of sinusoidal oscillations in a harmonic relation, Helmholtz triggers a sound that consists of nothing – periodic parts.

Thanks to its sound specification, Helmholtz is able to create a distinctive sound description using the relative energy of the given tones. Since all the oscillations which do not correspond to a simple pendulum motion are broken down into a number of simple tones in the perception of the ear, sounds with a different timbre and hence without a fundamental tone for the ear must be different by the different forces of harmonic supersonic.

1. The sound is full when the fundamental tone is stronger than the supersonic.
2. The sound is empty when the basic sound is weaker than the supersonic.

3. The sound is empty if it only has strange supersonic sounds.
4. A sound is nasal when it has a large amount of strange supersonic noise.

In order to empirically test his theoretical concept, Helmholtz had to subject the most varied sounds to a precise analysis. Various methods were available for this purpose:

1. ear analysis with resonators;
2. phonograph;
3. manometric flame.

Other methods became known only later and probably have not yet played a role for Helmholtz in deciding which analytical method to use:

1. Blondel's oscillograph (In 1893 Blondel returns to Bell's 1876 phone and Hughes' microphone transmission from 1878).
2. phonograph (Edison, 1877).
3. graphophone (Bell and Tainter).
4. graphophone (= recording of the graphophone on the zinc).
5. phonodeik (Miller, 1909).

Helmholtz decided to analyze complex shells with resonators, which Rayleigh later did as well. So Helmholtz used Mittonne's sonic analysis. Its resonators are empty air-enclosing bodies that are extremely sensitive to resonance.

Helmholtz, who initially worked with cylindrical resonators, was the first to experimentally prove the theory of sound in this way, which until now had only been treated mathematically. Helmholtz took the second important step towards speech synthesis by extending the analytical study of the overtones and overtones of vibrating bodies to the sound of speech produced by the human voice, in the spirit of his theory that all sounds are subject to the same laws.

However, Helmholtz described the isolation of parts of the human voice as relatively the most difficult to isolate. For this reason, he conducted experimental studies on the spherical shape of the resonator, which had optimal resonance sensitivity and seemed to be the most suitable for speech sound analysis. The best shape for these resonators are glass spheres with two holes, one of which ends in a very short funnel-shaped neck, the end of which fits over the ear canal. The resonance of these spherical hollow bodies is determined by their contents and the degree of opening.

By using ball resonators, Helmholtz was able to use his early analysis to compare whispered vowels, in which he compared the pitch heard in the mouth adjusted to the vowel, with a standard frequency.

The use of spherical resonators that create only one resonance and their principle of resonance implies the model of the extension tube as a spherical resonance chamber system.

The most important results of these studies include:

1. The human voice is so rich in supersonic that it is easy to hear them up to 12. The male emu's voice assigns Helmholtz to 16 supersonic.

2. The strength of supersonic determines the sound of a vowel, and also individually determines the overall tonal community: strong supersonic produces a sharp, bright voice, weak supersonic produces a soft, dull voice.
3. There are vowels with single (a, ɔ, o, u) and double resonance (i, e, ae, ø, y).

Double resonances made it necessary to revise the idea of a single resonant space responsible for vowels. The theoretical solution was obvious. As the vowels of the front tongue show a pronounced elevation in the front and middle of the back of the tongue, which can also be shown for ø and y, and since most of them assume a closed or half-open position according to the degree of jaw opening, the idea of presenting these sounds with an extension tube should not be viewed as a simple spherical resonator, but as a system of two coupled resonators of the same type. The first resonance should arise in the cavity behind the tongue elevation, the second through the narrowing of the middle part of the tongue and the hard palate. Strictly speaking, this coupled system should be considered to be composed of a front cylindrical resonator and a spherical coupled resonator behind it.

While, of course, Helmholtz could not argue for an arrangement of one or two conjugated, strictly spherical resonance spaces, he firmly clung to this basic idea because such cavities only have one resonance at a time. It was his closed vocal articulation system. He was no longer able to interpret the empirically obtained data substantially differently. Helmholzt's use of the Willis cylinder model would have given him the opportunity to model a tube closed on one side as the basis for a new vowel theology a century earlier. However, it was reduced to a spherical resonator.

It was natural that Helmholtz, who in his theory of sound (and thus also the vocal theory) stated that sound consisted of a finite number of sinusoidal tones, which seemed to be confirmed by analytical experiments, also took the opposite path to analysis as the basis of his theory for synthesis.

For Helmholtz, synthesizing sound could only mean combining a series of pure sinusoidal tones into a harmonious union. He produced sinusoidal sounds (because otherwise only blown-up bulbous bottles produce a quasi-clear sound) by tuning the resonators to the fundamental frequency of the voices and placing a tuning fork before opening the corresponding resonators. The clean base came out clearly, while the upper tones of the tuning fork were muted and remained unheard of to the base.

The integration of synthesizer models that are on the border of genetic and passive-genematic synthesis is practically not subject to inconsistencies, since the models can be understood as a function of appropriate reasoning.

Chapter 6. The last phase of mechanical synthesizers

In the pursuit of the further path of mechanical synthesis, until it is completely replaced by electrical and electronic synthesizers, development can be sketched according to the division into genetic groups and active and passive-genematic synthesizers. This allows you to use all the equipment you provide, if you do not lack all information about their principle of operation, for example in the Kessel apparatus mentioned in Gutzmann's "Sprachheilkunde" (it is said that the synthesis machine was shown in Munich in 1899, where Kessel showed all the vowels and a series of words).

In 1872, Alexander Graham Bell completed his language machine, inspired by the imitation of a Kempelen stem. He presented a good imitation of the natural speech apparatus with the trachea, larynx, tongue, soft palate, teeth and mouth, however, his mobility was not comparable to that of a Faber machine.

A year later, Potter described his experiments with the rubber model of the extension tube, which he had enriched with tiny freely swinging tongues. Despite the simplicity of this principle (spherical rubber cavity), this model offered for the first time in history the possibility of an artificial extension tube as a whole filter system to optimally express extreme freedom of movement based on natural conditions. The influence of articulators during the operation of speech on the quality of the emitted artificial signal was more visible than in the case of Faber. However, Potter did not draw any obligatory consequences from his attempts, neither in terms of resonance theory nor in any other theoretical sense. He was content to create a limited set of vowels and diphthongs: Λ, i, u, ͽ, ai, ei. On the other hand, physiologist Landois developed various models of synthesizers from castings of a clamping tube (with different vowel positions) in which he built a striking ivory tongue as an artificial larynx. Landois is said to have produced all the vowels using these models. Landois also showed a common weakness of purely genetic synthesizers: he contented himself with an optimal imitation of the genetic apparatus, without attempting to theoretically record the associated articulation or acoustic laws. Marage's synthesis experiments, in turn, did not show this deficiency. He developed a theory of the vocal articulation of wine – albeit incorrect – in connection with his models of genetic synthesis. In 1897, Marage perfected the first experimentally tested Doumer method of photographic recording of manometric vibrations of flames. The intense preoccupation with the flame images, which created a special pattern for each vowel, gave Marage the idea that the vowel sound had to be developed in the vocal folds (other representatives of this theory: Eijkman, Natier et al.).

Therefore, when designing his synthesis machine "Sirene a voyelles et reonateurs buccaiux", he focused in particular on the study of the optimal source form. Since he was involved with the mermaid, he was concerned with determining the shape and constellation of holes that, in his opinion, should be characteristic of a vowel. His experiments were based on Koenig's manometric forms of flames. Moreover, he referred to a specific glottic opening (narrow triangle) because, according to his ideas, the glottic opening formed the physiological basis of the characteristic vowel sound. The vocal folds are reached in the vowel e, while u and o have the shape of a triangle. Eventually he used isosceles triangles for "u", "o" and "a" whose sharpest angle was towards the center of the disk. He used a narrow gap for "e" and "i". For "u" and "i" he arranged the holes equally, for "o" and "e" in groups of two, for "a" in groups of three. If this theory were correct that the source is responsible for quality, Marage should consider the mermaid a suitable synthesis machine. It did not bother him that he was able to produce vowels only from the base of the siren (for a few sounds accidentally with reservations). With indecisive and bizarre arguments that are not worth mentioning here, he put forward the theory of the need for an extension tube, for which he made a series of models from casts of a natural extension tube when pronouncing various vowels. As with Landois, this mechanism produced intelligible vowels, although theoretically quite different bases from those announced by Marage.

Together with Russell and Paget, genetic speech synthesis has found its way back on track. With admirable persistence, Russell searched among 160 different models of reed for a mechanical solution that showed satisfactory compatibility with a natural generator. But he didn't find this solution. He turned to the material for the synthesizer with the same meticulousness and decided that in order to produce natural sounds it was necessary to have extremely soft tube walls. However, Russell's most significant contribution to speech synthesis is the fact that he categorically abandoned the idea of a conjugated resonator system. In his opinion, it is not individual depressions, but the entire space above the vocal fold that should be the basis of the theory of vocal articulation. To prove his views, Russell imitated vowels using sound models based on three-dimensional X-ray images of the lamp, thus creating easily understood vowels. As Russell was satisfied with the acoustically satisfactory result of his synthesis experiments, which he considered to confirm the rejection of the tube approximation by the coupled Helmholtz resonator, Paget reserved the right to derive general rights from his models. Russell's idealism was that each vowel should have a specific and unique configuration. So he rejected any form of articulation compensation. Russell strongly pointed out the danger of self-hypnosis in the production and interpretation of artificial sounds. We hear what we listen to. He demanded that all artificially generated signals be checked by intelligibility tests with neutral persons. Following this idea, Russell published a vowel perception test in 1931, which was synthesized in accordance with Paget's statement of 1923. From 1922, Paget was concerned with the problem of speech synthesis. Marking the resonance points of the tube during the articulation of whispered vowels and consonants led him to imitate a tube with indentations made of wood, rubber, cardboard and kneaded plastic. These models can still be viewed today at the Royal Institution in London. Paget first generated the frequencies of the first and second formants, which he determined with great accuracy in whispered speech, in plastic models according to the principle of trial and error, and depending on the resonance frequencies, he also obtained the desired vowels. In determining the resonance behavior of spherical hollow bodies, Paget found that the resonance height depends in the same way on the size of the opening as it does on the content of the resonators. He described the width as irrelevant because it did not affect the nature of the sound, but only the magnitude of the vibration amplitude: the wider the resonator, the greater the amplitude. In the Helmholtz sense, he saw the reaction of the vibrating potential air in the resonator as a function of the ratio of the laryngeal signal to the potential natural frequency. With optimal correspondence, the response and duration were maximum. Like Helmholtz, Paget saw the tube as a system of two coupled resonators, each with a specific resonant frequency at which the source develops a subtext. These resonators are changed by the position of the tongue and lips. Apparently Paget did not mind that some of his resonators could hardly be considered a system of two coupled resonance chambers, but rather a tube with a variable cross-section closed on one side. In this way he also introduced, without hesitation, the technique of hand imitating the articulation (which he mastered perfectly) as based on the Helmholtz principle – i.e. the resonator. Paget himself used a rubber diaphragm as a sound generator, or produced a fundamental tone using his mouth. He was able to generate the following sentences:

1. Hello, London, are you there?
2. Oh, Leila, I love you!
3. A happy New Year!
4. I hope you have all enjoyed the paper.
5. Easy there, you-re on the nerve!

The main goal of Paget's synthesis research was first and foremost to prove that all vowels have two formants, and only then to answer the second question about how they arose or what the relationships between the model geometry and the acoustic behavior of the resonator system are. For this reason, most studies have undertaken numerous experiments that either isolate the given two formants involved from the vowels, or allow the corresponding vowel to be identified from the two resonators. Paget was categorically against the results of Miller and Stewart, which preceded his synthesis experiments and in which one formant was considered sufficient for some vowels. In doing so, he argued wisely and with great skill. Paget tried to explain that Miller assumed only one resonance for some vowels, saying that at the resonant frequency n, the laryngeal frequency must not exceed 1/2 n because otherwise the resonance point is at the fundamental frequency and differentiation is not possible – even with Fourier analysis. This is the reason why all high "a's" show one resonance, but all show two resonances. Paget also proved in direct experiments that it is not possible to unambiguously identify vocally with one formant (e.g. Miller's ɔ = 781 Hz; u = 383 Hz). Developing his thesis about the necessity of using several formats for each vowel, Paget demanded that each formant for each vowel should be in a specific position and that it should be different from the position of the same formant in every other vowel: it is an obvious mistake to use the same upper or lower resonance for two or three different vowels, where by changing the spacing we can give each vowel an upper and a lower resonance that would be characteristic of that particular vowel:

> *It is an obvious mistake to use the same upper or lower resonance for two or three different vowels when, by a process of re-spacing, we might give each vowel an upper and lower resonance range which would be, both, characteristic of that particular vowel*[16].

Paget's error is symptomatic of his entire concept of work. Too valuable, subordinated to the idea of the resonance space theory, it failed, despite the interesting discoveries found, to capture complex right-wing laws and acoustics. In the case of W Nasal, Paget made an exception to the requirement that all vowels be complemented by resonances, the coupling of the nasal cavities would introduce other additional resonators into play. Overall, they produced a total of four resonances in each nasal vowel. In the experiments with the model unit policy, they were also not tuned to the articulation conditions in a systematic manner, which was intended to re-submit plastic models as long as prices are available and resonance could be generated ; Paget found some rules. For example, he found that the displacement of the modeled language within the cavity is critical to

[16] Paget S.R.S., *Human speech*, London 1930, p.72.

the movement of the formants. In detail, it meant that when the shift position was shifted forward, the first and second mold layers shifted towards each other, , the second time position shifted towards each other, the second time position shifted forward, the first position and the second layer of the mold shift towards each other. This is well suited to the articulation and acoustic conditions. Paget also imagined that he clarified the question of whether the form (round, ovoid, cylindrical) or its conjugation (serial, parallel) influences formation. From his experiments, he concluded that neither a certain form nor a certain coupling has a direct relationship with the later form, but that the timbre in the period depends solely on the actual values of the resonances to be through the pits. This insight has indirectly led to the recognition of compensation rights as it leaves the extension tube free to produce the same resonances with optimal articulation settings and to obtain acoustic performance data. Paget presented the simplest proof of his thesis through the interchangeability of series-connected resonance spaces.

In many cases, the experimental results are limited to a modest inventory of some vowels and diptychs. Continuous speech has not been developed, nor does it have good quality natural vocabulary. The root problem is unsolved, the theoretical foundations are wrong. It is the balance of 2000 years of mechanical speech synthesis. The only value is that it leaves the basis for an electrical hypothesis from which more insightful theses can be developed.

Mechanical synthesizers were developed in the 20th century. In the 19th century, they were replaced by electrical devices, and later electronic ones, which are in some respects more precise and have fewer restrictions. Nevertheless, even in the age of electronics, they have not yet lost all their rationale. They are still used in experiments today to shed light on certain functions of the speech mechanism, which is essentially a mechanical device. Mechanical synthesizers are very valuable in illuminating unbalanced sound effects that are difficult to analyze mathematically.

Chapter 7. Electric and electronic synthesizers

With the help of Riesz and Miller, mechanical speech synthesis was completed. The 20th century was stronger than the 19th century. There were many more interested in powerful machines that could take over the function of a person and facilitate his work. More attention was paid not only to the efficiency of the machines, but above all to how quickly they could do it. As the mechanics broke down at high speeds, better and better synthesizers were introduced, which focused more and more on help. After sixty years of rapid development of electrical and electronic synthesis, it has finally been possible to bring about high-level discoveries in the field of artificial speech, so that nowadays such speech is not inferior to human speech.

Recent successes in the field of research on speech synthesis in connection with its development from the decline of mechanical speech synthesis to the present day, the ability to point to its deficiencies, as well as the contribution to phonetic research, especially in the field of voice articulation, are the task of the further course of the work presented here.

The enormous increase in importance, which clearly distinguishes modern speech synthesis from mechanical synthesis, and the acceleration of its development, which has become a fact after replacing mechanical synthesis, results from the role that non-mechanical synthesis could take as a problem-solving tool in various fields of science, thanks to its ability to be applied well beyond the possibilities of mechanical synthesis. Compared to purely mechanical sound and loudness generation, electronic methods have the advantage of eliminating almost all restrictions on the range of tones, loudness, sound character and, in the case of musical instruments, playability. Only the introduction of electrical devices made it possible to transform any form of vibration into sound.

Speech synthesis is of particular interest to phonetics, even if until 1949, in the narrower sense of the word, access to phonetic sciences as an integral part of it was simply necessary. Until then, experimental phonetics, if scientifically operated at all, relied heavily on analytical tools. A description of the role of synthesis in phonetics in relation to the first generation of speech synthesizers can be found in brief outlines in Fischer-Jorgensen (1958) and Fant (1957), who discuss the functioning of the second generation of speech synthesizers in terms of their phonetics. With the beginning of the 1950s, the tasks of synthesis in phonetics clearly crystallized, and today the synthesizer is a central research instrument in many phonetic laboratories, the value of which for heuristic foundations in phonetic sciences is indisputable.

The use of synthesis machines in the context of phonetic-linguistic research concerns:

1. Controlling hypotheses concerning the mechanism and functions of the speech production apparatus.
2. Supplementary function within the analysis or precise analysis of the speech signal (synthesis analysis).
3. Systematic testing of parameters.
4. Development of perceptual regularities.
5. Demonstration of articulation and acoustic processes by manipulating the sound of speech in language teaching (teaching a foreign language) and in teaching speech and hearing to people with hearing impairments.

Due to the task of synthesis to control hypotheses about the mechanism and function of the speech production apparatus, it can be applied back to mechanical synthesis. Here at the Hermann and Helmholtz school, the main function of synthesis was to test theoretical concepts. To this day, this task is one of the most important tasks of synthesis: The primary goal of research on synthetic speech is to thoroughly understand the process of speech production. However, it is imperative that the synthesis machine mimics natural language as much as possible in this task. If one is satisfied with a quality similar to natural language, there is a danger that theoretically irrational ideas will be considered proven. Whether the use of artificial language research results can be unconditionally transferred to the natural language level has not yet been fully clarified.

In addition, Fant, however, indicated that the listener's reaction to artificial speech is the same as the reaction to natural speech.

He himself believed that the following should be done first:

1. Phonetic analysis of a given language.
2. Detect the socio-linguistic background of the speaker.
3. Find out as much as possible about the personal characteristics of the speaker.

It follows that an artificial language can be exchanged without problems for a natural language in experiments.

Another point of view is that the slightest deviation from the level of natural quality causes large changes in the perceptive area, so examination with artificial stimuli only allows conclusions to be drawn about artificial speech. Translating these results into the natural language is unacceptable.

Both opinions are consistent with each other, it is only necessary to make a sharp distinction between what they refer to in detail. Fant's judgment is based on an artificial stimulus that has a quasi-natural quality. It states that its stimulus not only meets all physiological requirements, but is a vehicle for the sociolinguistic and personality differentiating features of speech sound. Under these conditions, it is difficult to assume that the other party would not find such a stimulus fully operational in an attempt to solve the problems of natural language because it is not produced by the natural speech apparatus. On the other hand, it cannot be assumed that Fant would be responsible for the results of studies with extremely bad vowels as sufficient to be transferred to natural conditions. The limitation of the value of experiments with artificial stimuli as binding to natural conditions depends solely on the quality and nature (there have certainly been trials where bad quality may be desired) of the result together.

Also, synthesizers were used in flight simulators to simulate noise in all phases of flight. Strictly speaking, in the context of this work, only problems related to the synthesis of speech are of interest, insofar as they are related to the synthesis of natural language.

Chapter 8. The beginnings of speech synthesis

The special capabilities of electrical and electronic simulation have given synthesis a new status as an integrated part of phonetic research. It must be admitted that the increase in the recognition of a synthesis in the phonetic sphere is accompanied by the reference to data from modern analytical devices, which is more and more noticeable in phonetic research. The first great success of these devices was the discovery of more than two formants for each vowel. After Obata and Teshima were the first to discover the third form with electrical analysis devices, since vowels usually have four formants, although in the vowels analyzed he could also prove five or six. Lewis finally set five resonance frequencies for each vowel in 1936.

In 1906, Cahill built a new type of organ without pipes, based on an electric system, and successfully demonstrated it to the public. For the first time, the use of electrical means in sound production has been successfully demonstrated. Jaensch first transferred the electrical principle to speech synthesis in 1913 by replacing the mechanical sources commonly used in synthesizers with electric light.

The selenium cell underwent periodic changes in light due to the rotating profiled disks, resulting in resistance fluctuations that could be heard as sounds in the telephone receiver. But although the mechanical-acoustic phenomena were also transferred into the electrical realm, the basic principle remained the same as with mechanical synthesis: the inertia of the selenium cell generated an electrical oscillation composed of quasi-sinusoidal exposure variations produced complex electrical vibrations that showed the number of harmonic dominant tones. By converting the electric vibrations in the telephone handset, additional subtones were added, creating a full spectrum of energy, the nature of which depended on the type of exposure fluctuation.

No new knowledge of the vowel structure can be expected from this apparatus, since its structure obscured important factors rather than made them accessible. Moreover, due to Jaensch's theoretical starting point, there was no basis for a systematic synthesis of vowels.

After the still imperfect electrical apparatus of Jaensch and Lachmund, Stewart in 1922 provided unambiguous evidence that the vowel synthesis path used by mechanical synthesis can also be described by electrical means using the composition of harmonics. Stewart was the first to apply the modern circular theory to synthesis technology. He published a description of his theoretical considerations and practical experiments in his essay "Electric analog of vocal organs" (Stewart, 1922). The Stewart synthesizer was seen as an imitation of the mechanical-acoustic double resonator model created by Helmholtz with electric means, which Paget was honing at the same time on original foundations.

Analogous to the whistle on the tongue or vocal folds, Stewart generated a tonal alternating voltage by periodically charging the capacitor with a battery via a rotating switch (whispering vowels were generated by non-periodic gaps at a frequency higher than the frequency of the resonant circuits). All subtones have been muffled by two parallel formant wheels (extension tube function) made of adjustable resonance filters with dampers that do not fit in both passbands, and the air bubbles from the vocal cords stimulate air vibrations in the mouth. The vowel of two forms thus formed was then audible through the headphones.

According to Stewart, he synthesized all vowels and half-vowels with a 50% intelligibility using his synthesizer. This result was to be surpassed only by Dudley Voder. The prototype of the formant synthesizer was created with an apparatus that is analogous to the acoustic principle of a speech-generating mechanism that creates the source signal by filtering the final signal.

If the research of Jaensch and Lachmund, which could still be used in a relatively imperfect electrical synthesis apparatus, is ignored, Stewart's experiments are the earliest attempts to operate on electric speech devices based on new considerations about the essential properties of speech sound. In experiments with his device, Stewart determined the importance of the parameters underlying his system. Although he recognized the vowel sound as being independent of the fundamental frequency, he made the frequency of the formant circles and the intensity of the formants very important. Stewart and Miller answered the question about the number of formants involved in vowels using a single resonance characterized by a single sequence of recurring damped oscillations for "u", "ɔ" and "a", as well as for "i", "e" and "a" double resonance characteristic of two sequences of recurring damped oscillations.

Limited to generating static vowels, Stewart's machine remained within the disadvantages of mechanical models. Furthermore, Stewart did not attempt to connect his resonance filter in series. At about the same time as Stewart, Eccles conducted electrical synthesis experiments with two resonators connected in parallel. In 1923, he and Wagstaffe presented their equipment to the Royal Institution at a lecture titled "Demonstration of the Electric Method of Producing Vowel Sounds and Its Application to Wireless Telegraphy" (Eccles and Wagstaffe, 1924).

Eccles and Wagstaffe Stewart applied their theoretical approach. Only the mixing of components around the end signal was solved differently: while Stewart was simultaneously transmitting the output power of both resonant circuits through the loudspeaker, the Eccles switch acted on the transmitter so that the upper and lower resonances were alternately sent.

Plastic models with cross-sectional areas designated for individual vowels were installed. Scientists say that such artificially produced vowels were of good quality.

At the end of his Theory of Speech Production, Fant demonstrated the decidability of his theory by experimenting with models of a one-sided closed cylindrical tube and a double resonator. The models were excited by a condenser microphone (Fant, 1960). Ungeheuer (1962) also validated his theoretical prerequisites for the configuration of plastered cast pipes which he excited with an electric source.

Homer Dudley's Voder occupies a special position among the early devices for electrical speech synthesis. Based on knowledge from the first four decades of the twentieth century, which showed that it is possible to produce sounds and short words by electrical means, Dudley sought a synthesis solution from limitation to static sounds. The first such synthesizer which could produce connected speech was Voder, developed by Homer Dudley at Bell Telephone Laboratories and demonstrated at the 1939 New York World's Fair. Controlled by a keyboard and foot pedals, it was played like a musical instrument by operators who required a long training period.

If the Voder goes back directly to the Wagner synthesizer, it is directly related to the voice-controlled synthesizer, the vocoder, developed by Dudley in 1936. The Voder is therefore to be considered as the reproducing part of the vocoder in which the manual and pedal controls replace the filter outputs.

The device consisted of three main parts: the source, the resonant part, and the radiation part. As sources, one could choose a pulse generator (oscillating oscillator) or a noise generator (white noise), mixing was possible. The selection was made using a lever. The height of the pulse generator could be controlled by a pedal, which guaranteed a comfortable modulation of intonation in generating artificial speech.

Source signals were fed to a set of filters of ten successive bandpass filters, which covered the frequency range from 300 to 3300 Hz. The outputs of the ten filters were controlled by ten potentiometers that corresponded to ten buttons (there were also connected buttons, among other things; the 11^{th} key was used to reduce the volume by 20 phones for unstressed sounds). Filters released by the buttons supplied the appropriate energy to the amplifier that was connected to the loudspeaker.

The overall structure of the device thus simulated the natural process of speech generation (this device produced vowels and consonants in a manner closely resembling their production in the human mouth, the operator could create phonemes similar to those produced by the human voice channel). The vocal folds (tilt oscillator) generated a sonorous source spectrum that obtained its final shape in the clamping tube through the filter. The output spectrum was on the lips and emitted to the nose (loudspeaker). In voiceless speech generation (fricative, voiceless, whispering consonants), the noise source assumed the function of a natural source signal that was generated in the glottis (aspiration, whisper) or in a constriction in the mounting tube. There were additional filters for generating voiced and voiceless blasting sounds.

Of course, the analogy to natural speech generation was not complete, since in the model the filter effect for each formant usually consisted of several filters, while under natural conditions it functions as a whole. Moreover, the frequency range could only be changed every 300 Hz. The continuity necessary to understand speech should also be assigned to a user who has been trained for a year. It was an integral part of all equipment.

However, the acceptance of Voder by the public shouldn't have been particularly exciting as the quality of the generated language still had some drawbacks. Voder was not yet fully able to cope with the enormous difficulties of all dynamic language machines, the coupling of the individual elements together to ensure continuous speech. The technical equipment did not yet allow the necessary transitions. After all, Voder deserved to be the first electric synthesizer machine capable of producing dynamic speech. Thanks to this, it was possible to examine a number of old (optimal source, optimal filtering part) and new (dynamic laws) questions in the field of phonetics and, for example, also include the role of the fundamental tone, intensity and dynamic change of formants in speech. This benefit was also unqualifiedly recognized by most researchers. A manually controlled speech generator, such as a vocoder, is essential for basic phonetic research. It represented a standard source and made it possible to produce speech sounds with a specific spectral composition at any time with any accuracy.

Chapter 9. Schott Spectrogram Reader (Voback)

Through a device constructed by Schott (1948), it was possible for the first time to convert a spectrogram back to a spoken language.

Synthesis was no longer such a short-lived process, but pre-planned and repetitive.

For the main representatives of the group of spectrogram readers, the Schott model, due to its structure, can only be considered a precursor, as it has not yet applied the harmonic source spectrum directly to the samples. The device operated rather on the principle of a Voder and a Vocoder Dudley, which spread the entire spectrum over a small number of channels (bandpass filters), the energy of which was controlled independently of each other.

Usually, at least in the lower ranges, more harmonics fall into the same channel. Thus was the angular resolution of the Schott spectrogram reader to which later synthesizers of this group were subject.

In its approximate division of the spectrum, Schott's machine corresponded to the analytical spectrographic process of his time, which also performed extensive analyzes, registering the corresponding energy in higher frequency ranges through the brightness of the light bulbs. The Schott apparatus, however, did not rule out the use of spectrograms from later sonographs, because the accuracy of the spectrographic resolution could not be adversely affected in any way.

Schott used a linear light source, which he placed parallel to the spectrogram's frequency axis (inside the instrument) to illuminate the variable density spectrographic pattern (on photosensitive or electrochemically sensitive paper). Twelve photocells (except as indicated) arranged continuously throughout the pattern developed amplitude control signals for a group of bandpass filters such as Voder and vocoder control mechanisms. Voiced or unvoiced information about the sound selection and pitch have been included in additional tracks. This principle was not to record sonic vibrations, as was the case in Cooper's later spectrogram readers, but to record the strength of the different direct currents of the individual channels. This information was used in the generational part to impose some form of energy spectrum on the source signal which was variable in frequency but otherwise passive.

The main disadvantage of the Schott spectrogram reader was the approximate resolution of the speech spectrum (12 frequency ranges), which inevitably led to poor quality. Its great advantage over other spectrogram readers was the base tone modulation, which was a prerequisite for the study of problems in intonation and accent studies. The steady key tone was the substantial lack of later spectrogram readers.

Chapter 10. Cooper spectrogram reader

The first model of this spectrogram reader was developed in 1949 by Cooper. It was based on a simulation of a Fourier series speech wave. There are basically two types of Cooper spectrogram readers that differ from each other by different control principles. The oldest machine used as a template spectrograms (here = sonagrams) of special spectrographs (here = sonagrams), which show the results of the analysis on the footage. The foil negatives allowed the application of a simple principle of transmission of the generated light spots through the original. Of course, this system was only suitable for experiments in which the aim was to reproduce the results as faithfully as possible, and not to introduce multiple changes. The transmitted light energy was collected on the other side of the negative foil in a light collector and then used for synthesis. Although the originals could be retouched or produced entirely artificially, such procedures were a waste of time. About a dozen words could be synthesized in an hour in this way. Therefore, the Haskins laboratories soon switched to a different control principle, thanks to which several words could be prepared for synthesis in just eight minutes.

Chapter 11. The structure of the standard apparatus

In the long run, another principle of spectrogram reader control dominated, which was more accessible for subsequent amendments. With this control system, the standard form of a spectrogram reader is constructed as described below.

A spectrographic pattern applied with white paint to a transparent support (plastic band), which represents the intensity of speech sound as a function of time and frequency according to the type of short time spectrogram, is illuminated on the frequency axis by 50 adjacent light points.

Each of these points of light is modulated in its light intensity depending on its position in the spectral image at a specific, constant frequency harmoniously tuned to the fundamental tone. This modulation is done with a so-called "tone wheel". The wheel (perforated disc) has 50 concentric rows of variable density holes that are evenly spaced from axis to edge. The innermost ring has four holes, the next one is eight, the third is twelve, etc. The innermost ring has 200 holes and rotates at 1800 rpm. so that the fundamental frequency is 120 Hz. Harmonics show a linear distance of 20 Hz among themselves up to a maximum of six hundred Hz; they therefore represent a multiple of the fundamental frequency as integers. The device does not generate disharmony.

If the tone wheel is illuminated on the one hand in a slit-shaped cutout with a strong light source, on the other hand, 50 beams of light are generated at a predetermined frequency according to the rows of holes that catch the passing patterns through the mirror and illuminate them in the slit-shaped cutout. The illuminated areas of the patterns are reflected in a device in which a row of photocells corresponding to 50 light points is placed. Upon receiving light, these photocells transform the fluctuations of the light points into a corresponding electrical signal, which is mixed with signals from other photocells and, after amplification, emitted through the loudspeaker.

The instrument was built primarily for the production of vocal sounds as it can only work on the harmonic spectrum. Nevertheless, unbilled sounds can be simulated – even very imperfectly - by non-periodic modulation of the time / intensity of the frequency components, inconsistently recording patterns but combining them with dots and dashes.

Since the speed of rotation of the wheel is constant, it is not possible to change the pitch. Variations in intensity are limited to a limited extent by the width and complexity of the applied material.

The Cooper spectrogram reader was basically an optical matrix reader because it synthesizes speech from a simplified matrix that was applied to a plastic tape with a white color, analogous to a previously made sonagram.

Chapter 12. Efficiency of the spectrogram reader

In addition to a number of advantages that make this apparatus interesting for phonetic and linguistic research, the Cooper's spectrogram reader has a number of significant disadvantages.

Already in the description of the technical structure it was noted that the essential tone is constant. This was a significant limitation compared to natural language, in which there are no longer statements with a permanent fundamental tone. In addition to the limitation on the apparatus that it would never be able to generate the natural quality of artificially produced speech, this also made it impossible to use it as a research instrument in almost all fields, especially in the phonology related to sound. The spectrogram reader was thus able to perform its task of extracting the parameters of the various information levels of the speech sound and describing their form only to a limited extent. Moreover, the total intensity was not controllable at all and the intensity of the individual elements within the spectrum could only be insufficiently controlled, which further limits the scope of the device and strengthens the reservation of the theoretically achievable optimum quality.

While the intelligibility of language obtained with the spectrogram reader is said to be 90% for isolated sounds and 95% for continuous speech, the quality (compared to natural) was rough and metallic. The voiceless friction measures were indicated only because there was no noise generator and the patterns for the friction measures had to be compiled from random parts.

Chapter 13. Evaluation of analogues

The design of terminal analog synthesizers was based on a well-established basis with theories of vocal articulation and the technical possibilities of electronics. The overall construction (source, fixing tube, emitting) was based on the natural production of speech sounds and thus guarantees a system that, from the outer frame, had all the conditions for a good simulation of natural conditions.

The imperfections in the artificially generated language, which have not yet been eliminated, can therefore be found less in the general concept than in the details. Sufficient data was available for the realization of those details which are essential to natural quality. However, their respect for the technical system usually required a lot of effort, which in many cases was small in relation to the acoustic contribution, because non-linear or non-periodic phenomena were difficult to implement in electronic systems. The fact that the inclusion of analytically proven details was not carried out with greater force was often due to the fact that the economic structure of the artificial language was more interesting than its naturalness, which beyond a certain level seemed to be achievable only by investing more resources.

In combination with later electronic terminal synthesizers, there remains a need for a better source approximation which has an impact on energy distribution across the spectrum. The more precise shape of the impulse, which had to be controlled depending on the decisive information, such as basal height and volume, and time fluctuations in successive periods were the basic requirements of natural quality (for this, it was necessary to take into account the subliminal and supraglottic effects on the laryngeal wave: the fact that the source and the first approximation tube could be used independently of each other did not mean that such a simplification could be afforded by a sophisticated system.)

In control areas, more attention should be paid to banding and its dynamic changes, as well as to the different levels of the poles. The use of fixed parameters could only be approximate.

Smith (1965) estimated, for example, fluctuations in the frequency of the first formants to values of up to 30 Hz. In high-quality synthesis apparatus, such fine structures had to be taken into account. Also, the noise components in some formants should become automatic for vowels, making the source of the noise necessary even for synthesizers that are only used to produce vocal sounds.

Thanks to the proposed improvements, terminal analogs could be used, in addition to their tasks in the field of basic research, also to assess acoustic structures as carriers of linguistic and extra-linguistic information, and to assess the perception of acoustic stimuli and synthesize rules, which also opened the field for didactic use.

In 1970, Reynolds and Strong showed students how to make parts of the speech framework visible by systematically removing, adding, and distorting artificial speech stimulus parameters.

Chapter 14. Speech synthesis with computer systems

In the 1960s, the simulation of technical systems on computers (beginning: 1958 – David et al.) was started to be used also in the field of speech synthesis, after it turned out to be extremely fruitful for:

1. Extraction of perceptually important properties for better transmission and storage of speech sound signal.
2. The acceleration of the pace of the syllable or its reduction.
3. Removal of distorted speech sound signals or artificial distortions.
4. Automatic recognition of speech and speakers.

Thanks to the possibility of storing large amounts of data and arbitrary manipulation of them by programs, electronic computing systems are able to simulate all types of analytical and synthesizing equipment as a whole.

The use of a computer as a speech synthesis research tool also makes sense from a purely economic point of view, as theoretical concepts can be transferred to practical systems in a short time and refined in electronic simulation before allowing for costly hardware implementation. A computer can simulate a device that would take years to design in a very short time.

The link between the acoustic signals and the digital computing unit was initially established with analog-to-digital converters, while the reverse was used for digital-to-analog converters. Since manipulating data with programs was sometimes very difficult and required the use of specialists, there were options to solve this problem by developing special compilers. Training to work with such compilers was usually limited to a few hours. Later compiler models allowed for the direct programming of extremely complex processes.

In the case of full simulation of synthesizers on digital computers, a subroutine was configured for each analog model function (filtering, rectification, etc.). For example, the program described the resonant circuit during point-by-point simulation by calling the appropriate subroutines one after the other. The output signal resulting from the calculation is a series of numbers that could be converted into a quantized analog signal.

The computation time depended on the complexity of the simulated systems. Later, simulating the vocoder on larger computers took a few minutes for every second of speech.

Experience in simulating a synthesis apparatus has shown that it must be possible to control the program in real time in order to directly improve the computational process. It became possible thanks to multi-program computers. Newer I / O devices in cathode ray tubes offered good direct control capabilities through information visualization and direct intervention with a "magic pen" (Strong, 1966).

For the first time, Kelly and Gerstman (1961) showed relief in the importance of computer simulation of synthesizers with regard to the application of phonetic rules and the direct transfer and improvement of control signals in synthesizers. Their program was still very simple. Each sound has been divided into three memory units:

1. initial phase;
2. stationary part;
3. final phase.

However, the synthesis of a reasonably intelligible language required many additional ad hoc transitions.

A number of other synthesis systems of all types were then simulated on computers and used in the context of synthesis studies. However, since most of them have been used specifically to solve current speech problems, you can limit yourself to a few examples. Libermann's research (1959) has continued since the control synthesis which began with Ingemann's work (1957). The wealth of data that had to be manipulated simultaneously showed a significant improvement in a given area, taking over the control function for wire and terminal analogs or simulating an entire synthesis apparatus.

In 1964, Öhmann carried out tests with a computer-simulated cable analogue with physiological conditions of interval explosive sounds in Swedish and English. He assumed that in both Swedish and English, the phrase in the string V + Stop + V was not a sequence of three articulation gestures, but rather an explosive configuration based on the transitional part of a diphthong-like vowel. Based on this idea, the author designed a numerical model 1 that describes the V – C – V complex with a small number of control parameters related to the articulation of the tongue.

Henke simulated the dynamic analog tract model for the synthesis of vowels and explosive sounds in 1966. From phonemic data, he constantly created extension tube configurations in the midsagittal section. The model worked with:

1. states = data owned;
2. goals = target values to which the variables specifying the preceding sound must operate after they are met.

In 1969, Ruiz and Mermelstein described a simulation of an analog-to-wire synthesizer from 12 heterogeneous sections of a transmission line with parallel nasal coupling. The control concerned a discrete sequence of information on articulation, excitation parameters and time parameters. Movement between articulation information was achieved by linear interpolation of the elements that define each section. The speed was 40 times real-time, which enabled on-line synthesis and data modification in sentence sequences. The control program made it possible to quickly and easily correct the articulation information. It was possible to enter information about the section graphically and numerically.

Among digitally simulated analogs of terminals, De Clerk et al. described a formant synthesizer. The required frequency range for the control signals is only 25 Hz for this model.

The series-connected serial analog of Rabiner (1967) markedly influenced the development of rule synthesis in anatomy and physiology, which developed strongly in the late 1960s. Klatt (1970) simulated a series-connected five-pole analog clamp with which he developed a series of rules for generating English explosive sounds in the initial position. Due to the active noise, it was possible to use a five-wire fractional bank control to determine whether these efforts were based on genetics (coming from the construction of the instrument) or genetics (coming from medicine). If one of these poles was removed and replaced with zeros according to the pivot point of the generated consonant, very good explosive consonants could be produced. The source of the aspiration is said to have been of high quality. It does not identify rare sources or draw them from other sources. The basics were available for both approaches. But the question was, did it make sense? However, this should already be understood to mean that it was a decisive breakthrough in the production of language machines.

A look back at historical development seems to be consistent with both Kratzenstein and Mical or Kempelen: from their achievements, the anatomy and physiology of the extension tube were deduced. In the previous chapters, it was shown that in order to search for the origins of speech synthesis, it is necessary to go back to the beginning of human culture. Of course, the first primitive synthesizers interpreted synthetic instruments. It was not until the nineteenth century that there was a clear separation between gennematics and genetic synthesis (although Hooke's 17th-century gennemic experiments had no real counter-concept value and only little to do with the later continuous artificial language production apparatus). After the stage of systematic and empirical development of synthesizers, the time has come for theoretical and scientific researchers, thanks to whom linguistic sounds were produced mainly from unspoken signals. It is not possible to provide evidence of the systematic development that took place in the 18th century designs, nor even of a scientific effort to develop early primitive synthesizers. Their function as speech-like signal synthesizers may become less significant with progressive separation of singing and speaking, music and speech synthesis. The combination of theory and experimental review has reached the stage of modern scientific methodology.

Forced (Helmholtz) vibrations were basically closely related to each other, as they are based on the phenomenon of resonance, establish the theories of resonance as well as explain areas of acoustics and extension tube. They replace part of the acoustic signal that defines natural vowels in their own way. Since the natural signal is rich in redundant elements, the limited Helmholtz approximation of the vowel spectrum by two resonances is sufficient for an acceptable quality of artificial vowels, and Hermann's theory, which allows only one resonance (although in most experiments the spectrum must have had more), serves to achieve language in mouth.

Subsequent analog systems continued to successfully disseminate the wealth of information that had to be achieved in order to achieve the target and combined into an overall control process. While Hermann's experiments based on passive-gennematic synthesis would never achieve natural vowel quality, Helmholtz's active gennematics method would rather have this option.

Achieving the goal of absolutely natural quality would mean an immeasurable increase in data, and at the same time would lead to a complete failure of the speech synthesis models and their mathematical treatment. The absence of an imperfect source with respect to the frequency level could not have had an effect on this in the case of Helmholtz, as this feature is determined by the harmony energy of the artificial language in the given systems. The first language syntheses were established primarily for English – British, Angelic – American, Swedish, Russian and Japanese. Work on improving the synthesis of other natural languages is still ongoing. In discussing the problems associated with the production of artificial speech, it has become clear in recent years that synthesis is theoretically capable of achieving a high level of naturalness in the artificial signal. Naturalness has finally been realized, which was mainly aimed at linguistic intelligibility.

Chapter 15. Language analysis[17]

[17] Chapters 16-18 are based on: Gardner T., *Hauptstömungen der modernen Linguistik*, Göttingen 1973.

The mainstream linguistics discussed here should only be understood as the most important theoretical approaches to generative grammar. Of course, the full great structuralist movement has not yet been reached, and even today there are other significant theoretical directions in research, such as stratification grammar. Structuralism, for example, in foreign language teaching is almost the absolute master of the situation, even if the grammar of transformation has gained many foundations, mainly due to the growing support of language learning theorists. And other grammars, such as the stratification grammar that some famous representatives may reserve for themselves, have proved viable over time. Structuralism, however, cannot be considered the mainstream of our time, and other interesting non-generative theories (or quasi-generative theories have not been internationally accepted). Rome, to which all subsequent paths of transformative generative grammar lead in spite of increasing diversions and many critical constructions, is known from earlier works by Noam Chomsky. With these works, Chomsky established his standard theory and presented its theoretical motivation. The 1957 "Syntactic Structures" is certainly one of the most cited linguistic works of all time (though not the most widely read linguistic works of all time). However, in some places a short and almost straightforward booklet does not always reveal its exceptionally solid theoretical foundations to the layman, since even experienced linguists often skip or at least do not read it.

But it is these earlier theoretical foundations that constitute the generative principle, the wavelength of modern linguistics. It is clear that the grammar of transformation was not created traditionally and suddenly out of anything scientific. Considering the developments in the sister disciplines, it is almost possible to say that their development can still be expected. In retrospect, very early approaches are discovered that seem to be based on generally similar linguistic perspectives. Chomsky writes in Cartesian Linguistics, his omnium gatherum on the philosophy of language:

> *[…] it seems that after a long interruption, linguistics and cognitive psychology are now turning their attention to approaches to the study of language structure and mental processes which in part originated and in part were revitalized in the "century of genius" and which were fruitfully developed until well into the nineteenth century. The creative aspect of language use is once again a central concern of linguistics, and the theories of universal grammar that were outlined in the seventeenth and eighteenth centuries have been revived and elaborated in the theory of transformational grammar. […] I have tried to indicate, in this summary of Cartesian linguistics and the theory of mind from which it arose, that much of what is coming to light in this work was foreshadowed or even explicitly formulated in earlier and now largely forgotten studies. […] it would be quite accurate to describe current work as a continuation of the tradition of Cartesian linguistics and the psychology that underlies it […]*[18].

[18] Chomsky A.N., *Cartesian Linguistics*, New York 1966, pp. 72-73.

Thus, if Descartes, Locke, and Leibniz are cited sequentially on one and the same subject, the case must be extremely general or it must be quite superficial in various philosophical systems. Bertrand Russell meticulously argued with Cartesian philosophy:

> *Cartesian Dualism was rejected by Malebranche, Leibniz, Berkeley, Hegel and William James. I cannot think of any philosophers of repute who accept it in the present day, except Marxists and Catholic theologians, who are compelled to be old fashioned by the rigidities of their respective creeds*[19].

Either way, Chomsky's compelling psychological remarks about the innate premise of language learning are not directly influenced by innate learning based on traditional language criticism. Descartes, Locke, Leibniz, and many others became more embraced by such a predisposition, regardless of its fundamentally different philosophical conception. Whatever Cartesian linguistics should be, it cannot be said that syntactic structures can in some way become merely a continuation of them. The basic (generative) principle obviously comes from a different philosophy, namely from an analytical point of view. A certain model of acquisition with all its linguistic and linguistic-philosophical implications arises as a consequence in the broadest sense of this generative principle, and not the other way around. If Chomsky absolutely wanted a non-analytical philosopher as patron, he didn't need to go back to Descart. Edmund Husserl, who developed Cartesian philosophy, directly offers a theory about the whole and parts, and this, for example, would translate into a grammatical analysis of the component structure. This statement makes little difference if not followed by other important generative approaches. The importance of Husserl's category in the history of modern logic and linguistics is well known. Bar-Hillel said:

> "*Husserl's treatment of meaning categories is an important though not always adequate anticipation of the modern theories of syntactic (or semantic) categories*"[20].

Husserl's categories play a central role in his pure grammar. In particular, when it comes to a specific area of meaning, the most elusive thing is that we do not have the freedom to combine interpretation with meanings, and therefore we must not deliberately mix these elements in a given linking unit. Only in some predetermined manner do the meanings fit together and constitute meaningful, uniform meanings, while other combinatorial possibilities are legally excluded. The inability to merge is a fundamental legal impossibility. This inability to be more precise does not fit with the specificity of the meanings to be unified, but rather with the essential genres from which they come, that is, with the category of meanings.

For Husserl, categories of meaning are the most important basic elements in the production of meaningful linguistic strings as illustrations of uniform conditions. The possibility of forming a well-formed chain therefore depends on the positioning relationship to be linked at the level of a known component. Disregarding the phenomenological motivation and its biased terminology, it is factually consistent with Chomsky's theory. Following

---

[19] Russel B., *My Philosophical Development*, London 1959, p. 245.

[20] Bar-Hillel Y., *Hussrl's Conception of a Purely Logical Gramma*, [in:] *Philosophy and Phenomenological Research*, 17, 1957, p. 366.

Husserl's position, it can be said: The expression "this tree is green" has an unambiguous meaning. If we move formally from a given meaning (independent logical proposition) to a suitably pure form of meaning, to "sentence form", then we get the conclusion that it is a formal concept which deals with all independent interpretations in its scope. It is now clear that we are not completely free within it, but bound by fixed borders. It is not possible to replace one variable with any other meaning. It is true that we can list words in a sentence, but it is important to remember that there are also uniform permanent meanings or grammatical expressions, the meaning of which can only be completed in a uniform way. This will no longer happen as soon as the categories are changed. We can put the words side by side; we can replace figures in a relational statement, but we only have one word at a time. Despite the substantially different motivations of the two systems, this is exactly the same criterion of grammatical quality that Chomsky has. Categories in a string of a given form cannot be mixed, otherwise there is a system of unrelated components that are not grammatical. On the other hand, the strength of the systems allows the formation of any number of grammatical sentences of the same form, respecting the rules of categories, regardless of the fact that grammatical sentences may be a spider web of meaning. There are two different predictive levels, one is "concept" and the other is "concept concept".

Example:

Ideas are dormant.

This sentence can be extended as follows:

New ideas lie dormant.

Individual words can also be replaced with other words in the given categories:

** Green ideas are asleep.*

The above sentence is no longer meaningful. Thus, some words can theoretically be listed in sentences and replaced with other similar ones, however, not always a newly formed sentence makes sense, even though the words are still linked according to grammatical rules.

But even if we agree with Bar-Hillel, we only see a very indirect relation to the changes taking place in Syntactic Structures. To avoid absurd sentences, Chomsky introduced his rules for sub-categorization in later works. However, a more exact relationship with the level of transformation in the Chomsky sense can be found elsewhere.

Husserl's remarks on "modifying the meaning" sound surprisingly modern. He compared the process of modifying the meaning with, the name (= 'Mention') of a word, or with a given statement in a sentence. Intuitively, it is not far removed from a transformation process, though formally in fact, it is actually separate from it. All of this may be over-interpreted, but clearly according to Husserl's observations of the "change of meaning" (where the word, meaning 'must be understood according to its specific use' within

Husserl's theory), over-interpretation is almost impossible. About these changes in meaning, he clearly says:

They are rooted in modifications of meaning in a sense [...] aside from expressions, which is somewhat analogous to arithmetically speaking of "transformations" of arithmetic structures. There are a priori regularities in the domain of meanings according to which meanings are to be transformed into new meanings – in different ways while maintaining the essential core.

As an example, Husserl presented, among others extracting attributive adjectives from the predictive position (possible way of looking at things, as rationalists have already pointed out):

This adjective is predestined for a predicate function and consequently an attribute function, it functions normally, in its "original", unchanged meaning, something in the example is a tree is green. Only the form itself remains – apart from its syntactic function – when we say this green tree.

Of course, we are talking about the so-called "Meaning" that results from a specific philosophical attitude; Husserl, however, wanted the syntax to be understood as a clear reflection of this level. Syntactic "expressions" should be viewed as implementations (contextual "captions") of the abstract core forms obtained by transformation at the level of meaning. These basic forms would, of course, have different syntactic realizations in different languages, "partly following generally human, partly randomly changing empirical motives," said Husserl: if our research "examines primitive meaning structures, primitive types of structuring and linking, and the operational laws based on them, supplementing and modifying meaning, we also recognize the undoubted law of rationalism of the 17th century in the distinction between primitive structures of meaning, primitive types of structuring and linking and the operational laws based on them of supplementing and modifying the meaning of the 18th-century idea of universal grammar"; which was born at the end of the 19th century. It is clear that Husserl would be a serious candidate for the historical sponsorship of the generative movement (also in terms of generative semantics). His work lacks, above all, a well-structured formalization that would allow for an algorithmic approach to the system in such a way as to do justice to the creative process of producing sentences in a natural language. It was a little early, despite tradition, for this logical investigation. Truth, logic, culminating in the epochal work Principia Mathematica, at about the same time, built a bridge between logic, mathematics and linguistics. The time has come for the breakthrough work of the great Russian mathematician Markov, whose application of probability theory to Evgenij Onegin in 1913 was seriously debated in linguistic circles. Nevertheless, it took lively work in the field of group theory and programming language in the second quarter of the 20th century to experience the synthesis of these seemingly very divergent disciplines that would be useful for describing natural languages, at least in terms of generative grammar in the strict sense.

Intensive research into Thues systems, Turing machines, and recursive functions in mathematical logic in general took place almost simultaneously with the development of the American structuralist movement in linguistics. The analytical-paradigmatic distribution of sentences and parts of sentences into their component structure, allegedly according to purely systemic aspects, must at some point give a mathematically well-versed linguist the idea of making an attempt at a syntagmatic and productive description of sentences, especially since mathematical logicians interested in linguistics had made such attempts long ago .

Therefore, in retrospect, it can be said that Chomsky's Syntactic Structures (and, of course, his earlier thesis) were not a surprise from a scientific point of view. This is all the more true when you consider that his teacher Zellig Harris has already initiated progress in this direction. That Chomsky did in fact surprise most linguists is another point.

Chapter 16. Avram Noam Chomsky

Very early on, grammarians began to express their dissatisfaction with the traditional definitions of categories and grammatical functions. The vagueness of the division of sentence elements into those that denote things and those that denote the attributes of activities and the like has been emphasized many times. The anxiety they felt about the material criteria of formal honors encouraged great grammarians at the turn of the century to look for other criteria. Purely morphological criteria were even less sufficient for them, especially for languages such as English, where inflection is nearly extinguished, or for inflection-rich languages such as Sanskrit, where different kinds of words can inflect in the same way. The difficulties of a purely morphological division of categories are clearly seen in the trivial example of the genitive 's' in English:

1. The man's beard is unusual.
2. That man over there's beard is also unusual.
3. That's the man you were talking abort’s brother's dog.
4. That man's a boy I know’s uncle.
5. That's man I know’s mother-in-law.

In other words, the very far-sighted Hermann Paul clearly saw serious problems that fall into categories according to morphological criteria. Nevertheless, he found this type of division to be the most consistent possible. But Brugmann's partner Delbrück, who had to deal intensively with syntax problems in his part of the layout, disagreed.

For this reason, he believed that criteria should be introduced when different kinds of words could be used depending on the purpose of their use.

The time was ripe for such ingenious linguistic directions as that of Ferdinand de Saussure and Edward Sapir. De Saussure made the important distinction between 'la Langue' – the essential, systematically conditioned 'language' that is independent of the individual – and 'la Parole' – the accidental, individually conditioned 'speech'. He passed on the dualism of the rationalists in his distinction between linguistic form and linguistic substance. For him, la Langue consisted of a system of signs that are paradigmatic or syntagmatic denoting each other. De Saussure understood 'signs' as the connection of a denoted 'concept' with an acoustic form as a linguistic denomination. The elements of la Langue were thus determined by their form and their paradigmatic and syntagmatic 'value' in the system[21]. The value of a sign was determined through its opposition or its context-free exchangeability with elements of a certain distributional class (paradigmatic designation) and through its opposition or compatibility in a certain grammatical selection (syntagmatic designation).

De Saussure categorically rejected any attempt to describe grammatical categories independently of their syntactic roles.

Sapir also assumed that grammatical form and grammatical function cannot be separated from one another. In his opinion, every grammatical statement should be based on an analysis of sentence forms. He saw the sentence as a 'linguistic expression of a proposition'.

According to Sapir, a proposition receives grammatical form in that language-specific processes signal the connection between the primary and secondary 'concepts' of the

[21] Cf. de Saussure F., *Course de Linguistique Générale*, Paris 1916, p. 155.

proposition. As grammatical processes, Sapir wanted to understand such things as the arrangement of the sentence elements, affixation, composition, internal modification of a 'radical' (ablaut etc.), reduplication, emphasis etc. What is to be understood by the term 'grammatical concept' can be seen from the analysis of the following sentence:

The farmer kills the duckling.

Concrete Concepts:

First subject of discourse: farmer

Second subject of discourse: duckling

Activity: kill

analyzable into:

Radical Concepts:

Verb: (to) farm

Noun: duck

Verb: kill

Derivational Concepts:

Agentive: expressed by suffix -er

Diminutive: expressed by suffix -ling

Relational Concepts:

Reference:

Definiteness of reference to first subject of discourse: expressed by first the, which has preposed position.

Definiteness of reference to second subject of discourse: expressed by second the, which has preposed position.

Modality:

Declarative: expressed by sequence of “subject” plus verb; and implied by suffixes -s

Personal relations:

Subjectivity of farmer: expressed by position of farmer before kills; and by suffixed -s

Objectivity of duckling: expressed by position of duckling after kills

Number:

Singularity of first subject of discourse: expressed by lack of plural suffix in farmer; and by suffix -s in following verb

Singularity of second subject of discourse: expressed by lack of plural suffix in duckling

Time:

Present: expressed by lack of preterit suffix in verb; and by suffixed -s.

Sapir let the form of a sentence be determined by his grammatical processes and the distribution of his 'types of terms'. Such a way of looking at things can easily be described as structuralistic. This is particularly evident when one has in mind Sapir's subdivision of the basic elements. The symbolizations of words like thanks, thankless, thanksgiving, as A + (b), A + b and A + (b) + B + (c) correspond pretty exactly to the morpheme analysis of the later structural analysts[22]. What was not yet structuralist with Sapir was his method. What cannot be called structuralistic is Sapir's emphasis on semantic terms[23]:

What, then, are the absolutely essential concepts in speech, the concepts that must be expressed if language is to be a satisfactory means of communication? Clearly we must have, first of all, a large stock of basic or radical concepts, the concrete wherewithal of speech. We must have objects, actions, qualities to talk about, and these must have their corresponding symbols in independent words or in radical elements. No proposition, however abstract its intent, is humanly possible without being tied at one or more points to the concrete world of sense. In every intelligible proposition at least two of these radical ideas must be expressed, though in exceptional cases one or even both may be understood from the context. And, secondly, such relational concepts must be expressed as the interrelation of concrete concepts and construct a definite, fundamental form of proposition.

It did not take long before an attempt was made, in the grammatical analysis of a sentence, to get as far as possible from the seminal criteria. As is well known, the structuralists have tried to study the structural meaning 'of a sentence or utterance'.

As Fries put it: "The grammar of language consists of the devices that signal structural meanings"[24].

That meant that its abstract, purely structural linguistic level had to be postulated. A seemingly pointless sentence like:

**The Gnafel believes that the Inks are sniffy.*

it has, as it has been seen, a point of some sort. Such a sense of the parts of sentences can only be found in their functional or structural roles, whereby their functional roles can actually be viewed as functions of their form classes. From a structural point of view, the form classes are 'signaled' by different types of features: position, intonation, inflections (and other segmental morphological signals), structural words (= functional words), etc. Structural sense (in order to essentially retain the term used by the American structuralists) can be derived from this at Bloomfield as the sense of a form class[25].

The behavior of the concepts of the American structuralists can basically be called the sense of the class of form, following the example of Bloomfield. So you can make

22 Cf. Sapir E., *Language: An Introduction to the Study of Speech*, New York, 1921, p. 24 f.

23 Ibidem, p. 93.

24 Fries Ch., *The Structure of English: An Introduction to the Construction of English Sentences*, New York, 1952, p. 56.

25 Cf. Bloomfield L., *Language*, New York, 1933, p. 146 and p. 202.

a structural sense by replacing the lexemes with their form classes and interpreting a series of form classes (in one sentence). This leads to some kind of grammar that works primarily on the criteria for the distribution of sentence components – regardless of the lexical meaning of these components. The related "misinterpretation" or "wrong specification" places the analysis on the "Langue" level as opposed to the "Parole" level, because here it is only a matter of grammatical universals, i.e. the essential and not random rules of a given language.

Jespersen's claim that phrases are only specified at the "Parole" level is based on confusing a sentence, word, etc. using (= Utterance) a sentence. By "utterance", structuralists and earlier grammarians often seem to mean only one sentence which is stylistically considered to be a unit of spoken language. But it is obvious, of course, that the names of the sentences:

What should we do?

What to do?

both are only sentence names, i.e. II. it certainly has no more to do with "Parole" than I. Both can be "named" or "mentioned", as well as "used" orally or in writing. II. it has a structural significance which would depend on the particular speaker or act of speech. It should be obvious that super segmental structural features, insofar as they are relevant to a structural sense, depend as little on individually applied sentences as segmental features. They are no less and no more universal, no more and no less abstract. The same is true of the structural features of ambiguous sentences that signal meaning. They are two different structural meanings that share the same name (= 'Token').

It is important that a certain part of a sentence can appear in a sentence under normal distribution conditions, and not that it actually occurs at some point in a sentence spoken by a certain person. Fries considered as early as 1952[26]:

One cannot speak or understand a language without "knowing" its grammar – not consciously, of course, but in the sense of making the proper responses to the devices that signal structural meanings and also of producing the proper signals of their own structural meanings.

One of the earliest steps in learning to talk is this learning to use automatically the patterns of form and arrangement that constitute the devices to signal structural meaning. So thoroughly have they become unconscious habits in very early childhood that the ordinary adult speaker of English finds it extremely difficult not only to describe what he does in these matters but even to realize that there is any thing there to be described. One of the basic assumptions of our approach here to the grammatical analysis of sentences is that all the structural signals in English are strictly formal matters that can be described in physical terms of forms, and arrangements of order.

In order to graphically present the "patterns of form and arrangement, which constitute devices signaling the structural meaning", the designers divided the proposals into their

---

[26] Fries Ch., *The Structure of English: An Introduction to the Construction of English Sentences,* New York, 1952, p. 57-58.

component construction layers in the Immediate-Constituent-Bracketing process. This was to reveal the structural meaning of the sentence. This procedure, however, required knowledge of the classes of form of the individual sentence components. Such insight was obtained by identifying contrasts using paradigmatic substitutions, which allowed for the assignment of grammatical categories (or form classes) to groups of lexemes. This procedure can be used to name any word (an adjective) that can fill two blanks in the matrix below:

The ______________ man is very __________________ .

Words that contrast with the adjective naturally do not match the scheme, so they belong to different classes of forms. The classes of the proposition's component form can be represented by numbers (cf. eg Fries[27]) or more suggestively by category names (cf. e.g. Harris[28]). However, it remains an overwhelmingly paradigmatic inductive procedure that can only be used word for word, sentence for sentence, etc., and can develop. This leads to a poorly descriptive system that cannot unambiguously capture many functions and relationships. A generalization aimed at establishing the relationship between the following sentences for now:

Rainer pinched Rosa.

Rosa was pinched by Rainer.

Rainer, pinch Rosa!

Did Rainer pinch Rosa?

If we do not make such a system, a generalization can only be made if it shifts syntagmatically deductively. The basic elements of a syntagmatic deductive system, however, are at least partially contained in structural analysis. For structuralists, the most important were the rigid analysis of the component structure and the systematic description of the sentence structure through the description of the signals of its structures. It was also important to symbolize the abstract classes that make up the classes of forms. This symbolization of the abstract components of sentence types was certainly the starting point for discourse analysis, between symbolic English sentences, such as:

N1 V N2

and

N2 is V-en by N1

Harris defined his equivalence relation as follows[29]:

In particular, we will say that sentences of the form A are equivalent to sentences of the form B, if for each sentence A we can find a sentence B containing the same morphemes except for differences due to the difference in form between A and B.

---

[27] Op. cit.

[28] Cf. Harris Z., *Methods in Structural Linguistics*, Chicago, 1951.

[29] Cf. Harris Z., *Discourse Analysis, [in:] Papers in Structural and Transformational Linguistics,* Dordrecht, 1970, p. 334.

The definition of grammatical equivalent sentence classes corresponds to the primitive definition of grammatical transformations. They were only symbolically marked by Harris (known as Chomsky's teacher), but not yet formalized. "They were only formalized by Chomski within his formal grammatical model. The structuralist system thus provides abstract, structurally defined sentences and abstract classes of forms, which are defined by their place in the sentence. In addition, it provides an insight into the complex relationships between classes of forms of specific sentences through linear (segmental and suprasegmental) statements, describing the positions of the classes of forms in different component layers. Structural functions in a sentence such as modification, complement, co-ordination, subject, predicate, object etc. can also be defined relatively efficiently when considering stratified groups, but we are still dealing with a poorly descriptive system that can only "respond" analytically to given body. Sentence types and classes of forms are inductively determined, and certain relationships between different sentence types can be analyzed analytically sentence by sentence. Nevertheless, the relationships and functions resulting from the relationship between related but structurally different types of sentences cannot be clearly described. The two most important description axes are available – the linear axis and the hierarchical axis of the component structure description. However, as a result of using different analytical processes, covering different types of systems, they are related to each other in a systematic context, and this relationship is not formal.

Chomsky shifted two issues of description into a formal context. For this purpose, he used the principles of mathematical logic. In analytically obtained abstract sentence descriptions of structuralists, Chomsky saw possible criteria for the definition of a finite set of primitive schemas of structures that could be used as the basis for an infinite number of possible sentences of a certain language, especially the English language. Structurally defined form class configurations based on these schemas were directly well-formed criteria. The paradigmatic relationship that takes place within forms according to purely structural conditions between the elements of a sentence directly related to conditions in free variation has made it easy to think about reformulating into a formalized system in which class names and morphemes (as names) are grammatically defined basic elements , replaced by uninterpreted meta variables and terminal symbols of the formal alphabet. On the plane of Chomsky's linear description there was a well-formed chain of morphemes that corresponded to his lexical and grammatical "formative" chains. He found that they were juxtaposed with sentence schemas that were defined by their component classes. On the level of component structure analysis, Chomsky saw the possibility of algorithmizing the various structural layers analogously to the steps in providing evidence. Fries describes the levels of the structural analysis of an English sentence as follows:[30]

[30] Fries Ch., *The Structure of English: An Introduction to the Construction of English Sentences,* New York, 1952, pp. 267-268.

1. The first step is an identification of the parts of speech and the function words. This operation includes the noting of the inflectional forms and other similar formal features of the form-classes.
2. The second step is the marking of the special ties that are signaled by a concordance of forms or by particular intonation contrasts.
3. The third step is the identifying of the particular arrangement of the Class 1 and Class 2 words that signal the kind of sentence.
4. The fourth step is the identifying of the particular arrangement of the Class 1 words (not in word-groups with function words) before and after the Class 2 word.
5. The fifth step is to cut off of any "sequence" signals that stand either at the beginning or at the end of the sentence. The relation of these sequence signals is to the sentence as a whole, not to any particular part.
6. The sixth step is to cut off of an included sentence that stands at the beginning of the utterance in which it is included. The sentence included in this position is related to the whole of the unit that follows.
7. The seventh step is the cutting between the Class 1 word and the Class 2 word that form the basic arrangement of the sentence.
8. The eighth step is the cuttings separating the various modifiers of the Class 1 word that is "subject". These cuttings can proceed mechanically in accord with the use of word order in present-day English. With multiple modifiers the modification is cumulative and directional. Postmodifiers are cut off first, beginning with the last one. Word groups as modifiers are treated on this level as whole units in relation to the head to which they are attached. The analysis of the arrangement within the group is of a different structural layer.
9. The ninth step is the cuttings separating the various modifiers of the Class 2 word. These cuttings can also proceed mechanically in accord with the use of word order in present-day English. With multiple modifiers the modification here too is cumulative and directional. Premodifiers are cut off first. Postmodifiers are cut off next, beginning with the last. On this level, word-group modifiers are treated as whole units.
10. The tenth step is the cuttings, following a similar procedure, within the word groups that have been treated as whole units on the level above.

Fries word classes 1 and 2 correspond to classes N and V in a somewhat simplified and somewhat abridged English sentence analysis:

The man likes a woman.

and without taking into account the morphological plane, they look something like this:

Firstly, through paradigmatic substitutions, classes of forms of individual segments are identified:

a	b	c	d	e
the	man	likes	a	woman
one	boy	sees	one	boy
some	dog	hears	some	dog

these	students	hate	those	professors
*man	*the	*the	*the	*the
*likes	*likes	*man	*man	man
a	*a	*a	*likes	*likes
*woman	woman	*woman	*woman	*a

Items "a" and "d" as well as "b" and "e" are interchangeable. Contrasts (forms marked with an asterisk) show which positions cannot be changed. The substitution options allow for three classes of forms, denoted by the symbols, D ', N', V '.

This results in the following chain:

D N V D N

B. Then the components of a certain level – starting with the entire chain – are broken down into their successive lower components in a mostly binary procedure:

(D N) (V D N)

(D N) ((V) (D N))

(((D) (N)) ((V) ((D) (N))).

Of course, each component level – including sentence level – can be specified using the substitution and contrast options. Finally, the functions of the various components and the mutual relations are structurally defined (cf. "subject", "predicate", "complement", etc.), so that the parentheses have a systemic meaning.

In addition to using inductive discovery procedures, the steps in this decomposition system can be easily reformulated into steps in the deductive derivative process. The gradual specialization of parts of a sentence according to purely structural principles is very suitable for translation into sentence formation defined by the generation rules. The unit in parentheses can be understood as the form class (or category) at each component level and the corresponding – already interpreted – parentheses denoted by their names. Then the transition to the formal expression in parentheses applies to the following form:[31]

(((the)d(man)N)NP.((likes)V((a)D(woman)N)NP)VP)S.

The left labels are omitted because there is a clearly defined pair of square brackets[32]. Of course, a parenthesis expression can also be represented by an equivalent expression structure tree[33]:

[31] Cf. Gross M., Lentin A., *Mathematische Linguistik,* Berlin – Heidelberg, 1971, pp. 95 f and p. 114. f. Cf. also Maurer, H.: Theoretische Grundlagen der Programmiersprachen. Mannheim, 1969, p. 185.

[32] Cf. Gross M., Lentin A., *Mathematische Linguistik*, Berlin – Heidelberg, 1971, p. 116.

[33] Cf. ibidem, pp. 114 f.

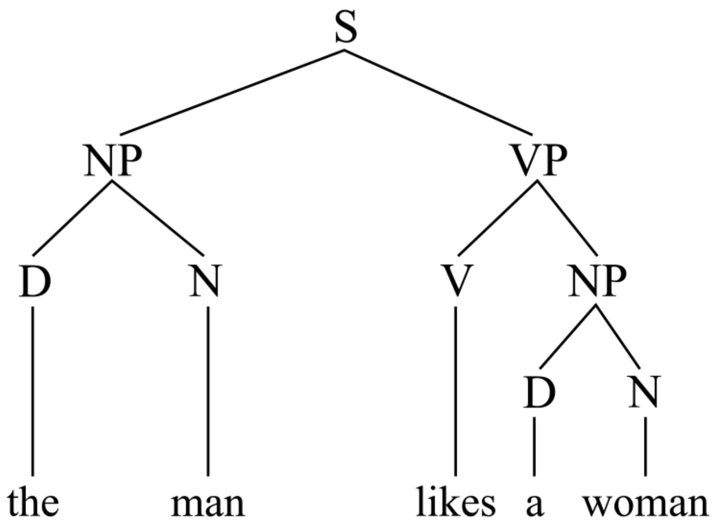

Figure 1. Expression structure tree,
source: Gardner T., *Hauptstömungen der modernen Linguistik*, Göttingen 1973, p 27.

Different symbols could be chosen for the two NPs and for the dominant D and N to insert the phrases "the man" and "a woman" in the appropriate places (many grammars do this with pointers). This can also be guaranteed by selecting the correct order of rules. This is a problem that is more or less irrelevant to the grammar of the sentence; because:

A woman likes the man.

it is also grammatically correct.

One of Chomsky's greatest achievements was to construct a grammar that gave one sentence an appropriate description of the structure of a phrase, i.e. assigns a structure tree or expression in parentheses. The context-free grammar of Chomsky's phrase structure is a special form of the Semi-Thue-System. You can define the grammatical model of G as follows:

1. A finite set of symbols

Mt = {the, man, likes, a, woman}

2. A finite set of metalinguistic auxiliary symbols

Auxiliary symbols Mh = {S, NP, VP, D, N, V}

3. The correct beginning of the symbol S (with S ∈ Mh)
4. A single binary concatenation operation (denoted by the symbol, + ')
5. A finite set of context-free sequence rules F = {S-> NP + VP, NP-> D + N, VP-> V + NP, D-> the, a, N-> man, woman, V-> likes} where all rules are either final (in the form A-> x, with A ∈ Mh and x ∈ Mt) or untimely (in the form A-> B, with A ∪ B ∈ Mh).

If we start with the phrase structure tree (23), the dominant nodes at each component level with their directly dependent nodes (S with NPVP, NP with DN, etc.) can be understood as ordered pairs. This partially results in a hierarchical arrangement of the rules – the arrangement of the rules at the component level must of course remain arbitrary: it is irrelevant whether NP or VP (both of the same type) is extended first. It is a formalization of a sentence. Discovery Procedures can therefore still be described as mostly structural: rules derive from description.

But you can also act purely formally. A partial rule hierarchy is already default when marking start symbol and when selecting different rule types. The fact that all rules are strictly sequential also suggests classification criteria, at least indirectly; because in their application all non-terminal symbols move to the left and disappear (this does not mean, of course, that the rule cannot be applied repeatedly); The rules of F can be simplified somewhat from the following general form: X1 -> Xj, with i <j and X1 ≠ Xj. To keep the number of rules as small as possible, i.e. to avoid unnecessary rules and to organize the grammar efficiently, look for a sequence that will speed up the gradual disappearance of symbols. This "acceleration process" can be promoted by introducing additional conditions that the rules must meet. The following examples can serve as examples of such conditions:

1. The symbol cannot appear in two different rules on the left.
2. Symbols cannot be empty.
3. There can never be more than one symbol on the left or only one symbol can be replaced with the rule.
4. The symbol cannot appear on the left and right sides of the same rule.
5. A terminal rule can have only one symbol on the right, and a non-terminal rule - only two.
6. Two different rules cannot have the same symbols on the right or left side, i.e. they cannot differ on one side only.
7. Start symbols appear only once and appear only on the left side.
8. The symbol does not appear twice in the same place in the rule (cf. A-> BC, C-> BD).

Mostly a formally oriented routine that is only mentioned here refers to a specific sentence, or at least to a specific type of sentence that needs to be generated. The grammar that (16) describes above and describes only structurally, i.e. what completely ignores its interpretation, is important. The description (16) here only means the above-mentioned assignment to (16) from the structural description in the formal grammar G. And this is a very important point; because we would be abusing the grammar of Chomsky's phrase structure if we understood a single symbol from the outset with its own meaning that could affect the grammar in any way. The rule, NP-> D + N 'must be "Replace symbol, NP" with a string of symbols, D + N ", not" Replace noun phrase with determinator combined with noun". The rule is also called N-> woman ":" Replace symbol, N "with symbol, woman" "and so on. Both interpretations of the corresponding structured description must be done using completely different kinds of rules.

We might as well use the symbols a, b, c, d, e, f as auxiliary symbols and the rule set F by the set F '= {a -> b + c; b-> d + e; c-> f + b; d-> the, a; e-> man, woman; f-> likes}. And of course, we could also use other symbols with appropriate terminal rules, provided we can assign them constant values necessary for interpretation by certain correspondence rules. However, these constant values would not play a role even in the grammar of the phrase structure.

'Woman' is an English feminine noun as long as it relates meaningfully to symbols and strings with a specific grammatical structure, and is not generated by grammar itself.

'Woman' refers to a female human being if she is to be regarded as metalinguistic expressions of grammar symbols.

The grammar of the phrase structure does not say about the meaning of its own symbols, it only defines the systems and combinations of these symbols that are considered grammatical, and it defines the syntactic relations in which they can meet each other. First, it generates grammatical sentences, not meaningful ones. The statements of such a grammar are purely syntactic as long as the generated language has not been interpreted, which would require a different kind of grammar. If we consider the interpretation as a translation into another – perhaps isomorphic – language using a set of correspondence rules, we can leave the definition of a different grammar. All of this does not mean that symbols cannot be names of certain entities that are important for interpretation. In fact, this is conducive to the adequacy conditions. Nor does it imply that the end symbols should not be identical to the words in the object-oriented language being described. This is the easiest and most economical way. It only means that the term "grammar" has nothing to do with the meaning or lexical meaning of symbols if it emerges from the grammatical rules of the phrase structure.

Even structuralists wanted to define the term "grammatical sentence" regardless of the lexical meaning.

**Uggles wuggle wumply.*

Having a structured sense in the English language was evidence that Chomsky insisted there was a difference between them:

**Colorless green ideas sleep furiously.*

and

**Furiously ideas green sleep colorless*.

Chomsky rightly pointed out that the ability of a native language user to recognize and form grammatical sentences cannot depend on his ability to determine statistically more likely words. Chomsky noted that an English speaker would find it easier to remember sentences like the above and be able to recall them even if he had never encountered any of these sentences before.

The likelihood of distribution depends on the grammatical category, not the lexical meaning, if it plays any role at all.

Suppose grammar symbols have no meaning of their own until they are interpreted by a separate semantic component and the term "grammatical sentence" is defined purely syntactically. It is striking that the breakdown can be quite different.

Chomsky's emphasis on descriptive adequacy as the goal of his phraseological-structural grammar – and as an evaluation criterion - forces him to combine informal inductive and formal deductive 'Discovery Procedures', by means of which he wants to arrive at the simplest, and in terms of their descriptive adequacy, the most acceptable set of rules, or rather, forces it to precede deductive formal 'Procedures' with inductive informal 'Procedures'. Since Chomsky did not believe that a theory of grammar could contain its own discovery procedures anyway, none of this bothered him at all. On the contrary, he emphasized the priority of presystematic identification of the corpus of grammatical sentences in the language to be described. As he put it:

> *[…] We may assume… that certain sequences of phonemes are definitely sentences, and that certain other sequences are definitely non-sentences. In many intermediate cases we shall be prepared to let the grammar itself decide, when the grammar is set up in the simplest way so that it includes the clear sentences and excludes the clear non sentences. This is familiar feature of explication. A certain number of clear cases, then, will provide us with a criterion for adequacy for any particular grammar. For a single language, taken in isolation, this provides only a weak test of adequacy, since many different grammars may handle the clear cases properly. This can be generalized to a very strong condition, however, if we insist that the clear cases be handled properly for each language by grammars all of which are constructed by the same method. That is, each grammar is related to the corpus of sentences in the language it describes by a given linguistic theory.*

In some ways, the theorist Chomsky is conventional. He wants to define the nature of his grammar within a given theory. The kinds of grammatical rules arising from this theory must not be questioned. On the other hand, a specific form of grammar is judged by whether the rules are simpler than the rules of another grammar of the same type. The types of rules are 'agreed' within the chosen theory, the form of the grammar depends on the effectiveness of its rule set, measured against other rule sets of the same kind. Two grammars are of the same type if the following conditions are true:

1. They produce the same set of end chains (= weak equivalence).
2. They have the same set of constants or operations (eg: ->, +}).
3. They have the same kinds of rules (e.g .: Normal symbols; A -> BC; linear, A -> xB; left linear, A -> xB; right linear, A -> Bx; finally, A -> x - where A, B, and C are the trailing metasymbols, ax, y).

The degree of relatedness between different grammars of the same kind can be measured by the strength of an average set of uniform rules.

However, regarding his 'Discovery Procedures', Chomsky's practice is a moderate constructivist:

> *we shall never consider the question of how one arrived at the grammar whose simplicity is being determined; e.g., how one might have discovered the analysis of the verb might have phrase presented in 5.3. Questions of this sort are not relevant to the program of research that we have outlined above. One may arrive at a grammar by intuition, guesswork, all sorts partial methodological hints, reliance on past experience, etc. Itis no doubt possible to*

> *give an organized account of many useful procedures of analysis, but it is questionable whether these can be formulated rigorously, exhaustively and simply enough to qualify as a practical and mechanical discovery procedure*.

Using these observed cases, he would then like to explain other less clear-cut cases within a construct in the system:

> *Notice that to meet the aims of grammar, given a linguistic theory, it is sufficient to have a partial knowledge of the sentences (i.e., a corpus) of the language, since a linguistic theory will state the relation between the set of observed sentences and the set of grammatical sentences; i.e., it will define "grammatical sentence" in terms of "observed sentence", certain properties of observed sentences, and certain properties of grammars. To use Quine's formulation, a linguistic theory will give a general ex planation for what 'could' be in language on the basis of "what is plus simplicity of the laws whereby we describe and extrapolate what is".*

Chomsky may be right when he says that an effective theory and a collection of clear cases can be used to show ways to describe obscure cases. And it seems that what generative grammar theory – at least in terms of syntax – can achieve to a large extent. But how do bright cases happen? In order to be able to determine which sentences can be included in the corpus of grammatical sentences, it must rely on the intuition of a *"native speaker". In order to define the properties of a grammatical sentence, it must refer to „precise reconstruction of large parts of the traditional notion 'parsing' or, in its more modern version, immediate constituent analysis".*

Chomsky himself doubted that it was possible to find algorithms that could promote the discovery of grammar for natural languages, as already mentioned above there are no external detection routines for him from the system. It is a strange coincidence that with an almost irrelevant trivial algorithm, a non-trivial structural description of the type of a given sentence can be obtained formally:

1. A series of six symbols was ordered (the number of necessary primary node labels plus the starting symbol), e.g. A, B, C, D, E, F.
2. A matrix with three rows and three columns was drawn and the first three symbols were successively entered in the first column:

1 2 3

1 A

2 B

3 C.

3. The first symbol was omitted (this gives the starting symbol) and the remaining symbols are entered one after the other, starting from the top, into the matrix from left to right, until the row of symbols is exhausted:

1 2 3

1 A B C

2 B D E

3 C F

4. The first symbol was omitted and the column was still open with a second symbol:

1 2 3

1 A B C

2 B D E

C F B

5. Follow the (hierarchically) ordered untimely rules as follows:

A -> BC

B -> DE

C -> FB

and lead the pre-term chain:

D + E + F + D + E

with description 1:

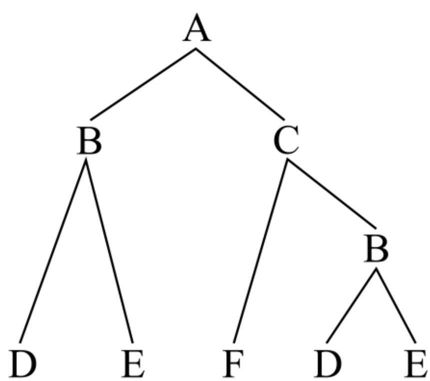

Figure 2. Description 1,
source: Gardner T., *Hauptstömungen der modernen Linguistik*, Göttingen 1973, p 46.

With appropriate terminal rules, the resulting pre-term string (i.e., from the string of symbols shown on the left in the rules) gives a structure description strongly equivalent to the above, it plays a central role in the grammatical model of grammatical structures. All this, completely by chance, thanks to an extremely trivial, completely arbitrary algorithm.

Other such arbitrary algorithms have led to the strangest results. There is no practical way to circumvent the empirically derived set of "bright cases" in the natural language being described. Chomsky was pragmatic enough to see it and at least be able to say yes and no. His actual march was from the above-mentioned translation of the component structure analyzes to the given formal system. The definition of a grammatical sentence is empirically secured from the very beginning in the formal system. Initially, the formalization could only cover fragments of the English language, namely fragments consisting of primitive, simple sentences that did not require any complicated special rules and did not show any visible irregularities. For Chomsky, these were obvious cases. By extending the grammar that generated these bits of English, he wanted to come to full grammar within general grammar theory. The type of extension depends on the structural properties of a given natural language and its implications for formal grammar.

Example:

If it is true that if it rains in Göttingen, then we will get wet, then we will get wet.

For this reason, internal improvements have been made to expanded grammar in relation to the criterion of simplicity. The internal structure of extended grammar is determined by comparing different grammatical models of the same type and, if possible, with the same descriptive performance, and selecting the model with the simplest structure. It should be clear that grammars derived from different grammatical theories do not allow for a systematic evaluation of simplicity if neither of them can be translated into grammar of the same type. The most primitive way to apply this criterion is to simply quantify the rule sets of different grammars. The following sets of rules are compared:

a)	b)
1. $S \rightarrow NP_1 + VP$	1. $S \rightarrow NP + VP$
2. $NP_1 \rightarrow N_1$	2. $NP \rightarrow N$
3. $VP \rightarrow V + NP_2$	3. $VP \rightarrow V + NP$
4. $NP_2 \rightarrow N_2$	4. $N \rightarrow$ *John*
5. $N_1 \rightarrow$ *John*	5. $N \rightarrow$ *Mary*
6. $N_2 \rightarrow$ *Mary*	6. $V \rightarrow$ *likes*
7. $V \rightarrow$ *likes*	

Thus, it can be seen that e.g. repeated use of rule b. 2 between b. 4 and b. 5 of the amounts of a and b generate the same terminal chains. It is also easy to see that a and b are strongly equivalent; they form the same tree of structure.

The seven rules correspond to the six rules of a different grammar.

If the descriptive performance of both grammars is the same, a grammar with fewer rules is preferred.

A higher form of applying the criterion of simplicity is to evaluate the internal structure of different rule-based grammars. Let G1 and G be two different programming grammars that generate the same set of terminal strings and the same set of F rules: F1, F2, ..., Fn represent the rule sequence in G1 and Fr, Fs, ..., FΔ the rule sequence in G2 , where r, s, ..., Δ of numbers 1, 2, ..., n are numbers in the same or different order. For each rule F1 there is a corresponding set of indicators Z1 which specifies the rule (s) to be considered immediately after applying F1: it is obvious that Z1 $\subseteq$ {1, 2, ..., n}. If no valid rule is established via Z1, the output is aborted. If the G2 index set is stronger – if it contains more elements than the G1 index set, G1 is simpler than G2.

Categories can be used as terminal symbols. The grammar must allow descriptions that include sentences such as:

John likes singing girls very much.

John likes singing songs very much captures that but sentences like:

John likes singing girls loudly.

John likes singing girls and songs

excludes.

Even with an analytical process in advance, the problems associated with the application of finite grammars to natural languages remained significant. The most important problem of the m-dependency (= relationship of the dependency between different loops in a model) cannot be solved, i. e. not without the grammar degenerating into an extensive and complicated list of possible sentence types.

Chomsky went a different way. He divided the sentences into primary and secondary or derivative types on the assumption that formation rules for the primary sentences are relatively easy to formulate. Deviating from Chomsky, we can define the primary sentences as those that are obtained from 'grammatical prime chains' by substitution. The grammatical prime chains are chains in which the individual formation categories cannot appear more than once: a symbol must not appear twice with the same 'rank' in a prime chain. A certain amount of the sentences that are extracted in this way (i.e. directly from prime chains) Chomsky called the core sentences of a language. Every derived sentence is based on at least one primary sentence. Especially since formation categories in the Chomsky system can only be defined by their occurrence in a derivation, one must extract all intrusive (second) occurrences of formation categories from other prime chains in prime chains. In short, all chains can be traced back to one or more prime chains. A chain with an 'intrusive' formation category goes e.g. B. back to two prime chains. This can be illustrated like this:

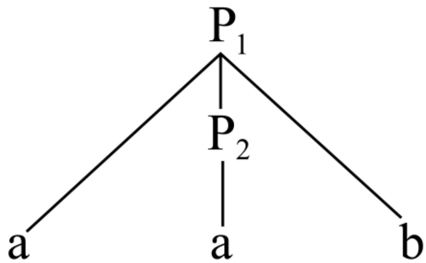

Figure 3: Chain with a formation category,
source: Gardner T., *Hauptstömungen der modernen Linguistik*, Göttingen 1973, p 70.

In "Syntactic Structures" Chomsky spoke of different 'levels of representation': the phrase structure level, the transformation level and the morphophonemic level. In this regard, he defined his core sentences as follows:

> *Suppose that we have a grammar G with a (E,F) part and a transformational part, and suppose that the transformational part has certain obligatory transformations and certain optional ones. Then we define the kernel of the language in terms of the grammar G) as the set of sentences that are produced when we apply obligatory transformations to the terminal strings of the [2, F] grammar. The transformational part of the grammar will be set up in such a way that transformations can apply to kernel sentences (more correctly, to the forms that underlie kernel sentences-i.e., to terminal strings of the [, F] part of the grammar) or to prior transforms. Thus every sentence of the language will either belong to the kernel or will be derived from the strings underlying one or more kernel sentences by a sequence of one or more transformations*.

This relatively general definition must be seen in the context of the search for systematic simplicity. Chomsky selected his core sentences according to whether they could be generated with a simple phrase structure grammar. For various reasons, some of which were traditional, he preferred simple positive declarative sentences to other simple sentence types. He was able to justify his preference for positive declarative sentences with the resulting simplifications in the description of pairs such as:

John chose sincerity.

Sincerity was chosen by John.

Chomsky formulated these terms differently in 1964:

> *Each major category has associated with it a "designated element" as a member. This designated element may actually be realized [e.g., it for abstract Nouns, some (one, thing)], or it may be an abstract "dummy element". It is this designated representative of the category that must appear in the underlying strings for those transformations that do not preserve, in the transform, a specification of the actual terminal representative of the category in question. In other words, a transformation can delete an element only if this element is the designated representative of a category, or if the structural condition that defines this transformation states that the deleted element is*

> *structurally indentical to another element of the transformed string. A deleted element is, therefore, always recoverable.*

This formulation assumes that the initial grammar of the sentence structure serves as the basis for the transformations: the emphasis here is therefore on a specific line in the derivation of a sentence and only indirectly on the derivation (or Phrase, Marker) in general. In other words, the transformations here are considered that they alter a certain linear string (as a result of the grammar of the phrase structure). It is therefore necessary to define the element types of this linear chain so that only some of them can be deleted. Chomsky's conditions are designed to avoid uncontrolled shortening with effects such as:

The car was stolen by the meanest thief in all creation.

→

The car was stolen.

Perhaps Chomsky fell victim to an illusory problem in this formulation. In this case, it is only about retrieving the erased element, and not about clarifying any semantic relationship between two sentences. The conditions should only be specified for the backing chain on which the sentence is based, more precisely, the backing chains should be specified in such a way that it is not possible to obtain both sets from the same backing chain. This seems to be guaranteed as soon as these two sets can be based on different kinds of backbones. It is also questionable whether the inclusion of designated pro-elements in the repayment terms is not superfluous if the rules are formulated in such a way that non-truncated pseudo-categories must be implemented by such elements. However, these elements can be considered as markers of certain types of sentences in the above sense, and therefore the need for special treatment does not apply in each case.

The set of different readings of a given sentence is the set of its interpretations. As a result of this procedure, a number of interpretations are created for each sentence, which are obtained in total by adding the interpretations of the individual components. As Weinreich noted, there is no compelling rule for the elements of the original reading order, nor for the elements of a combined reading order that motivate a hierarchical or other layout. You can see immediately how arbitrary the semantic characterization must be. Even if a set of fairly plausible universal semantic identifiers could be found, it seems unlikely that they would be clearly hierarchically ordered as it would likely indicate a hierarchical structure of the designated things themselves. Either way, it is easy to find examples that show a conversion of two reading sequences – after branching them to higher levels:

(Human)

fox → (Object) → (Animate) (Cunning) → …

(Animal)

The thesis that the interpretation consists of adding up disordered features also does not seem to be correct. Such a thesis would not be consistent with the obvious difference in the relations between adjectives and nouns, as illustrated in the following sentences:

?Nixon is a taller president than Harry Truman.

Nixon is a taller man than Harry Truman.

Tall pygmies are not tall.

And also sentences such as:

My bitch had puppies.

*My female dog had puppies.

*My she-dog had puppies.

As can be seen, the semantic component of the Standard Model is still a fairly open problem. It is also probably the most difficult problem for establishing a unified language theory.

While the generative movement can in principle be characterized as predominantly syntactic, Chomsky and his colleague Morris Halle outlined the phonological interpretive component.

Phonological rules have several functions: they define the distributive compatibility of different segments (they show for example that combinations like 'kltk' etc. do not occur in English, and combinations like 'kstr' etc. are only allowed in some positions); describe contextual (combinatorial) relationships between segments that occur together (they generalize, for example, the difference between combinations marked with a letter, "x" in "auxiliary" formations); they eliminate redundant features; they change grammatical symbols into phonological ones and remove those that cannot be represented in a phonetic representation; they define phonological expressions as fields for certain rule-related operations, etc. Typically, phonological rules are applied by analogy to transforming a syntactic component linearly from left to right at one level and cyclically (from the deepest parentheses out) from the smallest to the largest, i.e. from the lowest to the highest ranking ingredients.

Extracting words with a surface structure from the underlying abstract forms continues and formalizes the ancient morpho phonemics of the structuralists. The characteristic matrices go back directly to the matrices of the Prague School; are only slightly changed in terms of new systematic embedding. For purely systemic reasons, the term "phoneme" is only used informally in the Chomsky model.

Phonology is an extremely interesting field of transformational grammar. Chomsky and Halle have taken important steps towards an effective phonological theory. Nevertheless, this work is currently somewhat outside the mainstream of modern (generative) linguistic research. It is the "age of syntax" – perhaps in response to the age of phonology that preceded the 20th century. Therefore, we do not go into the details of generative phonology here. In the discussion so far, we have discussed the historical embeddedness and structure of the standard theory. Next, we will look at the major adjustments to this theory that may have been made by Chomsky himself and his critics.

Chapter 17. Review and Assessment

As early as 1965, linguists committed to applying generative theory began to exceed the formal limits set by Chomsky to such an extent that it was possible to predict a change of category for their type of grammatical description. Assuming it was still within the overall framework of Chomsky's theory, George Lakoff completed his PhD thesis this year, "On the Nature of Syntactic Irregularity". Lakoff's work, at least as originally intended, was merely an extension of Chomsky's theory However, this very strictly formalistic theory could no longer deal with Lakoff's higher level of abstraction in the syntactic base.

Lakoff wanted to explain some syntactic irregularities as only apparently irregular by tracing structures which, in his view, were only superficially different from deeper structures. He wanted to reduce the sentences like:

The sauce became thick.

The sauce got thick.

The sauce thickened.

to the same deep structure. He wanted to describe superficial differences with the help of new lexical rules that assumed them to be lexical in nature. Lakoff separated syntactically an important part of the lexicon (ie, the part containing complex symbols consisting of contextual features) into a lexical base and a lexical extension. In the words of Lakoff[34]:

> *The collection of syntactic features associated with each sense (semantic reading) of a lexical item can be thought of as being divided into two mutually exclusive parts. We define the "lexical base" of an item (with respect to a given sense) as the collection of all syntactic features that determine which deep structures the item may be inserted into without violation of the deep structure level. That is, the "lexical base" determines the deep structure distribution of the item. All other syntactic information associated with each sense of a lexical item, we will call the "lexical extension" of the item with respect to sense. The "lexical extension" will, therefore, contain syntactic features that have nothing to do with deep structure and will include the Boolean function of marked exception features (R features and SD features) associated with each sense of a lexical item*.

Lakoff's verification is based on the assumption that lexical insertion includes both grammatical and complex lexical symbols. He distinguishes between lexical and grammatical features – similar to Chomsky's first insertion procedure – but emphasizes that after lexical insertion, both kinds of features are present in the terminal string of the base element. Technically speaking, it is done as follows: each lexical category of the pre-end chain dominates the ordered pair of feature matrices; the left matrix contains grammatical features, the right matrix is empty until it is filled with lexemic features during lexical input. A grammatical matrix contains the subcategory specifications for the transformation rules (R) of the grammar and for the structural descriptions of the individual transformation rules (SD); at the level of the depth structure, there are

[34] Ibidem, p. 87.

negative indications for the time being in the R and SD specifications of the grammatical matrix. If a rule is satisfied in the structured application of the rules in the structure description transformation component, the meta rule converts the negatively marked SD specification into a positive binding specification, and then the negative R specification changes to a positive specification. If the specification R of the lexical matrix of a given lexeme does not correspond to the specification of the grammatical matrix, it is formally possible to speak of a lexical exception.

If there is a positive R specification in the lexical matrix, where the corresponding grammatical specification is negative, then the lexeme "regulates" the given rule.

The "row" of transformations with specific end-string lexemes had to postulate Lakoff to avoid the strange situation in which each complex string symbol that underwent a structural change would receive an R specification character for the corresponding transformation. If in some cases the order of the whole chain is not assumed, then the grammatically complex symbol of the verb "resemble" in the next sentence, and not just the one with "beat", would have to be marked with a positive specification:

The boy who resembled his dog was beaten severely by the teacher.

Up to this point, Lakoff's position depends on the assumption of the lexical insertion procedure (in Chomsky's sense). But what Chomsky wanted to achieve with the rules of sub-categorization, Lakoff wanted to make possible by generalizing exceptions to the rule in the lexical part of his lexicon. One might think that this should not bother Chomsky, as long as the generalization of the exceptions does not lead to a syntactic, purely lexically described difference between sentences such as:

John wanted me to leave.

John desired for me to leave.

For Lakoff, however, these two movements differ only in that "desire" and "want" govern different transformations. Lakoff would reduce the two propositions to a common basis and thus explain the difference by the influence of one of the lexemes "want" and "desire" which is supported in the lexical insertion procedure in each case.

Chomsky's reaction to generative semantics (and to Fillmore, who is generally involved in generative semantics) was at times very polemical. He dismissed many of the underlying allegations as irrelevant or trivial, “but he had to revise his standard theory fundamentally in the midst of increasing criticism”. He invented well-defined deep structures as the only semantic basis. The interpretations served to avoid the problem of quantifiers, to avoid quantifiers and other problematic elements that seemed to argue against the meaningful transformation theory:

Base: $(P_{1,\ \dots\ ,}\ P_1)$ (P_1 the K-initial, P_1 the post-lexical (deep) structure of the syntactic structure which is a member of K)

transformations: $(P_1, \dots, P_n)$ (P_n the surface structure; $(P_1, \dots , P_n) \in K$)

phonology: $P_n \rightarrow$ phonetic representation

semantics: $(P_1, P_n) \rightarrow$ semantic representation (the grammatical relations involved being those of P_1, that is, those represented in P_1)

(K here stands for a set of phrase markers).

This item is based on the assumption that intonation is the appearance of surface texture and that the range of logical elements such as negation particles and quantifiers is only for the surface texture level.

Chomsky summarized all grammatical elements influencing meaning, but which cannot be explained in terms of depth of structure, under the elements relevant for "presupposition" and "focus". In his paper "Some Empirical Issues" (Nov. 1970) he writes[27]:

Let us now modify the standard theory so that alongside of the rules of the categorial component (with S as the initial symbol) R, a grammar may have in addition the rules, where: f_1 and P_j are chosen freely from the categories of formal objects that serve as focus and presupposition, respectively:

$S' \rightarrow S\ f_1\ P_j$.

Thus the initial phrase marker P_1 generated by the categorial component of the base will be of the form, instead of PP_1, as before:

S'

↓ ↓ ↓

P_1 f_1 P_1

Chomsky describes the effect of this modification as follows:

Transformations now apply as before, paying no attention to f_i, P_j. Thus instead of $\sum = (P_1, \ldots, P_n)$ we have:

$$\sum{}' = (P_1\, f_i\, P_j,\ P_2\, f_i\, P_j,\ \ldots,\ P_n\, f_i\, P_j),$$

where each term $P_k\ f_i\ P_j$ is a phrase-marker. Finally, we add an output condition C that accepts $\sum'$ as well-formed just in case P_n, f_i and P_j are related by the rule R. This revision of the standard theory (call it "version II" of EST [= extended standard theory]) preserves the condition of the standard theory that deep structures alone determine semantic interpretation.

It does not appear to be able to solve Chomsky's procedure easily, here is a suitable example:

My dog had puppies.

* *My female dog had puppies.*

* *My she-dog had puppies.*

My bitch had puppies.

Perhaps, however, such problems can be eliminated by one of the inventories of features in the sense of Chomsky's remarks on Nominalizations (1968). In any case, the semantic meta-conditions would always have to remain in the speculative and informal realm. Some scientists have wondered how the formal basis of Chomsky could be narrowed down so that the transformations had a slightly less overall effect.

As can be seen from the previous considerations, many of the key basic questions of linguistics remain open. Despite the works of Katz, Fodor, Postal, etc., the semantic element of Chomsky's standard theory can still be viewed relatively simply as an empty array. Semantic universals of the kind required for the interpretation of syntactic foundations are not easy to find. Generative semantics has many problems with interpretive semantics without the need for constructivist discovery procedures. Even more problematic is elaborating the area to answer the very general question of whether the necessary meta-conditions must be semantic or syntactic. It is a bit complicated to then gradually formulate an initial set of constructive terms. Attempting to extend Chomsky's base to some extent so that some problems are resolved at the deepest level can also be undertaken relatively easily. However, in the field of generative grammar, future research must be largely empirical. In any case, it is necessary to read natural languages thoroughly (in the Fillmore sense). It is necessary to move away from the assumption of an "internalized system of an idealized speaker", and more to express the statements of real speakers and listeners in natural speaking situations.

Chapter 18. Natural language and gettering of information

Today the artificial intelligence – (AI) it is no longer science fiction for us. It is present almost every day and everywhere – starting from simple applications such as marking photos on FB or converting speech to text, machine translation (automatic) through online fraud detection, to breakthrough solutions in medicine, such as predicting injuries or even the most modern autonomous cars.

More and more common data storage in the so-called "Cloud", decreasing costs of data processing and almost ubiquitous access to the Internet – causes, inter alia, the generation of millions of data terabytes every day, the possibility of analyzing and drawing appropriate conclusions and creates great opportunities for further development. Cognitive technologies play a huge role in this matter, as tools helping to extract valuable information on the basis of analysis with the use of appropriate algorithms.

Appropriately created entire cognitive systems make it possible to organize and manage this huge amount of information/data. In addition, these systems can also be programmed with the help of appropriate algorithms for the so-called "Independent learning and prediction" which information may be of interest to a given recipient.

Cognitive systems are therefore a specific response to the growing amount of data generated, for example, by social media or by sensors located, for instance, in machines and vehicles.

Cognitive systems include such data processing methods as: advanced data analysis, natural language processing, as well as machine learning. One of the main tasks of cognitive software is therefore to support interpersonal communication and decision-making processes by ensuring greater accuracy and speed. By processing data from various sources and reports, cognitive systems more and more support decision-making in management, or also support process management through properly programmed so-called Workflow.

On the other hand, the example of the most dynamically developing machine learning techniques, which are widely used in business, shows that artificial intelligence has settled in our lives for good.

Machine learning, including one of its varieties – the so-called deep learning is a key method of extending the area of artificial intelligence applications.

Machine learning is primarily about creating such programs that make inference, planning, prediction, and remembering based on appropriate algorithms managing a whole series of various data, i.e. big data, thanks to which the system is able to make the most optimal (rational) decision. However, we can only talk here about the issues of quick accounting and calculation in a systematic way.

However, it should be remembered that the role of man does not remain completely eliminated, he should primarily supervise the course of the process and prevent it in non-standard situations.

The use of cognitive technologies is also the basis of the idea of Industry 4.0, which assumes creating a given product first in virtual reality and only then in the real world, thanks to which it can be properly modeled, configured, tested and checked in accordance with the needs of a given customer. As a result, it is possible to produce products even in single quantities while remaining cost effective. It is not possible at all without the use of integrated IT systems. Machine learning and big data analysis tools also play a huge role here. And only advanced cognitive technologies can process them. For data management, however, you also need an appropriate infrastructure that mimics the operation of almost the human brain, such as artificial neural networks (ANN).

> *ANNs represent modern computational tools that can be applied for the efficient modelling of objects, phenomena, and/or processes where, because of feedback complexity and not yet identified interrelations among the different processes (especially nonlinear processes) in different scales, the application of classical regression methods provides unsatisfactory results. In such cases, an ANN trained only based on a representative set of input-output data, can establish the data processing structure required to appropriately correlate the set of inputs and corresponding outputs, e.g. the effect(s), representing by themselves, hidden patterns of interrelations*.

Research on artificial neural networks (ANN) began in neurophysiological and biocybernetics research. The first DNA model of a neuron was officially made by Warren McCulloch and Walter H. Pitts in 1943. This description inspired later research into artificial neural networks.

Currently, ANNs are used, among others, in cognitive science to recognize images, patterns, sounds, speech, writing, data processing, forecasting and modeling various phenomena and processes. It can be said that the structure of AR's activity is very similar to the human brain. The universality of neural networks allows them to be effectively used in management, economy, finance, technology, medicine, geology, mathematics and computer science.

Artificial neural networks are defined as combinations of elements called artificial neurons that make up at least three layers: input, hidden and output, with multiple hidden layers.

Artificial neural networks are modern computational tools that can be used to efficiently model objects, phenomena and processes in which, due to the complexity of feedback and as yet unidentified relationships between various processes, it is difficult to perform calculations in a traditional way. In such cases, the ANN can only, based on a representative set of input and output data, establish the data processing structure required to properly correlate the input data set and the corresponding output data.

Different types of neural networks and their structures can be applied and recommended for various computational and modeling tasks. The most important aspect of ANN structure is the network architecture appropriate to the problem – that is, allowing other optional network-based computations (e.g. optimization). This applies in particular to the number of layers of neurons, the number of general neurons and their distribution.

An important feature of the AR model is its ability to learn and to generalize the acquired knowledge. Two basic types of learning can be distinguished here:

- supervised learning – also known as "teacher learning", which consists of comparing the network output signal with the known correct answers. In this case, the human presents the correct answers and the ANN creates the appropriate structure of the neurons;
- unsupervised learning, also known as "teacherless learning", in which the network has to create its own categories based on dependencies in its input data in order to correctly recognize the input signals.

Appropriate "training" of the artificial neural network is a key element of its proper functioning. An excessively trained network may generalize the results too much, which means that its operation will be similar to classical algorithms. On the other hand, an insufficiently trained network will make too many mistakes, making its usefulness negligible.

Chapter 19. NLP and the Evaluation of NLP

The history of computer understanding of text goes back to H. P. Luhn and concerns the automatic creation of literature summaries by IBM[35]. The successive evaluation assessments dedicated to the program's understanding of the message seem to be attempts from 1987[36]. Next, a project called 'Parseval / GEIG' com-pared the grammars of different languages with a phrase structure[37]. On the other hand, in the next project, "Tipster", series focused on tasks such as summarizing, translating and searching[38]. In 1994, the German markers were compared in the "Morpholympics" project in Germany. Then, the so-called Senseval and Romanseval campaigns aimed at semantic clarification of meanings. In 1996, Sparkle's evaluation compared parsers in four different languages (English, German, French and Italian). And a year later in France, the Grace project compared collections of 21 tags for the French language[39]. In 2004, during the project "Technolangue/Easy" 13 French parsers were compared.

Another large-scale evaluation of dependent parsers took place in 2006 and 2007 on the occasion of joint assignments at the Conference on Computational Natural Language Learning (CoNLL). In 2007, an evalita campaign was launched in Italy to compare different tools for an Italian website of the same name. In France, during the ANR-Passage project (at the end of 2007), 10 French parsers were compared on the website (http://atoll.inria.fr/passage/)[40].

Depending on the assessment procedures, there are several types of NLP assessment:

- internal and external;
- automatic and manual;
- and an assessment based on the principle of black box and glass box.

Internal evaluation examines a single NLP system and determines its performance primarily based on standard scores which are predefined by the evaluators. External evaluation, also known as in-use evaluation, instead examines the NLP system with more complex settings, either as built-in or to perform specific functions for the user. The external performance of the system is then determined based on usability, taking into account the overall task assigned by the complex system or user.

For example, considering a parser that is based on the results of new part-of-speech tags, on internal evaluation, the part-of-speech tag is triggered with respect to

the specified data and compares the system data obtained by the tag to the correct standard data. During external evaluation, the analyzer is started with a different part-of-speech tag, then a new tag, and then the accuracy of the analysis must be compared.

---

[35] Luhn, HP. (1958), *The Automatic Creation of Literature Abstracts'*, IBM Journal; April 1958, 159-165.

[36] Cf. Pallet D.S. (1998), *The NIST role in automatic speech recognition benchmark tests*, LREC, Granada.

[37] Cf. Black E.,Abney S.Flickinger D., Gdaniec C., Grishman R., Harrison P., Hindle D., Ingria R., Jelinek F., Klavans J. Liberman M., Marcus M., Reukos S., Santoni B., Strzalkowski T. (1991), *A procedure for quantitatively comparing the syntactic coverage of English grammars*, DARPA Speech and Natural Language Workshop.

[38] Cf. Hirshman L. (1998), *Language understanding evaluation: lessons learned from MUC and ATIS*, LREC, Granada.

[39] Cf. Adda G., Mariani J., Paroubek P., Rajman M. (1999), *L'action GRACE d'évaluation de l'assignation des parties du discours pour le français*, Langues vol-2.

[40] Cf. ANR MDCA Passage. Available at: http://atoll.inria.fr/passage/ [Accessed 27. Jul. 2020].

In many cases, automated procedures can be defined to evaluate the NLP system by comparing the results to a desired standard. Although the costs associated with the development of the standard itself can be quite high, the automatic evaluation can be repeated as often as needed without significant additional costs (with the same input). However, for many problems with NLP, setting a standard is a difficult task and may not be feasible when internal tag agreement is insufficient. Thus, manual assessment is performed by assessors who are responsible for determining the quality of the system or, most often, the quality of the results generated by the system based on the established criteria. Although the assessors may be regarded as a source of information on many tasks related to linguistic analysis due to their linguistic competence, there may also be discrepancies among their assessments. This is why automatic evaluation is often considered an objective evaluation as people tend to be more subjective.

The black box assessment, on the other hand, involves running an NLP system on a specific database and measuring parameters related to the quality of the process, such as: speed, reliability, resource use and, most importantly, the quality of the result, such as the accuracy of data determination and translation fidelity. Glass-based assessment involves checking the system design, introduced algorithms, used language resources, such as word count, etc. Given the complexity of NLP-related problems, it is often difficult to predict the performance of a given system based solely on the glass-based assessment box. However, this type of assessment provides more information regarding the analysis of errors or the future direction of the system.

Chapter 20. Tasks and challenges of NLP

Natural language analysis is theoretically one of the very attractive ways of "communicating with the computer". Initially, the world of scientists was overwhelmed with enthusiasm due to the initial work on simple word juxtapositions and in terms of limited language. However, enthusiasm quickly faded as the system collided with more real situations and the true complexity and ambiguity of words or phrases.

The very definition of understanding is a big problem in natural language analysis. Recognition of a natural language requires a broad knowledge of the outside world and the ability to transform it. For this reason, understanding natural language is often called the AI-complete problem, which is that in order to properly understand the human world and language, AI should perceive it in the same way as humans.

In most spoken languages, sounds are presented as consecutive and appropriately juxtaposed letters. Therefore, converting an analog signal to symbols is a rather complicated process. In natural speech, there are also quite few gaps between consecutive words. Where a given break occurs depends mainly on both the semantics and the grammar of the language, as well as the context of the sentence.

Some languages, such as Japanese, Chinese, and Thai, do not have specific word boundaries in the written language. At the same time, as you know, each syntactic analysis requires the marking of such boundaries, which is often a very complicated challenge.

It is well known that many words have more than one meaning. In such situations, therefore, you need to choose the word or phrase that best fits the whole context.

The grammar of natural language itself is ambiguous. Usually, there are many possibilities to parse a given sentence. Choosing the best option therefore requires more information about the context itself, or even semantics (meaning). A very important problem of syntactic ambiguity are often unclear sentence boundaries.

On the other hand, an obstacle may also be a foreign or regional accent or a speech defect, or even errors in the text, such as typos and incorrect grammar, or optical character recognition (OCR).

Often the interlocutor himself may perceive a given sentence as some kind of stimulus for a given reaction. The sentence structure itself may not contain enough information to identify this reaction. For example, a given question may be the sender's request for action by the recipient. This answer can be expressed both verbally and through facial expressions or some kind of mixture of both.

The question "Can you give me your name?" requires a simple answer (name). Instead, the question: "Can you pass me the flour?" is actually a request for physical activity. The answer "Yes, I can" without the accompanying activity is not appropriate (although the answer "No" or "No, because it is finished" fully explains the lack of any activity).

In turn, when we consider slightly longer sentences within the grammar of a given natural language, it turns out that most of them will not be unambiguous and may be analyzed in many different ways. To overcome such problems as those mentioned earlier, statistical analysis of natural language uses stochastic, probability and statistical methods. In addition, in such situations, also use is made of language corpora (sets of texts for linguistic research, such as determining the frequency of occurrence of word forms, syntactic constructions, contexts in which given words appear)[41], or Markov chains (sequences of events in which the probability of a given event depends only on the outcome of the previous event)[42].

In summary, NLP uses all available means to automate linguistic analysis, including probability models (e.g. the n-gram model[43]), information theories and linear algebra. This technology is derived primarily from machine learning and rapid data analysis, both of which are branches of artificial intelligence.

The main programs supporting NLP, working with texts and information collections are:

- Anaphora Resolution – this is a program that finds what a given anaphora refers to in the text;
- Automatic summarization – a computer program that automatically summarizes the text, containing the most important information in the summary;
- Foreign language reading aid – a program that helps in reading a foreign language;
- Foreign language writing aid – a program supporting writing in a foreign language;
- Information Extraction (IE) – a program that analyzes unstructured text in order to extract information from it; the program extracts facts from documents and the user analyzes them;
- Information Retrieval (IR) – a program that collects and retrieves information[44];
- Named Entity Recognition (NER) – is one of the parts of IE, it indicates which words in the text are proper names (in some languages not only proper names are written in capital letters, but also all nouns, such as in German);
- Natural language generation (NLG) – it is used to translate data from computer language into natural language;
- Natural language understanding – deals with the understanding of natural language by the computer;
- OCR – software for recognizing graphic characters, e.g. from scanned text;

[41] Linguistic corpora are widely used in contemporary lexicography. They are also used as training and test datasets in machine learning methods used in natural language processing. Some corpuses are sometimes referred to as balanced – this means that the text samples for the corpus were selected according to a special key so as to ensure the desired proportions between different styles or periods of text creation.

[42] Cf. Podgórska M. and Others (2002), *Łańcuchy Markowa w teorii i zastosowaniach*, Szkoła Główna Handlowa, Oficyna Wydawnicza, Warszawa.
And cf. Iwanik A., Misiewicz J.K. (2009), *Wykłady z procesów stochastycznych z zadaniami. Cz. 1, Procesy Markowa*, Zielona Góra: Oficyna Wydawnicza Uniwersytetu Zielonogórskiego.

[43] The n-gram model is the language model used in speech recognition. N-grams are based on statistics and are used to predict the next element in a sequence. They are mainly used for words as well as for example phonemes. Cf. Jurafsky D., Martin J.H. (2008), *Speech and Language Processing, 2nd Edition*, Prentice-Hall, Inc., New Jersey; And cf. Ziółko B., Skurzok D. (2011), N-*grams model for Polish,* [in:] *Speech and Language Technologies, Book 2*, InTech Publisher;
And also cf. Ziółko B., Ziółko M. (2011), *Przetwarzanie mowy,* Wydawnictwa AGH, Kraków.

[44] IR – it is a separate department in computer science (more similar to databases), however, dependent on NLP methods (e.g. reduction). Some contemporary research tries to bridge the gap between IR and NLP. The IR software finds a collection of documents and the user analyzes them.

- Question answering – a program that answers questions asked in that language using natural language. These can be questions with one answer (such as "What city is the capital of Poland?") Or with many (such as "What is the meaning of life?");
- Speech recognition – this program can save the text of a conversation from an audio file (it is the inverse of text-to-speech processing, the inverse of speech synthesis, i.e. an operation consisting in processing text into speech);
- Spoken dialogue system – it is a system that recognizes speech and has a speech synthesis module;
- Text simplification – this is software that simplifies the grammar and structure of the text.

At this point, we should also mention various types of computer translators – programs that automatically (machine) translate text from one language to another, as well as commonly used software that corrects the text, that is, marks errors in written text.

Chapter 21. Selected problems of the natural language understanding system

It is worth paying attention to the fact that currently it is almost impossible to transcribe a spoken language completely without human intervention. This takes place, for instance, during the preparation of transcripts of conversations, discussions, when, for example, people speak simultaneously. In such situations, programs "get lost" and are not able to adequately reflect the meaning of the dialogue. Of course, people who deal with this type of translation use the available software, but they have to correct the written text to convey the right meaning. Currently, this is the case, for example, when translating conversations shared on YouTube channels.

Moreover, sentences like "We gave the animals fruit because they were hungry" and "We gave the animals fruit because they were old" have basically the same grammatical structure. However, the verb "were" is used in one sentence to refer to animals and the other to fruit. Without general knowledge of a given situation, it is impossible to say in which case the verb "were" refers to which noun. Such a string of words can be interpreted differently.

For example, the English phrase "Time flies like an arrow" can be understood in several ways:

First, it may be the traditional comparison that time flies so fast like an arrow.

In addition, the word "flies" in English also means flying insects, and the noun "time" can also be a verb. Therefore, the meaning of the expression above can also be interpreted as: "(You should) time flies as you would (time) an arrow" (in this case, the expression is read in the imperative).

With this assumption regarding the meaning of the words "time" and "flies", the analyzed statement can also be understood as "Time flies in the same way that an arrow would (time them)".

Another meaning relates to the comparison of flies and arrows: "Time those flies that are like arrows".

"Time-flies" can also be read as kinds of flying insects, which in the sentence above would mean that all insects of this kind like a single arrow or fruit flies like an apple.

"Time-flies" in the above sense can also mean that each type of flying insect likes a different arrow separately.

Another meaning may be that a specific object, such as a Time magazine, flies through the air in a manner similar to an arrow.

Analytical natural languages, such as modern English, which have practically no inflectional morphology to help distinguish between parts of speech, are particularly demanding in this regard.

In English and in some other languages, there is no indication of which word the adjective refers to. An example would be the phrase "pretty little girls" school.

Does "little" refer to "girls"?

Does "little" refer to "school"?

Does "pretty" refer to "girls"?

Does "pretty" refer to "school"?

In spoken language, additional information is often suggested by accentuating individual words.

For example, the sentence "I never said she stole my money" shows how important the role of stress in a sentence is and indicates the significant issue that the natural language analyzer has to deal with when performing parsing. Depending on what word is stressed by the speaker, this sentence can have several different meanings:

"I never said she stole my money." – I just didn't say it.

"I never said she stole my money." – I just said that she probably only borrowed the money.

"I never said she stole my money." – I might have suggested it in a different way, but I never said it directly.

"I never said she stole my money." – I said she stole someone's money.

"I never said she stole my money." – I said she stole something but not my money.

"I never said she stole my money." – I said someone took the money, but I didn't say it was her.

"I never said she stole my money." – someone else said it, but not me.

The purpose of the efficient functioning of NLP is therefore to measure one or more features of a given algorithm or system to determine whether and to what extent this system meets the assumptions of its creators and whether it meets the needs of its users. Research on the evaluation of NLP is currently very popular, because determining the appropriate evaluation criterion remains the only way to precisely define contemporary NLP problems beyond the general nature of tasks defined as understanding or generating language without additional human intervention. The exact set of evaluation criteria, which mainly includes its data and metrics, allows you to compare the solutions for a given NLP analysis problem.

Chapter 22. Information retrieval and data mining techniques

Information retrieval may be understood as finding in a large collection of unstructured documents (usually expressed as natural language text) documents that meet a specific information need. There are a number of applications of search systems, such as searching for legal acts, searching for books in a library or searching for websites[45].

Data mining, data acquisition and data extraction – are some steps in the process of obtaining knowledge from databases. The idea of data mining is to use the speed of a computer to find hidden (precisely because of the limited time possibilities) regularities in the data stored in databases.

There are many data mining techniques that come from well-established fields of science, such as statistics (statistical multivariate analysis) and machine learning.

Techniques and methods for data mining derive mainly from the field of artificial intelligence research. The main examples of applied solutions belong to the following areas:

- evolutionary methods;
- fuzzy logic;
- machine learning methods;
- neural networks;
- rough sets;
- statistical methods;
- visualizations on charts.

In data mining, various processing methods are developed, plants are tabulated, using banded algorithms, and the results presentation mode. Among them stand out:

- classification;
- grouping;
- qualitative data analysis;
- quantitative data analysis;
- searching for associations;
- summaries.

There are many areas of data mining application, they include those places where information systems are used, inter alia, to collect the obtained data in the form of databases. We are witnessing a real explosion of databases in terms of their number and volume. Huge data collections are stored in databases. Due to the simplicity of database construction and acceptable prices, data collection systems are used in almost all areas of life. However, wherever a database already exists, there is a need to analyze this data in order to discover previously unknown knowledge. The fields in which data mining is widely used are astronomy, economics, medicine, technology, broadly understood business and others.

> *Knowledge Discovery in Databases (KDD) refers to the broad process of finding knowledge in data, and emphasizes the "high-level" application of*

[45] Cf. Manning, Ch. D.; Raghavan, P.; Schütze, H. (2008): Introduction to Information Retrieval. Cambridge University Press.

particular data mining methods. It is of interest to researchers in machine learning, pattern recognition, databases, statistics, artificial intelligence, knowledge acquisition for expert systems, and data visualization.

The unifying goal of the KDD process is to extract knowledge from data in the context of large databases.

It does this by using data mining methods (algorithms) to extract (identify) what is deemed knowledge, according to the specifications of measures and thresholds, using a database along with any required preprocessing, subsampling, and transformations of that database[46].

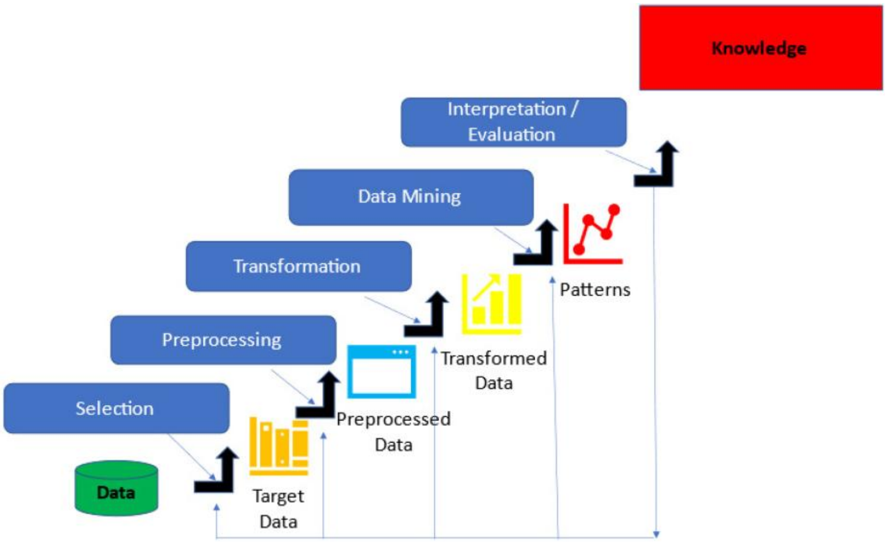

Figure 4: An Outline of the Steps of the KDD Process,
own figure based on source from: Fayyad, Piatetsky-Shapiro, Smyth, 1996, p. 4.

KDD is used for:

- analysis of banking operations;
- creating targeted advertisements;
- credit risk assessment;
- customer segmentation;
- designing a data warehouse;
- fraud detection;
- genetic research;
- internet traffic data mining;
- recognition of image, speech and handwriting signals;

46 Fayyad, Piatetsky-Shapiro (1996), *From Data Mining to Knowledge Discovery: An Overview*, [in:] Fayyad, Piatetsky-Shapiro, Smyth, Uthurusamy, *Advances in Knowledge Discovery and Data Mining. AAAI Press/The MIT Press*, Menlo Park, CA, pp.1-34, here p. 4.

- sales forecasting;
- supporting medical diagnostics.

An example may be the extraction of a dependency in the sales data of a given product, consisting in the fact that customers who have purchased a given product also usually buy other specific products.

Examples of data visualization are:

- box plots, the aim of which is to visualize the properties of individual features in order to select those variables that are characterized by the greatest shifts in quantiles, maximum and minimum values and medians;
- charts of empirical distributions for selected variables;
- histograms – for selected predictive variables.

The following data can be read from the charts:

- density;
- feature value;
- frequency;
- maximum;
- median;
- minimum;
- mutual correlation of variables;
- number;
- quartile.

Summary

Today's translations are promoted by modern programs to support the translation process.

Machine translation systems are applications or online services that use machine learning technologies to translate large amounts of text from and into one of their supported languages. The service translates a "source" text from one language into another "target language".

Although the concepts behind machine translation technology and the interfaces to use it are relatively simple, the science and technology behind them are extremely complex and bring together several cutting-edge technologies, in particular deep learning (artificial intelligence), big data, linguistics, cloud Computing and web APIs.

Since the early 2010s, a new artificial intelligence technology, deep neural networks (alias deep learning), has enabled speech recognition technology to achieve a level of quality that has enabled the Microsoft Translator team to introduce speech recognition with its core text translation technology (new language translation technology).

Historically, the primary machine learning technique used in the industry has been statistical machine translation (SMT). SMT uses advanced statistical analysis to estimate the best possible translations for a word that are in the context of a few words. SMT has been used by all major translation service providers, including Microsoft, since the mid-2000s.

The advent of neural machine translation (NMT) led to a radical change in translation technology, which led to a much higher quality of translation. This translation technology started for users and developers in 2016.

Both SMT and NMT translation technologies have two elements in common:

1. Both require large amounts of pre-human translated content (up to millions of translated sentences) to train the systems.
2. They do not act as bilingual dictionaries, but rather translate words based on a list of potential translations, according to the context of the word used in a sentence.

References:

Adda G., Mariani J., Paroubek P., Rajman M. (1999), *L'action GRACE d'évaluation de l'assignation des parties du discours pour le français*, Langues vol-2.

Anon., *Űber Herrn Kempelens Schachspieler und Sprechmaschine,* [in:] "Der Teutsche Merkur", November 1784.

ANR MDCA Passage. Available at: http://atoll.inria.fr/passage/ [Accessed 27. Jul. 2020].

Babik W., *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*, Stowarzyszenie Bibliotekarzy Polskich, s.: Nauka, Dydaktyka, Praktyka, Warszawa 2002.

Babik W., *Język naturalny w wyszukiwaniu informacji. Problemy jego przetwarzania*, 2013, Available at: http://bbc.uw.edu.pl/Content/1872/z2013_1_03.pdf [Accessed 27. Jul. 2020].

Bar-Hillel Y., *Hussrl's Conception of a Purely Logical Grammar,* [in:] *Philosophy and Phenomenological Research*, 17, 1957.

Bates M., *Models of natural language understanding,* [in:] „Proceedings of the National Academy of Sciences of the United States of America", Vol. 92, No. 22 (24.10.1995), 1995.

Black E., Abney S., Flickinger D., Gdaniec C., Grishman R., Harrison P., Hindle D., Ingria R., Jelinek F., Klavans J. Liberman M., Marcus M., Reukos S., Santoni B., Strzalkowski T., *A procedure for quantitatively comparing the syntactic coverage of English grammars,* DARPA Speech and Natural Language Workshop 1991.

Bloomfield L., *Language*, New York, 1933.

Brücke E., *Grundzüge der Physiologie und Systematik der Sprachlaute*, Vienna 1856.

Charniak E., *Introduction to artificial intelligence*, p. 2. Addison-Wesley, Boston 1984.

Chomsky A.N., *Cartesian Linguistics,* New York 1966.

Fayyad, Piatetsky-Shapiro, *From Data Mining to Knowledge Discovery: An Overview,* [in:] Fayyad, Piatetsky-Shapiro, Smyth, Uthurusamy, *Advances in Knowledge Discovery and Data Mining. AAAI Press/The MIT Press*, Menlo Park, CA, 1996.

Fries Ch., *The Structure of English: An Introduction to the Construction of English Sentences*, New York 1952.

Gardner T., *Hauptstömungen der modernen Linguistik*, Göttingen 1973.

Gross M. Lentin A., *Mathematische Linguistik,* Berlin-Heidelberg, 1971, pp. 95 f and p. 114 f. Cf. also Maurer H., *Theoretische Grundlagen der Programmiersprachen*, Mannheim, 1969.

Hand D., Mannila H., Smyth P., *Eksploracja danych*, WNT, Warszawa 2005.

Harris Z., *Methods in Structural Linguistics,* Chicago, 1951.

Harris Z., *Discourse Analysis,* [in:] *Papers in Structural and Transformational Linguistics*, Dordrecht, 1970.

Hirshman L., *Language understanding evaluation: lessons learned from MUC and ATIS*, LREC, Granada, 1998.

Iwanik A., Misiewicz J.K., *Wykłady z procesów stochastycznych z zadaniami. Cz. 1*, *Procesy Markowa*, Zielona Góra: Oficyna Wydawnicza Uniwersytetu Zielonogórskiego, 2009.

Język naturalny, [in:] *Słownik języka polskiego* [online], PWN, Available at: https://sjp.pwn.pl/szukaj/j%C4%99zyk%20naturalny.html [Accessed 27. Jul. 2020].

Jurafsky D., Martin J.H., *Speech and Language Processing, 2nd Edition*, Prentice-Hall, Inc., New Jersey 2008.

Kempelen W. v., *Wolfgangs von Kempelen k. k. wirklichen Hofraths Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine/Le mécanisme de la parole, suivi de la description d' une machine parlante et enichi de XXVII planches*, Vienne 1791.

Koenig W. et al., *The sound spectrograph*, JASA 17, 1946.

Köster J.-P., *Historische Entwicklung von Syntheseapparaten mit Erzeugung statischer und vokalartiger Signale nebst Untersiuchungen zur Synthese deutscher Vokale*, [in:] von Essen O., Gutknecht Ch., Köster J.-P. (all Ed.), *Hamburger Phonetosche Beiträge – Untersuchungen zur Phonetik und Linguistik*, Vol. 4, Hamburg, 1973.

Kuzior A., Kwiliński A., Tkachenko V., *Sustainable development of organizations based on the combinatorial model of artificial intelligence*, Entrep. Sustain., 2019, Vol. 7 iss. 2, pp. 1353-1376, DOI: 10.9770/jesi.2019.7.2(39).

Kuzior A., Postrzednik-Lotko K., *Natural language and gettering of information*, [in:] *Sustainable economic development and advancing education excellence in the era of global pandemic. Proceedings of the 36th International Business Information Management Association Conference (IBIMA)*, 4-5 November 2020, Granada, Spain. Ed. Khalid S. Soliman. International Business Information Management Association, s. 13479-13486.

Kwiliński A., Tkachenko V., Kuzior A., *Transparent cognitive technologies to ensure sustainable society development*, 2019, DOI: J. Secur. Sustain. Issues. Vol. 9 No. 2, pp. 561-570.

Larose D.T., *Odkrywanie wiedzy z danych*, Wyd. Nauk. PWN, Warszawa 2006.

Liénard J.S., *La machineparlante de Kempelen*, 1968.

Luhn H.P., *The Automatic Creation of Literature Abstracts'*, IBM Journal; April 1958, 159-165.

Lyons J., *Natural Language and Universal Grammar*, Cambridge University Press, New York 1991.

Machej K., Piotrowski K., *Application of artificial neural network for description of mass crystallization process*, [in:] "14th International Symposium on Industrial Crystallization". Cambridge 1999, CD: theme 3, article 0006.

Manning Ch.D., *Hinrich Schutze Foundations of Statistical Natural Language Processing*, MIT Press, 1999.

Manning Ch.D., Raghavan P., Schütze H., *Introduction to Information Retrieval*, Cambridge University Press, 2008.

Nowak I., *Przemysł 4.0 – rozwój technologii identyfikacji, lokalizacji i komunikacji dla czujników i systemów kognitywnych*, [in:] "Logistyka", 2018, No. 1, pp. 71-72.

Paget S.R.S., *Human speech*, London 1930.

Pallet D.S., *The NIST role in automatic speech recognition benchmark tests*, LREC, Granada 1998.

Piotrowski K., Piotrowski J., Schlesinger J., *Modelling of urea synthesis process equilibrium with the use of artificial neural networks*, [in:] "Polish Journal of Applied Chemistry", 1999, XLIII, 1-2, pp. 69-76.

Podgórska M. and Others, *Łańcuchy Markowa w teorii i zastosowaniach*, Szkoła Główna Handlowa, Oficyna Wydawnicza, Warszawa 2002.

Russel B., *My Philosophical Development*, London 1959.

Sakiewicz P., Piotrowski K., Ober J., Karwot J., *Innovative artificial neural network approach for integrated biogas – wastewater treatment system modelling: Effect of plant operating parameters on process intensification*, [in:] "Renewable and Sustainable Energy Reviews", 2020, Vol. 124, pp. 1-14, DOI: Available at: https://doi.org/10.1016/j.rser.2020.109784 [Accessed 27. Jul. 2020].

Sapir E., *Language: An Introduction to the Study of Speech*, New York 1921.

de Saussure F., *Course de Linguistique Générale*, Paris 1916.

Sieci neuronowe, Available at: http://kognitywistyka.uwb.edu.pl/component/k2/item/406-sieci-neuronowe, [Accessed 27. Jul. 2020].

Słownik języka polskiego [online], PWN, Available at: https://sjp.pwn.pl/ [Accessed 27. Jul. 2020].

Thagard P., *Cognitive Science*, [in:] *Stanford Encyclopedia of Philosophy*, Available at: https://plato.stanford.edu/entries/cognitive-science/ [Accessed 27. Jul. 2020].

Willis R., *On the vowel sounds, and on reed organ pipes*, Cambridge, 1830.

Ziółko B., Skurzok D., *N-grams model for Polish*, [in:] *Speech and Language Technologies, Book 2*, InTech Publisher, 2011.

Ziółko B., Ziółko M., *Przetwarzanie mowy*, Wydawnictwa AGH, Kraków 2011.

European Funds
Knowledge Education Development

Republic of Poland

NAWA
POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE

European Union
European Social Fund

Research, preparation of materials and preparation of the textbook were carried out under the project – grant no. PPI/KAT/2019/1/00015/U/00001 "Cognitive technologies – second-cycle studies in English" and were carried under the KATAMARAN program Polish National Agency for Academic Exchange (NAWA). The program is co-financed by the European Social Fund under the Knowledge Education Development Operational Program, a non-competition project entitled "Supporting the institutional capacity of Polish universities through the creation and implementation of international study programs" implemented under Measure 3.3. Internationalization of Polish higher education, specified in the application for project funding no. POWR.03.03.00-00-PN 16/18.

Wydawnictwo TYGIEL

www.wydawnictwo-tygiel.pl

ISBN 978-83-66489-54-7

9 788366 489547