Palacký University Olomouc
Faculty of Arts
Department of Asian Studies

MA THESIS

Describing and Assessing Linguistic Distance
The Case of Czech and Korean

OLOMOUC 2024 Patrícia Bohušová

Supervisor: Dr. Andreas Schirmer

I declare that I have written my thesis independently and have listed all sources and literature used.

Olomouc, 21 June 2024

……..

Annotation

Author:	Bc. Patrícia Bohušová
Name of the Faculty and Department:	Faculty of Arts – Department of Asian Studies
Title of the thesis:	Describing and Assessing Linguistic Distance: The Case of Czech and Korean
Supervisor:	Dr. Andreas Schirmer
Number of pages:	59
Number of characters:	118 554
Number of sources:	59 sources (internet sources included)
Key words:	*linguistic distance, Czech, Korean, English, linguistic similarity*

This master's thesis is focused on describing and assessing linguistic distance between Czech and Korean while using English as a bridge between the two. First, the thesis explains linguistic distance and its nature. It describes how the distance can be perceived between multiple languages and what determines the length or closeness of this distance. It introduces the three studied languages within this thesis and illustrates different factors that influence their linguistic systems. Then, Korean, Czech, and English are being assessed within their respective language families. The thesis describes relevant historical backgrounds and linguistic developments. Subsequently, each language is examined in relation to different linguistic dimensions. The languages are then compared in pairs within one specific dimension, that is phonology. Various common attributes and contrasting characteristics are presented and compared between the language pairs. Finally, other methods for assessment of linguistic distance are briefly addressed. The thesis evaluates several different approaches to describing linguistic distance while applying this knowledge on the language pair in the main focus, that is Czech and Korean.

Anotácia

Autorka:	Bc. Patrícia Bohušová
Názov fakulty a katedry:	Filozofická fakulta, Katedra ázijských štúdií

Názov práce:	Opis a posudzovanie lingvistickej vzdialenosti: prípad češtiny a kórejčiny
Školiteľ:	Dr. Andreas Schirmer
Počet strán:	59
Počet znakov:	118 554
Počet použitých zdrojov:	59 zdrojov (vrátane internetových zdrojov)
Kľúčové slová:	*lingvistická vzdialenosť, čeština, kórejčina, angličtina, lingvistická podobnosť*

Táto diplomová práca je zameraná na opis a posúdenie jazykovej vzdialenosti medzi češtinou a kórejčinou, pričom angličtina slúži ako most medzi nimi. Po prvé, práca vysvetľuje jazykovú vzdialenosť a jej povahu. Popisuje, ako možno vnímať vzdialenosť medzi viacerými jazykmi a čo určuje dĺžku alebo blízkosť tejto vzdialenosti. Predstavuje tri študované jazyky v rámci tejto práce a ilustruje rôzne faktory, ktoré ovplyvňujú ich jazykové systémy. Potom sa v rámci príslušných jazykových rodín posudzuje kórejčina, čeština a angličtina. Práca popisuje relevantné historické pozadie a jazykový vývoj. Následne sa každý jazyk skúma vo vzťahu k rôznych jazykovým dimenziám. Jazyky sa potom porovnávajú vo dvojiciach v rámci konkrétnej dimenzie, ktorou je fonológia. Medzi jazykovými pármi sú prezentované a porovnávané rôzne spoločné atribúty a kontrastné charakteristiky. Nakoniec sa stručne spomínajú ďalšie metódy hodnotenia jazykovej vzdialenosti. Diplomová práca vyhodnocuje niekoľko rôznych prístupov k opisu jazykovej vzdialenosti a popri tom aplikuje tieto poznatky na hlavnú jazykovú dvojicu, ktorou je čeština a kórejčina.

Acknowledgements

I would like to thank my supervisor Dr. Andreas Schirmer for his patience and guidance in the process of creating this thesis.

I would also like to thank my family and friends who supported and encouraged me throughout my studies.

Editor's note

The thesis contains Korean expressions with transcription in McCune–Reischauer system in [ ] parentheses, followed by English translations in quotation marks “ ”.

Czech expressions are written in ⟨ ⟩ parentheses.

And vowel and consonant phonemes are written in / /.

Table of contents

1. Linguistic distance

Linguistic distance characterises the mutual relations between all foreign languages to varying degrees. For example, in the case of Korean and Czech (which are also the subjects of this study), just the fact that the two are from two significantly diverse language families suggests that there is a considerable linguistic distance between the languages.

According to Ferdinand de Saussure who pioneered modern linguistics,[1] "language is made up of a collection of units, all related to each other in very particular ways, on different levels. These different levels are themselves related in various ways to each other." The units' mutual connections and relations can be the indicators of the linguistic distance and the level of depth between them. Depending on whether the units of the given languages are considerably closer to each other or further apart, the linguistic distance can be assessed. According to Saussure's theory of Language, the linguistic system of every individual is constructed from experience and this process of construction depends on the associative principles of contrast, similarity, contiguity, and frequency.[2]

The principle of contrast makes linguistic units distinct from each other which helps to avert confusion or interference between them. The principle of similarity captures aspects of rule-governed behaviour by defining classes of interchangeable units which compel combinatory processes. The principle of contiguity describes the formation of more complex units that are created from the combination of simpler units. This principle also enables the integration of the products of the linguistic system into the system in a similar manner as mathematical formulae (products of mathematical systems) become components of mathematical knowledge. Lastly, the principle of frequency allows frequently occurring units to be recognised as stable units. This principle also allows formation of abstract categories if a sufficient number of units share similar properties.

It is crucial to consider the influence of these principles because they allow Saussure's Theory of Language to adopt rule-based and experience-based perspectives on language while giving the theory a certain psychological plausibility. These perspectives can provide a more complex description of linguistic knowledge that in turn allows a better assessment of its differences and similarities in relation to other languages.

---

[1] His seminal *Course in General Linguistics* was published in 1913 by his students.
[2] Chipere 2003.

Language users have the capacity to oscillate between using both rule-based and experience-based perspectives. However, every individual displays different extent to which they utilise these perspectives. Apart from individual preferences and inclinations towards the two, there are certain situations where one perspective may be generally more prevalent than the other. For example, when L1 users learn their native language, they will use experience-based perspective because associative learning is the only learning approach they know at the age. On the contrary, when it comes to L2 users, they mostly use rule-based perspective since it usually involves the language that is being taught at school by rule-following approach. Therefore, when assessing linguistic knowledge, both perspectives need to be integrated based on linguistic and psychological considerations.

The same principles may be applied to a collective linguistic system. If every individual possesses their own linguistic system, the same thinking can be applied on larger scale, meaning every group of the same language users shares a collective linguistic system. If every individual's linguistic system (e.g., within the same culture, country, or native tongue) is as unique as its user, every group's collective linguistic system will be unique as well.

Culture can be one of the important defining factors when it comes to linguistic system. People from same culture usually tend to share similar thinking processes. These thinking processes are then reflected upon every individual's unique linguistic system, and vice versa, every individual's unique linguistic system forms a collective linguistic system of the whole group of people within the same culture.

The influence of culture can be perceived when comparing linguistic differences of the same language that is being used in different locations. The users of language grow up in different cultures and even though they share the same language, it does not inevitably mean their linguistic systems will be the same. Language forms and develops accordingly to its environment in which it is being used. For example, English is a native language of multiple countries, but every country has some slight variations in their collective linguistic system (e.g., accent, intonation, spelling of some words, slang).

The other language that is being assessed in this thesis, Korean, can be examined within the two linguistic systems that are being used in two countries. South Korean linguistic system is heavily influenced by traditional Korean culture, K-pop culture, American pop culture, or consumerism whereas North Korean linguistic system is more

influenced by communism or Juche ideology.[3] These cultural influences are subsequently being reflected in the respective linguistic systems.

Both groups of Korean language users can generally understand each other but there are some differences in the usage of the language that can sometimes cause miscommunication. Korean language used in South Korea is more open to global influence and its trend of merging cultures that can lead to infusion of directly borrowed English words or new words that are being adapted into Korean.[4]

The influence of English is strengthened even more through the South Korean educational system. With the global labour market's requests, the demand after English speaking employees has been growing throughout the world. The South Korean educational institutions, too, react to global market's demands and put a lot of emphasis on teaching English. Therefore, not only the internal influence of culture but also the external influence of other cultures can have effect on a specific linguistic system.

The North Korean linguistic system is not exposed to these influences on the same level. The country is predominantly closed off from the outside world and the cultural influences come mostly from traditional Korean culture, Juche ideology and the recurring themes of martyrdom during revolution, happiness of society, and brilliance of the leader. Instead of incorporating some English vocabulary, North Korean language tends to follow the Pyongyang dialect while occasionally incorporating Russian words (or varied adaptations of the words) into its linguistic system. Overall, North Korea puts heavy emphasis on linguistic purism, consciously controlling the language and purging any foreign influences that could lead to the introduction of new words.

Misunderstandings resulting from communication between South Koreans and North Koreans can be observed when North Korean defectors try to assimilate into South Korean society. According to a study of National Institute of the Korean Language from 2012,[5] North Korean defectors in general understood only half of South Korean. The vocabularies used in daily life were different only in 30–40%, but in professional settings more than 60% of the vocabularies differed.

---

[3] Juche ideology is the state ideology of North Korea that includes the historical materialist ideas of Marxism–Leninism while putting a strong emphasis on the individual, the nation state, and national sovereignty.

[4] English is currently considered as a global lingua franca, so that is the reason for mentioning the result of a global influence being the infusion of English words into the (South) Korean linguistic system.

[5] Language Magazine online, s.v. "Bridging the North and South Korean Language Divide." [Retrieved 23 April 2024.]

As for a more recent example of certain issues in communication, during the 2018 Winter Olympics South and North Korean women's ice hockey teams decided to join forces. Even though the teams generally understood each other, it was necessary to mediate, especially for North Korean players, some English expressions that were implemented into South Korean.[6] Considering a lot has not changed when it comes to the isolation of North Korea from the outside world to this date, it can be assumed that the numbers from the above mentioned 2012 study have not drastically changed.

Lastly, in the case of Czech linguistic system, both external and internal cultural influences play significant role in shaping it.[7] Since Czech is a native language of only one country, the influence of culture on the linguistic system will be examined by comparing different regions. The variations might be even smaller as it was in the previous case, because it still relates to one country.

Standard Czech is an official form of the language but there are other varieties that are part of Czech linguistic system. For example, different regions contain different dialects that reflect the subculture of people living there (most well-known being Bohemian or Moravian dialects).

The Bohemian dialects do not differ excessively from Standard Czech, and they are mostly influenced by dialects of the Prague region. They are relatively homogeneous in terms of pronunciation, and they tend to have more open vowels (compared to Moravian region).[8] According to the 2015 survey,[9] the Bohemian culture is not as distinct as Moravian. Considering Czech Republic's capital is situated in Bohemia, the region is more prone to global cultural influences and mixing of cultures that can lead to disappearance of some regional traditions and customs. This is then reflected in their dialects, too.

Moravian dialects are more widespread and diverse, which is evident thanks to a political movement for Moravian linguistic revival that has been active since the 1990s. This in return confirms some claims from the 2015 survey[10] that users of the Moravian dialects see themselves as more patriotic (in comparison to their Bohemian counterparts). Stronger sense of patriotism, paired with their perceptions of being welcoming and

[6] Los Angeles Times, s.v. "Koreas' unified women's hockey team has exposed a key difference between South and North — their language." [Retrieved 23 April 2024.]
[7] However, the focus will be predominantly on the internal cultural influences.
[8] Šimáčková et al. 2012.
[9] Kafkadesk online, s.v. "Bohemians and Moravians: Two bickering nations under one roof." [Retrieved 23 April 2024.]
[10] Ibid.

friendly are all cultural influences that help conserving not only regional traditions and customs but also Moravian dialects. The warm nature of people is especially prevalent in rural locations where people organise multiple traditional festivals and celebrations. These contribute to preserving of dialect because of masses of people gathering and using the dialect in direct communication. On a larger scale, if the cultural influence on the Czech linguistic system is being examined within the whole country, this influence is manifested in unique proverbs, sayings, idioms, or expressions that come from Czech traditions and customs.

In conclusion, when it comes to measuring linguistic distance, it is important to consider the cultural influences that shape language as well. Taking in consideration different habits and customs of the specific culture is a crucial part in assessing linguistic distance because these aspects are reflected in the collective and individual linguistic systems.

In the case of measuring linguistic distance between linguistic systems of two or more individuals, it is needed to keep in mind particular fluctuations and differences in each linguistic system (e.g., one individual might have wider knowledge of the specific language, other might possess larger vocabulary, or have higher level of language proficiency). Conversely, when measuring distance between collective linguistic systems, these systems tend to be generalised and any individual deviations are neglected.

Linguistic distance can manifest itself (and seemingly increase) during communication when participants in a discussion reveal that they adhere to divergent concepts.[11] If the concepts do not intersect, there is no common point resulting in possible comprehension obstacles and subsequently a further widening of the pre-existing linguistic distance. The cause of this outcome might be the heterogeneity of the respective cultural backgrounds. The cultural connotations of some concepts might be hard to grasp if a similar concept does not exist also in the cultural background of the other side. Since the assessment of the linguistic distance in this hypothetical case happens through real communication between language users, the language is not being evaluated statically but with all the varying factors that influence the perceived and the measurable distance.

Apart from culture, there are many other influences that help to shape language in a specific country such as historical background, religion, politics, societal values and

[11] This applies especially in cases when more complex topics are the subject of communication.

issues. All these factors need to be equally considered when assessing linguistic system and thereafter measuring linguistic distance.

2. Identification of language families and its influence on linguistic distance

Language identification can be considered as an effective way to determine linguistic distance between assessed languages. Language distance and language identification can be viewed as the same dual aspects of the same concept that are interconnected because "the more difficult the identification of differences between two languages is, the shorter the distance between them."[12]

Even though identification of contrasting qualities between languages can be helpful in the assessment of linguistic distance, observing linguistic similarity proves to be just as useful – as can be seen in various branches of linguistics, such as dialectology, historical linguistics, second-language learning (serving as a measure of learners' proficiency), or psycholinguistics (explaining lexical "neighbourhood" effects, where neighbourhoods are specified by similarity).[13]

Using classification of languages into different language families helps identifying shared similarities within larger groups. The entire process of observing linguistic distance within this classification is then simplified because it becomes easier to identify shared similarities as well as slight differences.

In the case of languages coming from the same language family, the units of the observed languages will be situated in closer proximity and some of them may be even mutually shared. On the contrary, the languages that are not parts of the same language family will have much wider linguistic distance between each other since they do not share any units that would be mutually related.

Any assessment of the linguistic distance between two languages based on a system of certain measures is a challenging task, taking in a consideration that each language has its own vocabulary, grammar, syntax, phonetics, and many other defining aspects. These aspects are unique for each language and even though some languages can be intuitively assessed as being closer to each other, or having smaller linguistic distance,[14] how can be

[12] Gamallo 2017.
[13] Nerbonne and Hinrichs 2006.
[14] E.g., languages that sound similar or share many similar expressions.

the distance between the two measured in specific units or a certain amount? Identifying and categorising languages into specific language families can however help with this process. If certain languages share some similarities between one another, this might be because they are a part of the same language family.

Gathering all languages that have descended from a same proto-language within the group of language family makes assessment of linguistic similarity much easier thanks to the shared systematic similarities. The classification makes it evident that certain similarities cannot be attributed to chance or effects of contact with another language (e.g., borrowing, convergence). The comparison of languages in turn becomes easier because similarities within one language family become more prevalent and small individual differences between respective languages[15] become more noticeable.

However, a few challenges might arise in the process of assessing linguistic similarity, too. One of the obstacles can be dialectal variation that can get in the way of recognising the level of similarity between languages. Dialects reflect regional or social variations of a language which can manifest in terms of different pronunciation, vocabulary, and grammar. For example, regional variations in language can be observed within one country (as is the case of Czech)[16] and social variations can be seen in different countries with different social norms and cultural influences (as is the case of English). Therefore, it is crucial to recognise core features that define a language and discern them from the variations occurring within them.

Another challenge takes form of borrowed words and loanwords. Distinct cultural influences, trade globalisation, interconnecting countries through establishment of international relations, or even colonisation in the past – these all contributed to rise of borrowed words in the collective linguistic systems. The problem arises when assessing linguistic similarity because these borrowed words can present lexical similarities between languages, even if their grammatical structures or language families are vastly different.

Lastly, language evolves and changes constantly in time. If there are too many changes (especially the ones that influence the core features that define a language), the assessment of linguistic similarity becomes more challenging. Language evolution

[15] Differences that otherwise might have become lost if they were being observed together with a language from vastly different language family.
[16] See more detail in the previous chapter, in connection with the comparison of Bohemian and Moravian dialects.

happens through several processes like modifications in sound, grammatical alterations, and semantic shifts. Over a period of time, these changes can build up and result in notable distinctions between related languages.

Language evolution is also affected by external factors such as historical events, cultural influences; and technological advancements (e.g., invention of printing press, expansion of mass media) that have resulted in simplifying language standardisation and reducing of dialectal variations. Understanding these evolutionary processes is important in order to then distinguish between common ancestral features and new changes.

Keeping these factors in mind, assessing linguistic similarity through language family categorisation offers a useful manner of measuring certain relationships between languages. Even though the categorisation does not provide the entire assessment of language and its connections to other languages, it can still be considered as a valuable stepping stone.

There are different methods that used for expressing categorisation of languages into language families via diagrams. The most common method is designing a family tree that does a good job in capturing ancestral language and all the daughter languages that were formed from it while briefly portraying historical development. One of the downsides to this method is, however, that it cannot depict mutual relationships between languages[17] because the trees tend to become too large to accommodate these additional details.

Another method for depicting a wider image of language families is creating a wave model. This model is supposed to simulate a ripple effect of water where languages families are depicted in a non-linear pattern in groups that share similar characteristics. The wave model is mainly used in dialectology because it portrays well smaller variations in a language but can become rather chaotic when used for classifying languages.

2.1. Comparing Czech, English and Korean within the language families

All three compared languages are classified into different language families: Czech (West Slavic group, Indo-European language family), English (West Germanic group, Indo-European language family) and Korean (Koreanic language family). Since Czech and English are both part of the bigger Indo-European language family and have been in direct

[17] Especially non-genetic relationships such as cross-language influencing or borrowing.

and indirect close contact for centuries, an intuitive assessment can be made that these two languages will share more similarities than it will be in the case of Czech–Korean or English–Korean.

Nonetheless, assessing family relationships between language pairs and tracking down the origin of these languages back to a collective ancestral language is often performed when distinguishing the languages. This traditional approach helps in better defining the extent to which languages differ from each other. In comparative linguistics, there are even speculations about interconnectivity of languages and possibility of all languages sharing the same proto-language but overall, they are not very well supported.

Even though related languages within the same family share the same ancestral language, they all differ separately on multiple dimensions. According to one paper observing relationships among linguistic distances of three different language families,[18] the results showed that the linguistic dimensions (in this case lexical distance, phonetic distance, and syntactic distance) were generally correlated, and these correlations were stronger for pairs within families than when the pairs were examined together with all the other languages.

This thesis also observes the relationships between languages belonging to three different language families. To better observe the nature of these languages and their relationships between each other, the following subchapters will be focused on discerning these attributes within the families of the respective languages. After observing the similarities and differences of Czech, English and Korean in relation to their families, the thesis will proceed to examine their possible mutual relationships with each.

[18] Heeringa, et al. 2023.

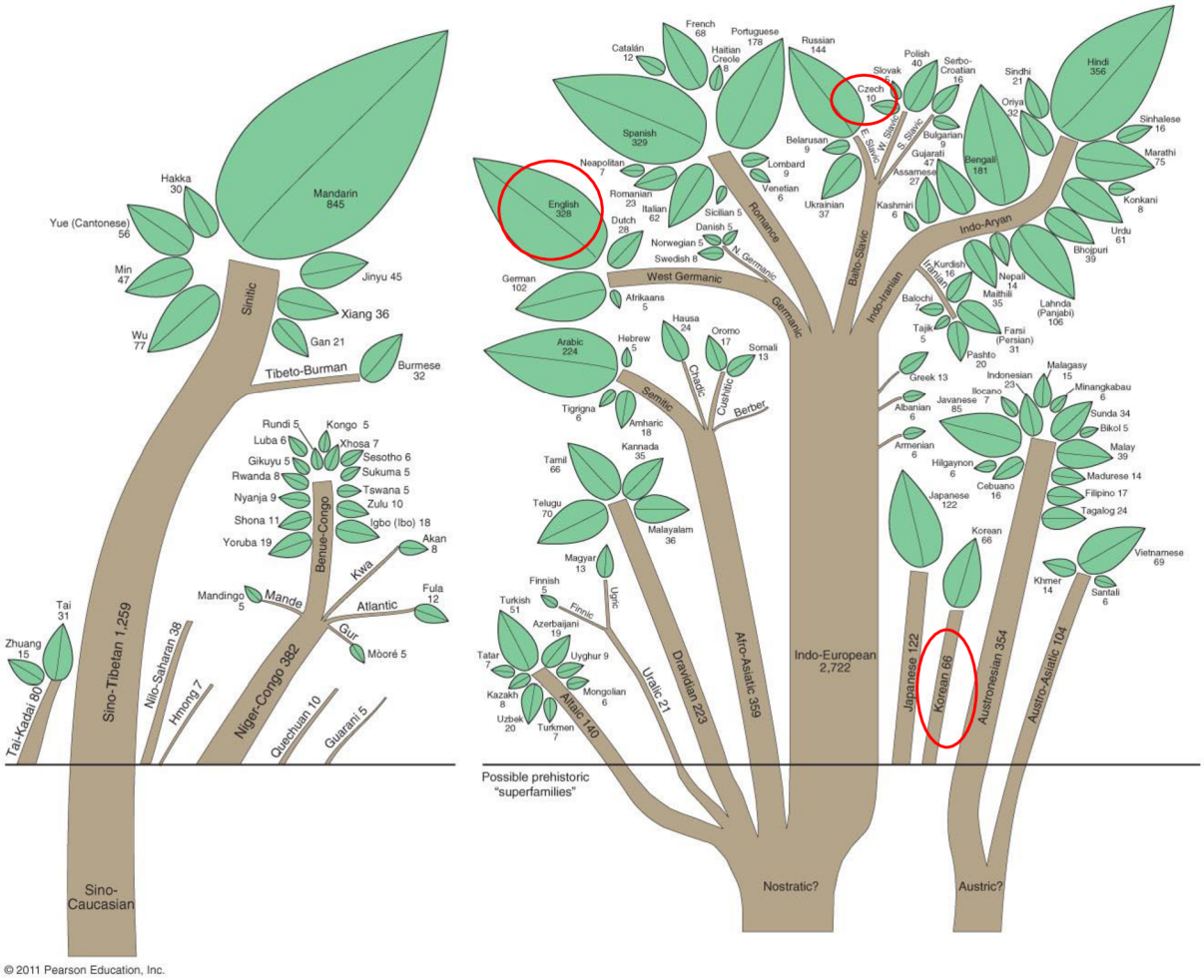

Fig. 1: Graphic representation of Czech, English and Korean classified in their language families

2.1.1. Czech as a part of Slavic languages

First language that will be compared within its language family is Czech. Comparing Czech with a language that it has a small linguistic distance with can help in recognising shared features within the entire family (that in turn will later help with better assessment of Czech in relation to another one of the observed languages of this thesis).

For example, Czech and Slovak are remarkably similar; they use Latin script, share certain grammatical principals and vocabulary, and have a common history since there was a time when the countries were united as one country.[19] The two languages share many similar aspects, therefore the linguistic distance between them can be considered as fairly small. This estimation can also be done on an intuitive basis because of the number

[19] Former Czechoslovakia.

of similar aspects between the two. However, if one wants to make a more precise assessment of the linguistic distance between Czech and Slovak, the categorisation of languages into various language families will make the process easier.

Language families are groups of languages that not only share many similar linguistic aspects (e.g., lexis, phonology, morphology, syntax), but also experienced a similar historical development. Czech is a part of the West Slavic language group. Slovak is also part of this language group; hence, it can be intuitively presumed that these two languages will be close in regard to linguistic distance (especially in comparison to other languages belonging to another language family). Both Czech and Slovak share similar linguistic rules and language structures, which might be also attributed to their geographical proximity and common history in the same empire they had shared for centuries. This principle applies to Slavic languages in general; their mutual linguistic proximity is tied to their historical development and geographical locations in relation to each other.

The language predating Slavic languages probably originated in today's Central or Eastern Europe (between the Dnieper River and the Oder River) and from there, it spread across further lands – territory of Balkans (Bulgarian, Macedonian, Slovene, Serbian, Bosnian, Croatian), Central Europe (Czech, Slovak), Eastern Europe (Belarusian, Ukrainian, Russian), and the northern parts of Asia (Russian).[20]

Despite the geographical movement that caused gradual increase of various Slavic dialects, a core similarity remained among those languages. The historical development can be seen "as a process in which tendencies to differentiate and to reintegrate the dialects have been continuously at work, bringing about a remarkable degree of uniformity throughout the Slavic area."[21] In a model communication situation where there are participants whose native tongues belong to Slavic family, and they both communicate in their respective native languages, it should be relatively easy for them to understand one another.[22]

Phonetics of Slavic languages is one of the linguistic features responsible for similarity in relation to each other. Many languages from this family share similar vowel and consonant sounds that make communication between the users of Slavic language understand the other party based on familiar pronunciation.

---

[20] Encyclopedia Britannica online, s.v. "Slavic languages." [Retrieved 1 March 2023.]
[21] Ibid.
[22] Meaning the closer the distance within language family, better the chance at understanding each other.

Another key linguistic feature facilitating communication between users of Slavic language is morphological similarity. Considering inflectional morphology, these languages share complex systems of noun declensions, verb conjugations, and adjective agreement but similar patterns and rules of the inflections apply on these languages within the group. The similar morphological structure makes it easier for users of one Slavic language to navigate the grammatical patterns of another.

2.1.2. English as a part of Germanic languages

Second, English language will be compared within its language family. English (just like Czech) belongs to the Indo-European language family but is a part of West Germanic branch. Germanic languages have encountered distinct changes over the course of time, but they still possess enough similarities that classify them as part of one language family. For example, these languages tend to follow a subject–verb–object word order, so users of Germanic language tend to have less difficulties understanding syntactic structure of another language from this group. Also, Germanic languages share a large number of cognates. These words facilitate comprehension of foreign vocabulary from another Germanic language, so users of these languages should be able to assume the correct meaning of words in mutual communication.

Germanic languages originated from hypothetical Proto-Germanic that was deduced by the comparative method of reconstruction because Germanic languages did not have any written records of their parent language. There were various historical developments of this earlier single parent language occurring. Some of historical evidence for Germanic languages includes for instance inscriptions of runic alphabet (whose later versions were sporadically used in England, Germany, but mostly in Scandinavia),[23] the incomplete Gothic translation of Bible or Old English (Anglo-Saxon).[24]

According to archaeological evidence, Germanic languages originated in the location of southern Scandinavia, along the North Sea and Baltic coasts from today's Netherlands to the Vistula River. Gradually, they spread out further and formed three branches: West Germanic (English, German, Dutch), North Germanic (Danish, Swedish,

[23] Although all later Germanic texts use instead adaptions of the Latin alphabet.
[24] Encyclopedia Britannica online, s.v. "Germanic languages." [Retrieved 2 March 2023.]

Icelandic, Norwegian, Faroese) and the now extinct East Germanic languages (Gothic, languages of few tribes, such as the Vandals or Burgundians).[25] [26]

English language was brought to British Isles by Anglo-Saxon migrants from today's territory of northwest Germany, southern Denmark, and Netherlands. The language originated as a group of languages used by settlers replacing Celtic languages (and possibly British Latin) that had dominated the Isles before.

When comparing modern-day English with another language within the West Germanic language group, shared core features of the language family can be observed. For example, German is another language from the West Germanic branch and the development of English had been influenced by this language in the past. Some of the similarities found between the two can be the already mentioned syntactic structure or shared cognates (where German language shares around 60% lexical similarity with English).[27]

Germanic languages share in general multiple distinctive characteristics that set them apart from other Indo-European languages, such as the development of a strong stress on the first syllable of the word (which led to significant phonological reduction of all other syllables), large numbers of vowel qualities (where English has approximately 11–12 vowels in most dialects, and standard German having around 14) or verb-second (V2) word order.

However, English can have some unique characteristics that become more prevalent when being assessed in comparison to German. For example, English is considered a relatively uninflected language whereas German is inflected (same as the common ancestral language). English used to share similar inflected variable forms with German but gradually shifted towards invariable forms – only nouns, pronouns, adjectives, and verbs are inflected. As the only European linguistic system, English possesses uninflected adjectives. While German words can additionally be inflected with four cases (for nouns, pronouns, and adjectives) and three genders, English uses cases only on nouns and pronouns, and has no variable forms to express gender of the word.[28] [29]

---

[25] Ibid.
[26] Encyclopedia Britannica online, s.v. "English language." [Retrieved 5 March 2023.]
[27] Day Translations Blog online, s.v. "German Interpreting: Key Differences Between German and English." [Retrieved 24 April 2024.]
[28] Conversely, similar linguistic features can be found in Czech.
[29] Encyclopedia Britannica online, s.v. "English language." [Retrieved 5 March 2023.]

Other important distinctions between English and German worth mentioning would be differences in the alphabet – even though both alphabets contain 26 letters, German also possesses a set of umlauted characters (ö, ü, ä, as well as the double S or scharfes S, represented by ß); phonology – German and English share many similarities in regards to intonation, stress patterns or sounds, however, some English phonemes are unique to itself and cannot be found in German (e.g., the sound of /θ/ and /ð/, or /w/ that is in this case usually replaced by /v/ sound); verb tenses – English has various tenses that do not have the same alternatives in German (e.g., present progressive or present perfect tense).

2.1.3. Korean as a part of Koreanic languages

Third, Korean language will be assessed within the scope of language families. Korean language has a special place in the global categorisation of language families because it belongs to a separate branch of Koreanic languages that is considered one of the primary language families. This language family consists of only two languages: Korean and Jeju,[30] although Yukjin dialect has been suggested for admission, too. Since Korean language was made in the 15th century, there are numerous records documenting the language's development in time, contrary to other languages. Even older records of Korean written by Chinese characters currently provide no information because they are difficult to decipher.

Apart from Koreanic language family, portmanteau families and language isolates are very common in Northeast Asia. There have been attempts to connect Korean to other of these languages and find some potential external relationships there, but none have been successful. The language families that are often considered for this pairing are Altaic (Tungusic, Mongolic, Turkic) and Japonic.[31]

Altaic languages share with Koreanic family multiple similarities, such as agglutinative morphology, subject–object–verb order or postpositions. In the process of defining Altaic family, various cognates were proposed and multiple attempts of remodelling a proto-language have been made. Nonetheless, Altaic has caused a mild controversy between linguists and is now considered a hypothetical language family.[32] It

[30] Jeju is regularly described as a dialect of Korean; however, its distinct traits are enough to consider it a separate language. There are other expressions used for this language ("Jejueo" or "Jejuan"), but this study will be using the term "Jeju."
[31] Wikipedia online, s.v. "Koreanic languages." [Retrieved 15 April 2024.]
[32] Starostin 2016.

was rejected because the relationship between Altaic languages has been generally accepted as a result of linguistic convergence (due to their geographical proximity) rather than shared ancestry.

While Korean scholars predominantly focused on potential linguistic bonds with Altaic group, other scholars were more inclined to study possible links to Japonic family. Both Koreanic and Japonic ancestral languages share similar phoneme inventories, comprising of a single series of obstruents, a single liquid consonant and six or seven vowels. When assessing possible relation to Japonic group, hundreds possible cognates with sound correspondences were proposed.

Despite all these factors, Koreanic and Japonic had had many interactions throughout their histories, so it is probable they could have influenced each other's language to certain extent. This could explain their mutual grammatical similarities, yet it equally makes the distinguishment of inherited cognates from ancient loanwords more arduous. Nonetheless, the Koreo-Japonic hypothesis has some valuable points, even if it still remains unproven.[33]

Considering the Koreanic family's portmanteau nature and Korean being a language isolate, it is quite hard to assess the ancestral language of this group. There are only two languages to provide the needed information, both of them historically developing within the borderlines of one relatively small territory (especially when compared to the size of territories other language families were evolving in).

Moreover, numerous sources are being neglected in the process of Proto-Korean's reconstruction. Since these sources had been written before creation of Hangul, they were recorded by using Chinese characters that are presently hard to identify (including sources on Early Middle and Old Korean). Even records that are written phonetically often get misinterpreted resulting in acceptance of some occurrences found solely in Middle Korean as "Proto-Korean."[34] Therefore, the Koreanic group offers a rather incomplete record of its linguistic development to properly reconstruct Proto-Korean language because of scarce implementation of sources predating creation of Hangul.

Old Korean, that originated during the Unified Silla period (7th century), is taken as a predecessor of all varieties of Modern Korean. There is not much known about other languages that were being used on Korean Peninsula before the unification, but it is

[33] Vovin 2017.
[34] This phenomenon has currently been changing and pre-Hangul materials are slowly being more recognised in Korean historical linguistics, too.

assumed that apart from Koreanic languages, there were also Japonic languages present in central and southern parts. Modern Korean variations display limited changes and can be treated as derived from Late Middle Korean (15th century), but earlier exceptions can be found, too. By using internal reconstruction from Middle Korean altogether with philological analysis of the fragmentary records of Old Korean, it is possible to roughly identify these exceptional variations that appeared after the unification of Silla.

All aspects of Korean language were significantly influenced by Classical Chinese, and multiple borrowed words and loanwords are still contained within its vocabulary. Sino-Korean vocabulary is approximated to make about 10% of basic vocabulary,[35] however, it supposedly borrowed more than half of the Korean lexicon in total (including especially technical terms). This is one of the factors that makes assessment of Korean and its language family harder, because of the multiple ancient loanwords and borrowed words that influence the language and possibly its relationship to other language families.

When comparing Korean language with another language from the same family, Jeju, certain similar features can be perceived.[36] Both languages are classified as head-final, agglutinative, and suffixing language. Jeju consonants share similar characteristics with Seoul Korean but in the case of vowels, Jeju possesses a larger and more conservative vowel inventory. The languages also used to share a back central unrounded vowel /ʌ/ that used to be expressed in Hangul as (ᆞ). Even though this vowel is still present in Jeju, it cannot be found in Modern Korean because it merged with other vowels in Korean dialects, thus ceasing to exist.

Nonetheless, Jeju is not mutually intelligible with Korean which is also reason for its classification as a separate language. The two languages notably vary in their verbal paradigms. For example, Korean does not have the continuative aspect marker of Jeju or the mood or aspect distinction of many Jeju connective suffixes. Yet, Jeju shares most of its lexicon with Korean while additionally preserving several Middle Korean words that got lost in Modern Korean.

[35] Sohn 1999.
[36] This assessment can be done intuitively based on the fact that Jeju has been previously considered a dialect of Korean.

3. Assessing languages and their mutual relationships: English, Czech, Korean

The previous chapter was focused on observing Czech, English and Korean throughout their historical development. It divided these languages into their respective language families and perceived the differences and similarities within these relationships. This chapter, however, is going to put emphasis on linguistic background of the languages. It will also work with the previously acquired information about historical development of each language and use their classifications into language families to better assess their linguistic attributes.

This part aims to introduce linguistic dimensions (e.g., grammar, vocabulary, phonology) of English, Czech and Korean respectively and then observe their mutual similarities and differences in relation to each other. These observations should eventually help in more accurate assessment of linguistic distance. It has already been established in the previous chapter that to assess linguistic distance, it is helpful to focus also on linguistic similarity between languages. This chapter will strive to determine shared similarities within linguistic dimensions with regards to some anticipated challenges Some of them being high numbers of borrowed words in vocabularies that could potentially mislead to false conclusions that these languages are closer to each other than is true or cognates with different meaning creating false sense of recognition in the users of language.

To describe and assess linguistic distance, this study will examine several grammatical units of Korean, English and Czech and the various mutual levels and relations. When it comes to the case of linguistic distance between Korean and English, there are comparatively much more academic articles and resources than in the case of Korean and Czech. That is also one of the reasons, why this study included English into the comparing language pairs.

3.1. Linguistic dimensions of English

The dimensions of English that will be observed in this chapter are phonology, vocabulary, and syntax. Starting with phonology, in the case of UK English, there exist various forms of standard speech within the English-speaking countries. However, British Received Pronunciation is recognised as the standard speech (used in London and

southeastern part of England). Conversely, US English with its own various forms (such as Inland Northern also referred to as General American) differ in the pronunciation of some individual vowels and diphthongs. English belongs to the group of strongly stressed languages that can be expressed by four degrees of accentuation (primary, secondary, tertiary, and weak which can be further indicated by three accent marks or by the breve).[37]

When it comes to pitch or musical tone, this can be either level, falling, rising, or falling–rising, depending also on the rate of vibration of the vocal cords. Another important linguistic aspect that is expressed in English is intonation that defines the tone of sentence. In general, intonation in British English is described as more singsong, compared to American, and it contains a wider range of pitch. Intonation is also influenced by geographical location because different patterns can be observed in regional accents.

The following dimension of English language is vocabulary. Modern English vocabulary is composed of wide range of languages. Approximately one quarter consists of Germanic (Old English, Scandinavian, Dutch, German), two-thirds are Italic or Romance (Latin, French, Spanish, Italian) and there is a large number of Greek expressions related to science and technology that keeps growing. Apart from these, a considerable number of words is being borrowed from more than 300 other languages.[38]

The final dimension mentioned in this part will be syntax. Sentences can be divided into three categories: simple, having one clause and predication; multiple or compound, including two or more coordinate clauses; and complex sentences, having one or more main clauses and one or more subordinate clauses. In the case of simple, declarative, affirmative sentences, two main patterns with five subsidiary patterns within each apply.[39]

3.2. Linguistic dimensions of Czech

Within the different dimensions of Czech, this part will be focused specifically on phonology and grammar. In Czech phonology, some quite unique phonemes can be distinguished. Starting with vowels, Standard Czech has ten basic vowel phonemes (⟨a⟩ /a/, ⟨e/ě⟩ /ɛ/, ⟨i/y⟩ /ɪ/, ⟨o⟩ /o/, ⟨u⟩ /u/; plus, their long equivalents ⟨á⟩ /a:/, ⟨é⟩ /ɛ:/, ⟨í/ý⟩ /i:/,

[37] Encyclopedia Britannica online, s.v. “English language.” [Retrieved 1 March 2024.]
[38] Ibid.
[39] Ibid.

⟨ó⟩ /oː/, ⟨ú/ů⟩ /uː/), and three diphthongs (⟨ou⟩ /ou̯/, ⟨au⟩ /au̯/, ⟨eu⟩ /ɛu̯/).[40] The letter ⟨ě⟩ has a special function; it signalises that every previous consonant is palatalised, but after a labial it marks the sound /jɛ/ with few exceptions (e.g., ⟨mě⟩ /mɲɛ/).[41]

When it comes to Czech consonants, they are divided between "hard," "ambivalent," and "soft." General rule usually applies on these consonants with a couple of exceptions. For example, after "hard" consonants do not follow ⟨i/í⟩ and vice versa, after "soft" consonants should not follow ⟨y/ý⟩.[42] As for the "ambivalent" consonants, they tend to be used with both characters indiscriminately.

Czech has many distinctive phonemes but ⟨ř⟩ is probably the most unique of them all. It is a rare sound and often presented as belonging solely to Czech language, although presumably it was also found in some dialects of Kashubian and in older variation of Polish.[43] When comparing the sound within Czech language, a variation of ⟨r⟩ and ⟨ž⟩ seems to imitate the sound the closest to reality. Another peculiar aspect of the language (that is shared within the group of Slavic languages) is that the consonants /r/, /l/, and /m/ can be vocalic when they are placed in words formed only with consonants. In that case, the consonants will be replacing a vowel – acting as syllable nuclei.[44]

The primary stress is in Czech put on the first syllable of a word. The only exceptions are minor, monosyllabic syllables that are in their nature unstressed. In the case of longer words, when there are two or more syllables, secondary stress is put on every odd-numbered syllable. Stress serves to mark boundaries between individual words, but it has no additional lexical or phonological function, nor does it discern the word meaning. Stress can be put on any vowel no matter its length; both long and short vowels can be stressed or unstressed and its absence cannot make vowels reduced in tone. Also, stress can be moved from the first syllable. If a monosyllabic preposition is put in front of the noun, it merges with the following word and the preposition ends up being stressed instead.[45]

Next linguistic dimension being assessed is Czech grammar that is inflected (as is a common quality across other Slavic languages as well). Phonological processes inflect

[40] Dankovičová 1999.
[41] Wikipedia online, s.v. "Czech language." [Retrieved 25 April 2024.]
[42] Except for loanwords.
[43] See Klein 2018.
[44] Naughton 2005.
[45] Wikipedia online, s.v. "Czech language." [Retrieved 25 April 2024.]

nouns, verbs, and adjectives[46] to adjust their meanings and grammatical functions. Czech is a rich language that uses various parts of speech including adjectives, adverbs, numbers, interrogative words, prepositions, conjunctions, and interjections. The language has many similarly formed words. Many Czech words are derived from other words e.g., adverbs are mainly created from the root of adjectives and new words are created by using affixes e.g., negative statements are made by adding the prefix ⟨ne-⟩ to the main verb of a clause.[47] [48]

Czech has a flexible word order; it does not have to rely on it to communicate clearly because all the essential information is already contained within grammatical case which conveys word function in a sentence. The opposite is the case for Englich which depends on word order to correctly convey information. Czech can, therefore, omit subject pronouns and form sentences consisting of only a verb because it already carries the information about the subject, too. The first syntactic unit of a sentence tends to be a subject or an object, such as a main form of a verb, an adverb, or a conjunction,[49] while auxiliary verbs and pronouns are generally placed as the second unit. Syntax has a subject–verb–object sentence structure (same as English) but because of the flexible word order this structure can change in some cases (e.g., in colloquial style, passive voice can be replaced by switched words[50]). The flexible word order helps to distinguish topic and focus of a sentence since known referents typically precede new information, hence the sentence does not need to maintain a strict structure.[51]

3.3. Linguistic dimensions of Korean

There are various linguistic dimensions contributing to the entire characterisation of Korean language. In this brief part, grammar and phonology will be described. First, Korean grammar shares some similar features with Japanese, what is noticeable when observing different sentences and their structure. However, when it comes to the phonological attributes of words, the similarity between the two does no longer apply.

---

[46] More specifically, they are inflected for case, gender, number in nouns and tense, aspect, mood, person, subject number, and gender in verbs.
[47] The only exception being the verb "to be" ⟨nebýt⟩ in third person ⟨není⟩.
[48] Ibid.
[49] Excluding the conjunctions "and" ⟨a⟩, "and even" ⟨i⟩ or "but" ⟨ale⟩.
[50] Meaning the sentence structure can also occur in a format object–verb–subject.
[51] Ibid.

Korean language possesses modifiers that always precede what they modify in a sentence. The unmarked order of the sentence structure would consist of subject + indirect object + direct object + predicate. However, only predicate is crucial because other information can be omitted without disturbing the meaning of the sentence. Processive predicates (verbs) are used to express actions and descriptive predicates (adjectives) are used for characteristics.[52] While descriptive predicates are only used to make statements or questions, processive predicates can also make commands and suggestions depending on intonation.[53]

Korean possesses specific particles that are attached to nouns to demonstrate their role in the sentence; if it concerns a subject, 이/가 [-i/ga] particle is attached after a noun, however, for the direct object, particles 을/를 [-ŭl/lŭl] are used. There are also endings that modify a following noun (predicate) that work in a similar way as English relative clause does. However, the level of formality or appreciation for a person who is a subject marked by predicate in the sentence can be in Korean expressed by inserting honorific marker (으)시 [-(ŭ)si-]. Moreover, there are different ways to mark the predicate also for tense and aspect.

It is important to keep in mind that when observing Korean grammar, there are multiple levels of formality in which the information can be conveyed depending on situation or level of respect the speaker has to the person that is being addressed.

Next linguistic dimension is the phonology of Korean, more specifically assimilation happening when two syllables are put together. The structure of Korean syllables is rule-governed and generally simple. Each syllable ends either in a vowel or in one of the voiced consonants. When two syllables are connected together, several changes in pronunciation can occur in a place where they meet one another. For example, if a syllable ending in a stop is followed by one that begins with a nasal, the stop assimilates.[54]

Some of these syllable assimilations can result in multiple combinations that sound almost identical. For this reason, Hangul spelling differentiates such combinations by writing the basic forms but in some cases, the assimilation is ignored in both spelling and

[52] There can also be a closely attached copula (linking verb) present. It is a special kind of descriptive that predicates nouns.
[53] Encyclopedia Britannica online, s.v. “Korean language.” [Retrieved 26 April 2024.]
[54] Ibid.

transcriptions. For example, " both ㄴ [-n] + ㄹ [l-] and ㄹ [-l] + ㄴ[n-] are pronounced like [-l] + [l-], so for the sound [-l:l-] one must know what is in the word to decide which (…) Hangul spellings to use."[55]

Finally, this thesis will look at aspirated and reinforced consonants within Korean phonology. Compared to English, which has a two-way distinction of voiceless and voiced stops, voicing in Korean is automatic (e.g., sounds [p] and [b] create a single phoneme, and both are transcribed with the same Hangul letter). Korean recognises two other kinds of obstruents – heavily aspirated ([p'], [t'], [k'], [ch']) and reinforced ([pp], [tt], [kk], [tch]). The language also possesses a tense [ss] and relatively aspirated [s], but the different pronunciation is not commonly distinguished in communication, unless the phoneme is located at the beginning of a word.[56]

3.4. Linguistic distance between Korean and English

There are many linguistic dimensions that could be considered when comparing Korean and English. Since including all of them would be too extensive,[57] this chapter will highlight only briefly some of the basic differences between the languages. For example, Korean standard syntax follows the basic order subject–object–verb, while English syntax follows in principle the subject–verb–object order. Even though the difference is easily perceivable at the first observation, this difference alone can only hint towards the linguistic distance between the languages, it cannot be considered as a defining difference proving the depth of the linguistic distance between the languages. Another such difference can be perceived in spelling. Korean has a non-linear writing system where spelling mostly follows consonant-vowel-consonant position (e.g., 김 [kim]). By contrast, English has a linear writing system where some words can have multiple consonants positioned next to each other (e.g., three).

A major difference relevant for translation and interpretation concerns indefinite and definite articles. Korean has no equivalent for the English indefinite and definite

[55] Ibid.
[56] Ibid.
[57] Some of the academic literature approaching this topic includes for example Wang, M., et al. (2006), *Korean–English Biliteracy Acquisition: Cross-Language Phonological and Orthographic Transfer* or Lee Amuzie, Grace, and Patti Spinner (2013), *Korean EFL Learners' Indefinite Article Use with Four Types of Abstract Nouns.*

articles “a/an” and “the” and on the contrary, Korean uses particles such as 은/는 [-ŭn/nŭn] and 이/가 [-i/ga] that do not exist in English language. While English indefinite and definite articles come before the nouns, the Korean particles are attached to the end of the noun. The Korean particle pairs are attached to the topic or the subject of a sentence where 은/는 [-ŭn/nŭn] represents the topic marker, while 이/가 [-i/ga] is the subject marker. This contrast between the languages suggests that there is some sort of linguistic distance, and it just needs to be assessed to which extent that is.

When comparing linguistic differences between the two languages, it is important to consider the full scope of Korean language with regards to its historical evolution. A prominent change can be seen when comparing the language used in North Korea with the current form of South Korean. As already said above, North Korea is not that open to global influence as South Korea; the language consists mainly of pure Korean words and words taken from Russian, while South Korea is implementing many words from US English. Some of the original English words Korean has accepted into its vocabulary are for example: 에어컨 [eŏk’ŏn] air-conditioner, 아이스크림 [aisŭk’ŭrim] ice cream, 와이프 [waip’ŭ] wife.

For a more detailed comparison, the next part will be observing a phonological analysis of Korean–English structures and processes.[58] The analysis observes comparison in phonemic inventories of both English and Korean while taking into consideration uniqueness of every phonemic inventory. These observations were made based on the results of measuring linguistic distance in the field of interpreting. First, vowels will be compared in the following part.

The Korean vowel inventory can be composed of either seven or ten vowels, depending on the individual stance of every researcher. The nonuniform vowel inventory is a result of a debate about the phonemic status of the front rounded vowels /ü/ and /ö/ and the fusion of /e/ and /ɛ/ vowels.[59] In this study, the seven-vowel system will be used.

In this system, there are rounded and unrounded vowels further categorised based on their position. The unrounded vowels include two front vowels: high /ɪ/ and non-high /ɛ/; and then there is also the non-front (mid or back) low vowel /ɑ/. Other two unrounded

[58] Cho and Park 2006.
[59] Yu Cho 2016.

vowels are the high back vowel /ɯ/[60] and the mid back vowel /ɣ/ that are not often present in other languages (and are equally absent in the English phonemic vowels, as seen below). Next, the rounded vowels include high back position /u/ and mid back position /o/ vowels but because of frequent diphthongization, these front rounded vowels cannot be commonly found as pure monophthongs in most contexts. There are ten diphthongs: /ja/, /jɣ /, /jo/, /ju/, /jɛ/, /wɑ/, /wɣ/, /wɛ/, /wi/, /ɯi/, and they are all combinations of a glide and a vowel.[61]

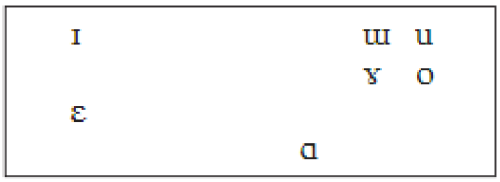

Fig. 2: Korean phonemic vowels

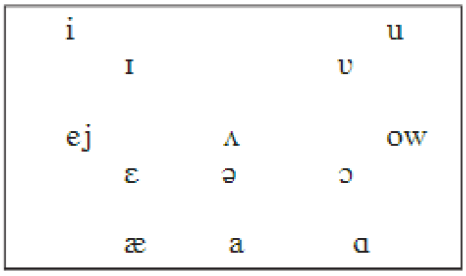

Fig. 3: English phonemic vowels

Conversely, the English phonemic system has a lot more vowels with more subtle distinctions (e.g., tense/lax distinction). Cho and Park identified these unassuming distinctions and compared them within minimal pairs to better present these differences. For example, the unrounded vowels /i/ and /ɪ/ are both in high and front position, but /i/ is a tense version of a lax counterpart /ɪ/. These distinctions become more visible within those minimal pairs, such as “beat” (for /i/) vs. “bit” (for /ɪ/). Same applies to /u/ which is the tense version of /ʊ/ (e.g., “pool” vs. “pull”). The other vowels that have the tense/lax distinction are in mid front and mid back positions; two mid front unrounded vowels, tense /ej/ (e.g., “bait”) and lax /ɛ/ (e.g., “bet”), and two mid back rounded vowels, tense /ow/ (e.g., “boat”) and lax /ɔ/ (e.g., “bore”). Nonetheless, the tense/lax distinction is not the only defining distinction for the mid front and mid back English vowels.[62]

All the tense versions of the mid front and mid back vowels are also diphthongized. Thus, instead of a simple /e/ there is diphthongized /ej/ and instead of /o/ there is /ow/. English diphthongs are the exact opposite sequence of Korean because of their vowel and glide combination. Apart from the already introduced diphthongized vowels /ej/ and /ow/,

[60] Which often ranges from central to back.
[61] Cho and Park 2006.
[62] Ibid.

there are also other diphthongs, such as /aj/ (e.g., “sigh”), /aw/ (e.g., “how”) and /ɔj/ (e.g., “joy”) that the analysis mentioned. However, the high front and back unrounded tense vowels can be sometimes considered as slightly diphthongized, too, and thus transcribed as /ij/ and /uw/.[63]

Next, the mid central vowels include /ʌ/ (e.g., “cut”) and schwa /ə/ (e.g., “banana”) whose occurrence depends on the presence or absence of stress. In the case of schwa /ə/, it is important to note that this vowel can occur in various spelling combinations and its sound will remain generally the same. Schwa /ə/ is commonly found in weak syllables so its sound does not change because of the absence of stress on the given syllables. However, the sound could change in accordance with the spelling, if the syllables were pronounced as strong syllables.[64] For example, schwa /ə/ spelt with “a” would change to /æ/ with strong pronunciation (e.g., “attend”). These alterations occur within content words (e.g., nouns, adjectives, verbs, adverbs that communicate the main information of the sentence) or function words (e.g., prepositions, auxiliary verbs, articles that connect the sentence together). Apart from the already mentioned spellings with “a,” schwa /ə/ can be found in the spellings with “e” (e.g., “water”), “i” (e.g., “council”), “o” (e.g., “complete”), “u” (e.g., “support”) vowels, too.

Lastly, Cho and Park introduced the front low vowel /æ/ (e.g., “bat”) and the low back vowel /ɑ/ (e.g., “pot”) in the English phonemic system. There is also the low central vowel /a/, but this one occurs sporadically and even if that is the case, it mostly occurs as the first part of diphthongs (/aj/, /aw/). In the following part of the analysis, a detailed comparison was created in order to identify dividing differences between English and Korean vowel systems.[65] Some of the observations included identification of /ɣ/, /o/, /ɯ/ Korean vowels as possibly difficult to pronounce for the English native speakers learning Korean since they are not present in their native language.

Even so, comparably more English vowels are not present in the Korean phonemic system suggesting that Korean native speakers might encounter more difficulties when pronouncing English vowels than vice versa. In particular, these English vowels include /i/, /ʊ/, /æ/, /ej/, /ow/, /ɔ/, /ə/, /ʌ/, /a/. Identifying these potentially challenging vowels that do not occur in the other assessed language can serve as a stepping stone in recognising

[63] Ibid.
[64] Roach 2009.
[65] The comparison was made with the objective of identifying common mistakes that native Korean and native English interpreters make when using the other language.

aspects of Korean that can be contributing to widening of the linguistic distance in relation to this other language (in this case English).[66]

The struggles with pronunciation that were analysed in the study became apparent, for example, in the case of /æ/ vowel that was put into the comparison with the /ɛ/ sound. The /æ/ vowel is in the front non-high tongue position that Korean does not recognise. The most similar counterpart to this vowel that can be found in Korean is /ɛ/ vowel, but it ranges over the mid front and low front positions. Because of this, Korean native speakers may struggle to differentiate between the two English vowels that can lead to some errors in pronunciation (or comprehension) especially evident in the case of minimal pairs that are distinguished by these vowels (e.g., "pen" vs. "pan", "set" vs. "sat").

Another defining difference that can potentially cause problems with pronunciation of some English vowels for Korean speakers is the tense/lax distinction. The tense/lax distinction can be found in the vowels in high tongue position, such as in pairs /i/ vs. /ɪ/ (e.g., "eat" vs. "it", "scene" vs. "sin) and /u/ vs. /ʊ/ (e.g., "pool" vs. "pull", "fool" vs. "full"). Korean vowel system does not include these distinctions; it only has one high front unrounded vowel /ɪ/ and one high back rounded vowel /u/. Despite that the high vowels in English are also distinguished by their length difference[67] so this distinction can at least initially help in better imitation of these vowels.[68]

Second, the analysis focused on observing the consonants, starting with the ones belonging to the Korean phonemic system. There are no voiced consonants on the phonemic level because Korean consonant inventory does not include voiced obstruents – all stops, fricatives and affricates are voiceless. Even though many obstruents do not have the voicing contrast, it is generally replaced by aspiration contrast and glottis constriction contrast. For example, aspiration contrast can differ based on the word position of the obstruents. Plain obstruents will be slightly aspirated if they are placed in the beginning of a word but in medial position, they will become voiced[69] and considered as tense. In the final-word position, the obstruents have no laryngeal features because they are all unreleased. In general, despite Korean consonants not being voiced, they have a three-way laryngeal distinction (fortis, lenis and aspirated obstruents) that involves

[66] Cho and Park 2006.
[67] Where the tense vowels are longer.
[68] Ibid.
[69] Except for /s/.

aspiration and tenseness. Finally, an important distinction of the Korean consonant inventory is that it includes flap [ɾ], but the consonants [l] or [ɹ] are not present which can lead to potential difficulties in communication.[70]

	Bilabial	Labio-dental	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Labio-velar	Glottal
Stop	p pʰ pʼ			t tʰ tʼ			k kʰ kʼ		
Fricative				s sʼ					h
Affricate					tʃ tʃʰ tʃʼ				
Nasal	m			n			ŋ		
Approximant						j	ɯ	w	
Flap				ɾ					

Fig. 4: Korean phonemic consonants

	Bilabial	Labio-dental	Dental	Alveolar	Palato-alveolar	Palatal	Velar	Labio-velar	Glottal
Stop	p b			t d			k g		ʔ
Fricative		f v	θ ð	s z	ʃ ʒ				h
Affricate					tʃ dʒ				
Nasal	m			n			ŋ		
Approximant				l ɹ		j		w	

Fig. 5: English phonemic consonants

Conversely, English phonemic system includes obstruents that all have voiced or voiceless counterparts – except for the consonants pronounced with glottis. These exceptions can apply to fortis consonants, usually made with open glottis, that often occur in fricatives. Compared to Korean inventory, English possesses multiple fricatives that can be found in different positions of articulation. The other exceptions are consonants made with closed glottis, such as in stops (marked as /ʔ/).

Both Korean and English consonant inventories include certain consonants that are unique[71] for them. These unique consonants might hint towards potential pronunciation struggles and comprehension challenges. For example, Korean consonants /pʰ/, /pʼ/, /th/, /tʰ/, /kʰ/, /kʼ/, /sʼ/, /tʃʰ/, /tʃʼ/, /ɾ/ cannot be found in the English inventory thus representing possible difficulty for English speakers when they would try imitating these sounds. The

[70] Ibid.
[71] From this point forward, when the thesis will be referring to "unique consonants," it will be meant in the context of comparison between Korean and English inventories.

obstruents with constricted glottis in particular seem to represent the biggest challenge since English glottal obstruents are typically made with open or closed glottis.

As for the English unique consonants that could lead to errors in pronunciation/comprehension for Korean speakers, these include /b/, /d/, /g/, /f/, /v/, /θ/, /ð/, /z/, /ʃ/, /ʒ/, /dʒ/, /l/ and /ɹ/. Apart from the unique pronunciation, the English voiced consonants can be placed in any position in a word (and still be pronounced as voiced) but the same does not apply for the Korean consonants. For them to become voiced, they need to be placed between voiced segments. Therefore, if Korean speakers encounter the English voiced consonants that are not within these voiced segments, they can struggle with their reproduction (even if they are normally able to pronounce these consonants when they are put between voiced segments). Hence, the first voiced consonants placed in the beginning of the word tend to be pronounced as their unaspirated voiceless counterparts. For example, the first consonant in the name "Bob" is usually pronounced with /p/[72] instead of /b/ and, similarly, the initial consonant in the name "John" tends to be pronounced with /tʃ/ instead of /dʒ/. Korean has only a few fricatives compared to English, therefore, many English fricatives, such as /f/, /v/, /θ/, /ð/, /z/, /ʃ/, /ʒ/ tend to be replaced by stops or affricates that are closest to them. For example, /f/ tends to be pronounced as /p/ (e.g., "fine"), /v/ is pronounced as /b/ (e.g., "video") by /b/, /θ/ as /s/ (e.g., "three"), or /z/ as /tʃ/ (e.g., "zebra").[73]

Korean speakers do not distinguish between liquids /l/ and /ɹ/ such as English speakers do but instead, these sounds are replaced by flap /ɾ/. However, this distinction is important in English communication because mispronunciation can result in the usage of words with a vastly different meaning (e.g., "lice" vs. "rice."). Moreover, Korean speakers are usually inclined to use English /l/ more often than /ɹ/ because the English retroflex /ɹ/ is not present in Korean. Korean speakers perceive the consonants /l/ and /ɹ/ as variations of only one sound which is represented by flap /ɾ/. In English phonemic system, the Korean flap /ɾ/ could be replaced by phoneme /l/ with three different allophones (an apical flap /ɾ/, a lateral /l/ and a geminate /ll/), therefore suggesting more troubles for Korean speakers when reproducing the English consonant /ɹ/ than the consonant /l/.

---

[72] This might happen because in Korean language consonant ㅂ [b/p], that would be used to rewrite name Bob, expresses both sounds based on its position to other phonemes. This principle often applies in Korean transcription.
[73] Ibid.

Possible problems in communication can also arise on the other side of the communication channel where English speakers are receiving the message from Korean speakers. When either of the English consonants, /l/ or /ɹ/, are pronounced as Korean flap /ɾ/, English speakers usually tend to assume that the speaker is trying to imitate English /ɹ/ instead of /l/. Therefore, the words containing the consonant /ɹ/ are generally better understood that the ones with the consonant /l/ (e.g., “row” vs. “low”).

Lastly, an incomplete transcription of English consonants can lead to some difficulties in reproduction, too. For example, the inherently rounded English consonants (e.g., /ɹ/, /ʃ/, /ʒ/, /tʃ/, /dʒ/) are not often marked with a superscript describing them as rounded (e.g., /ɹ^{w}/) because the inherit feature is considered obvious for native speakers even without such transcription. Nonetheless, this does not necessarily apply to non-native speakers. They can have some of these consonants in their native language, too, but as unrounded instead of rounded consonants. Especially for Korean speakers, they commonly replace these English consonants with their unrounded counterparts, and even if they learn to imitate the sound closely, they often leave out the rounding of lips.

Last aspect that will be assessed in this part is comparison in phonotactics. Considering the numbers of allowed constants in both onset and coda position of syllables, there are vast differences between Korean and English. For example, Korean has a maximum of only one consonant in the onset position; however, English allows up to three consonants. In the coda position, Korean again allows a maximum of only one consonant (exceptionally two but that applies only to a small number of combinations) while English allows three consonants.[74]

- Maximal syllable structure in Korean: CVC(C)
- Maximal syllable structure in English: CCCVCCC

The significant difference between the Korean and English syllable structure adds to the linguistic distance between the languages. Furthermore, it pushes Korean speakers to use multiple syllable structure preservation tactics when dealing with certain English syllable structures that are foreign to them. Another aspect that potentially contributes to the linguistic distance between the two languages is that Korean has restrictions regarding its syllables. There cannot occur any fricatives in coda position, and it does not allow

[74] Ibid.

certain consonant + vowel combinations that are present in English. These factors can be partially responsible for Korean speakers having difficulties reproducing some aspects of English pronunciation.

3.4.1. Konglish

English and Korean have a very specific relationship since English has been historically present in South Korea for a considerable amount of time. During the Korean War, South Korea had accepted military aid from the United States and strengthened the ties between the two countries even more. Ever since, the influence of the USA, and hence the influence of English language itself, can be perceived in the gradual development of Korean language. There is an evident increase of implementation of multiple English words in Korean vocabulary – as is also the case with other foreign languages due to the influence of globalisation and strong position of English language as a lingua franca.

However, in Korean this influence of English language is so strong, it resulted in creation of "Konglish." Konglish is a blend of Korean and English with Korean version of English vocabulary that has become widespread throughout South Korea. Generally, it has a rather negative connotation since its origin is perceived as a result of diminished knowledge of English. Conversely, there are certain articles that focus on this phenomenon and put focus on analysing its more positive attributes and potential for future studies.[75]

The expression Konglish is in general used to point out false friends between Korean and English, but it might also be seen as a special interlanguage of Korean learners of English. The function of language does not serve uniformly to fulfil only basic linguistic needs, but it can also serve as a medium to communicate culturally determined connotation and metaphor. Same applies to Konglish, not only linguistic level needs to be examined, but sociolinguistic and pragmatic functional deficits of Konglish should be considered as well.

When it comes to Konglish, there are multiple disputed opinions about its nature; whether there is any potential to use it as a useful tool in observing linguistic bonds between Korean and English. In the next part, the thesis will observe two applications of

[75] See Nam 2010.

Konglish within the L2 acquisition environment.[76] The dilemma is that it is hard to decide whether the use of Konglish words in English communication can be perceived as the use of Korean resources instead of evidence of English-based communication strategies. Konglish words are classified as loanwords, therefore it could be assumed that these resources are stored as L2 entries and in the process of accessing a target L2 word, an L2 competitor might get selected instead. If this was the case, the use of Konglish words in English might be then a part of L2-based communication strategies where it is a manifestation of learners' insufficient practice, and not evidence of L1 access.

Conversely, if Konglish words are stored as L1 items in Korean and retrieved through L1 entries in the L2 communication, Konglish may become a useful means in further studies exploring whether L1 is activated in L2 access.

Konglish vocabulary contains several types of words, such as: words whose meanings have been changed, borrowed words or phrases that have been made to mean something different from the original, words with altered pronunciation, and abbreviated words or phrases. Apart from common loanwords, this vocabulary also includes "pseudo loanwords" which are essentially English words used in Korean after making a direct translation from Korean to English. As one of the examples can be used even the name "Konglish" that was created as a combination of the first syllable of the word "Korean" and the word "English" (without the initial vowel e). These "pseudo loanwords" can often be adjusted so that the original form of the word becomes shorter.

Not only does Konglish contain "pseudo loanwords", but "false cognates" can be found within its vocabulary, too. For example, the word 서비스 [sŏbisŭ] "service" retained its phonological similarity to English, but the semantic feature has changed. In Korean language, the word 서비스 [sŏbisŭ] does not only mean "service" but can be used in to express something is "free of charge" or "on the house." Moreover, the word can appear in financial sphere, too, when requesting a cash advance with a credit card at a Korean bank.

Some traditional linguistic approaches considered cognate pairs only if it concerned languages that were etymologically related but nowadays, many studies focus on cross-language similarities within cognate pairs that do not have to share the same language

[76] Where L1 is Korean and L2 is English.

family nor mutual historical development. In that case, Konglish words could be to certain extent considered as cognates.[77]

The other type of Konglish words that has been mentioned in the beginning are abbreviations. Korean has a tendency to create new words by choosing certain syllables from the word/s (typically the initial one) and putting them together, so they can fabricate a short form of the original expression. This process happens both with Korean words as well as borrowed English words. In the case of English words, they are first transformed into Korean version of the word (based on its perceived pronunciation) and then they get shortened (e.g., happens regularly with slang words).

Since these borrowed English expressions undergo de facto two transformative processes and the abbreviated product generally does not resemble the original English word, it can be debated to what extent these Konglish words become uniquely Korean. Some of the examples include 에어컨 [eŏk'ŏn] air conditioner, 아파트 [ap'at'ŭ] apartment, 리모컨 [rimok'ŏn] remote control, 이벤트 [ibent'ŭ] sale/promotion.

Borrowing of words has been suggested to be connected with the presence of a clear base language, but code-switching is linked to the presence of two languages interacting in communication.[78] Therefore it is worth reflecting, whether certain Konglish words that have undergone transformative processes, while being borrowed from English vocabulary, could be a part of code-switching rather than borrowing. Following that it would be understandable to reconsider if some Konglish expressions might be taken as a part of unique Korean vocabulary.

However, this aspect also makes it more demanding to assess linguistic distance between Korean and English. It is important to distinguish between specific characteristics of Korean and English to differentiate their mutual distance. But when there is present merging of languages to the extent where an interlanguage is created as a byproduct, the distinctions become less apparent. Conversely, this interconnection may serve as a bridge between Korean and other languages that share certain similarities with English.

[77] However, considering that many Konglish words can have different meaning than the original English word it was derived from (despite their phonological similarity), it is better to mark them as "false cognates."
[78] See Nam 2010.

3.5. Linguistic distance between English and Czech

Even though Czech belongs to the Slavic languages and English to Germanic, they both form part of the large Indo-European family tree. This fact means that both languages might share more similar features than any of them would with Korean language. Czech and English share same syntax structure subject–verb–object, but the word order of Czech is flexible. When it comes to assessing the linguistic distance between these two languages, this thesis will observe the findings from the article focused on phonology of Czech words.[79]

The phonological corpus of Czech is constructed from phonologically transcribed databases including different word lists and texts from Modern Czech. The study that is used for observation in this chapter is based on quantitative analysis, focusing on phoneme frequency, phoneme combinations, and the syllabic structure of words. Even though phonology examines word from two perspectives – as a sequence of phonemes and a bearer of prosodic features (e.g., accent, stress), the article focused primarily on observing phonology of Czech from the former perspective.

Czech and English possess a similar number of phonemes and many of them can be found in both language inventories. Even though these languages share certain phonemes, it does not necessarily mean that their phonology is similar. One of the reasons is different combinability of the phonemes. For example, /m/ and /l/ phonemes can be found in both, Czech and English. However, their combinability differs in each language suggesting that the phonology of these two is not as similar as it may have appeared at the first observation.

In Czech, there are many various phoneme combinations that follow specific rules regarding their structure. Especially in the case of combining consonants, the number of combinations is high; there are over 440 syllable-initial combinations and over 100 syllable-final combinations in Standard Czech.[80] Another aspect of comparing Czech and English phonology depends on better understanding of the sound structure that can be fulfilled by considering the co-occurrence of multiple syllables and its conditions. When using a language for communication, the syllables tend to be naturally grouped in larger units (such as words) rather than being separated one by one. By observing the co-

[79] Bičan 2014.
[80] Ibid.

occurrence of syllables in words a more complex representation of consonantal combinations is provided.

There are some rules regarding the number of Czech consonantal combinations that is permitted across two syllables. In general, considering multiple syllables offers an even greater range of consonantal combinations than was possible within one-syllable. According to the results of lexical corpus analysis,[81] the most of consonantal combinations (approx. 70%) within the Czech lexemes occurred word-internally, followed by combinations in the beginning of the word (approx. 25%) and the least combinations (approx. 5%) were found in the final position of the lexemes. Based on the results, it can be presumed that even though the combinations of phonemes are conditioned by certain rules, these are less strict for word-internal combinations.[82]

Czech contains some phonologically distinctive consonantal combinations. These combinations of consonants cannot be found in English, moreover, they are rather "uncommon" in Czech vocabulary, too. For example, the names of Czech municipalities (e.g., Břvany, Polště, Kfely) illustrate this fact and attest to the peculiar nature and richness of the combinations of Czech consonants. The other aspect that needs to be considered is the allowed combinability of Czech phonemes in inflected forms of words (e.g., ⟨vojsk⟩, genitive case with plural form of ⟨vojsko⟩ "army," ⟨pomst⟩, genitive case with plural form of ⟨pomsta⟩ "revenge"). Some of the mentioned consonantal combinations can be found even in the nominative case but only within two syllables. Nonetheless, if the same combinations are to occur within one syllable, they can be permitted only in the inflected forms of words.

The results of the article about word phonology in Czech by Bičan show that in the observed lexical corpus the most common phonemes were as follows: /o/, /a/, and /i/ while the most common consonants were: /t/, /v/ and /r/. Conversely, as the least common phonemes were noted diphthongs: /au̯/ (e.g., ⟨auto⟩ "car") and /ɛu̯/ (e.g., ⟨euro⟩).[83] When concluding phoneme classes, front vowels were more prevalent (compared to mid and back vowels), together with stops (in comparison with fricatives and nasals) and alveolars (compared to labials, palatals and velars) occurring frequently. Meanwhile, diphthongs and velar consonants were noted as the least common.[84]

---

[81] Ibid.
[82] In this word position, almost 900 consonantal combinations can occur.
[83] These were imported to Czech vocabulary through loanwords, so it is understandable that they were not as frequent.
[84] Ibid.

Overall, there was recorded more consonants (60%) than vowels (40%). Distinctive features such as syllabic /r/ (e.g., ⟨prst⟩ "finger") and /l/ (e.g., ⟨vlk⟩ "wolf") were found in less words (0.5%) than the unique /ř/[85] (e.g., ⟨řeka⟩ "river") which occurred more often (in 1.4% of words). The average word consisted of nine phonemes where the most frequent patterns[86] were accounted the following ones:[87]

- CVCVCVCV
- CVCVCCV
- CVCCVCVCV

The lexical corpus further proved certain restriction affecting the distribution of vowel length in Czech language – nearly 80% of all syllabic nuclei consisted of short vowels meanwhile 17% consisted of long vowels.[88] Despite the results, the ratio of short vowels to long vowels differs depending on the word positions of syllables: long vowels are more often found in the word-internal syllables while short vowels occur in 95% before word-final combinations of consonants. The study also concluded new discoveries, such as the observation that the word position seems to affect the frequency of short and long vowels.[89]

3.6. Linguistic distance between Korean and Czech

Czech language, just like English, has syntax consisting of the subject–verb–object composition. So, when it comes to native Koreans learning Czech, it can be assumed that they will encounter similar difficulties when dealing with sentence composition as they did when they were learning English. Moreover, Czech word order has a very loose and flexible structure thanks to declension and conjugation. It is also a null-subject language, meaning the subject (including personal pronouns) can be omitted, if known from context, and the person can be expressed by the verb.

[85] That is generally considered even more unique.
[86] With C representing any non-syllabic phoneme and V marking any syllabic phoneme.
[87] Ibid.
[88] Others were diphthongs and syllabic /r/ and /l/.
[89] Ibid.

Korean falls under the null-subject language category, too. In this case, "not only the subject of the sentence but also any or all of the nominal arguments of a predicate (i.e., verbs, adjectives, and the copula), case markers attached to argument noun phrases, and even the predicate itself, can be dropped."[90] The omitted part is presumably "old information" that the listener will be able to understand from the previous context according to the speaker.

When comparing Korean and Czech, there is also a similar cultural attribute when it comes to using different speech levels to address someone. Even though Korean possesses up to seven different speech levels that vary in their form, Czech has a similar variation; it distinguishes two speech levels (⟨vykání⟩, ⟨tykání⟩) that are used depending on the formality of situation or nature of a relationship between the participants of communication.

In the case of assessing linguistic distance between Czech and Korean based on their linguistic dimension (more specifically from a phonological perspective), as was done in previous chapters, the shortage of sources and articles written on this topic poses as an obstacle. When it comes to comparison of language pairs used for illustrating their linguistic distance from Korean, there is rarely any other language used other than English. Therefore, this chapter will aim to primarily compare Czech phonological system with Korean with the help of previously accumulated information from the last two subchapters. Additionally, there will be some rough estimations of possible similarities and differences within the two phonological systems. Also, English will be considered as a bridge between the two languages where needed.

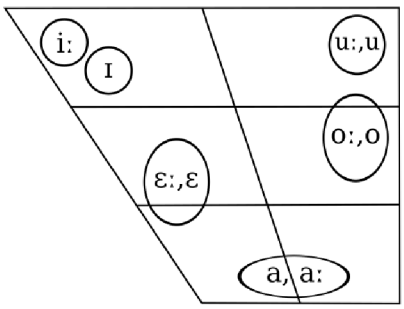

Fig. 6: Czech phonemic vowels

		Labial	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Nasal		m ⟨m⟩	n ⟨n⟩		ɲ ⟨ň⟩		
Plosive	voiceless	p ⟨p⟩	t ⟨t⟩		c ⟨ť⟩	k ⟨k⟩	
	voiced	b ⟨b⟩	d ⟨d⟩		ɟ ⟨ď⟩	(g) ⟨g⟩	
Affricate	voiceless		t͡s ⟨c⟩	t͡ʃ ⟨č⟩			
	voiced		(d͡z)	(d͡ʒ)			
Fricative	voiceless	f ⟨f⟩	s ⟨s⟩	ʃ ⟨š⟩		x ⟨ch⟩	
	voiced	v ⟨v⟩	z ⟨z⟩	ʒ ⟨ž⟩			ɦ ⟨h⟩
Trill	plain		r ⟨r⟩				
	fricative		r̝ ⟨ř⟩				
Approximant			l ⟨l⟩		j ⟨j⟩		

Fig. 7: Czech phonemic consonants

[90] See Lee and Shin 2008.

First category will be focused on comparison of Czech and Korean phonemic vowels.[91] In the Czech vowel inventory, there can be perceived some shared vowels with Korean that are placed at the same positions within the images, such as /ɪ/ (e.g., ⟨litovat⟩ "to be sorry;" or ⟨vykonat⟩ "to perform"), /ɛ/ (e.g., ⟨letadlo⟩ "airplane"), /u/ (e.g., ⟨učesat⟩ "to comb") and /o/ (e.g., ⟨koně⟩ "horses"). For Korean speakers learning Czech, these are the vowel sounds that would be probably the easiest to imitate in pronunciation.[92]

However, Czech phonemic vowels can be divided into two variants: short and long phonemes. These phonemes typically create a pair where the change of length from short to long phoneme is correlative but there are also pairs where some alterations are disjunctive (meaning these pairs can differ in more features).[93] The changes take place in word roots during inflections and derivations, and they also impact prefixes in derivations.[94] In the opposite case, which is when Czech speakers learn Korean, they can face certain difficulties in pronouncing /ɣ/, /ɑ/, and /ɯ/vowels. If they also can speak English, it can become easier for them to correctly imitate at least one vowel – /ɑ/ but the other two are, within the three compared languages, unique to Korean.

The next category is examining mutual bonds between Czech and Korean consonant inventories. In the Czech consonant inventory, there seem to be shared consonants but not all of them are pronounced the same way. For example, [p], [m] can be found in Czech inventory but while in Korean, the sounds are bilabial, in Czech they are labial. Thus, if Korean speakers wanted to imitate these sounds, they would not be foreign for them, but they would need to adapt a new articulation technique. However, consonants [t], [s], [n] are in both languages alveolar, and [k] is also in both cases velar, so these sounds should be easier to produce for Koreans. Conversely, all the remaining Czech consonants from image 7 represent new sounds that are not present in Korean consonant inventory, so it can be assumed that these would be the most difficult to imitate for Korean speakers.

In the case of Czech speakers imitating Korean consonants, most of the phonemes present in the table are foreign. Nonetheless, Korean consonants /pʰ/, /p'/, /th/, /tʰ/, /kʰ/, /k'/, /s'/, /tʃʰ/, /tʃ'/, /ɾ/ may be especially difficult to imitate since some of them represent

[91] See fig. 2 and 6.
[92] Wikipedia online, s.v. "Czech phonology." [Retrieved 1 May 2024.]
[93] These changes in phonemes (/oː/ → /uː/, /uː/ → /o̯u̯/) can be attributed to historical events, similar like the Great Vowel Shift in English.
[94] Ibid.

the same consonant with only altered pronunciation technique. Also, these consonants were marked as challenging to pronounce for English speakers who happened to have comparatively more shared consonants with Czech than with Korean. Thus, it can be assumed that Czech speakers will in this case also share similar challenges at pronouncing Korean consonants, as English speakers did.

Finally, a comparison in phonotactics will be observed. For this comparison, the data was taken from two previous chapters and put together to better examine the differences and/or similarities better. The syllable structures of the three languages are as follows:

- Maximal syllable structure in Korean: CVC(C)
- Maximal syllable structure in English: CCCVCCC
- Maximal syllable structure in Czech: CVCCVCVCV

Amongst the three languages, Korean has the lowest number of allowed consonants while English has the highest, leaving Czech in the middle. The structure of every language's pattern of consonants and vowels is also diverse. Korean follows a regular pattern of consonant–vowel–consonant(consonant) while English possesses larger groups of consonants next to each other before they get divided by a vowel. In the end, Czech also follows mostly regular pattern where every consonant (or at most two) gets separated by a vowel. In this particular case, it may seem as if Czech and Korean have a closer distance between each other than they have with English.

4. Different approaches to assessing linguistic distance

As was examined in previous chapters, linguistic distance can be assessed through different methods. However, there are still other ways in which the linguistic distance can be examined. This final chapter will briefly introduce two more approaches that can help in assessing linguistic distance and mutual similarities present between the respective languages.

Instead of comparing results from different studies, these following subchapters will be more theoretical in nature. Their aim is to present additional possible approaches to measuring linguistic distance that have not been primarily examined in this thesis but

deserve a mention, nonetheless. Also, it should provide a simplified representation of the nature of language and its intricacies from different study perspectives.

4.1. Principle of frequency and computational approach to linguistic distance

As was mentioned in the beginning, there are typically two approaches for the assessment of linguistic distance: rule-based and experience-based. For example, experience-based approach can be seen in experiments where linguistic distance is being measured based on proficiency and knowledge of the given language among language users. Rule-based approach can be used when assessing linguistic distance between two languages based on their attributes and rules.

In order to assess linguistic distance in its complexity, it is inevitable to use both approaches. An integration of both rule-based and experience-based approaches could provide a more in-depth representation of linguistic distance, however, implementing both approaches at once would have been too broad to establish into a model of language processing.[95] Still, a consistent number of experimental findings combined with a mathematical formalisation processed through computer should make the establishment of the model of language processing relatively easier. Especially when working with frequent regularities within the languages, the assessment of linguistic distance should become more accurate and easily accessed.[96]

When observing language, its rules and structure, certain frequency of patterns can be noticed. For this type of predominantly predictable patterns, assessment through mathematical calculations might be useful. This is the focus of computational linguistics.

Computational linguistics is an interdisciplinary field that observes how human language might be automatically processed and interpreted. Research in this field focuses on computational modelling of natural language while taking into consideration its mathematical and logical characteristics. It studies appropriate computational approaches to linguistic questions and develops algorithms and statistical processes for automatic language processing. Some of the fields that computational linguistics relies on are for example linguistics, computer science, artificial intelligence, or mathematics.[97]

---

[95] Chipere 2003.
[96] Ibid.
[97] Wikipedia online, s.v. "Computational linguistics." [Retrieved 4 May 2024.]

Measuring language based on the principles of frequency and contiguity relies on the fact that natural language shows certain statistical regularities and that language users react to these regularities. When measuring linguistic distance with the use of computer programs and mathematical equations, received results should be more precise, without the added factor of human error.

For example, when experiments are performed with language users to assess linguistic distance, many variables can affect the experiment, such as underperformance of participants due to stress. Conversely, when the results are being analysed by computer, the program only works with the available data without any other factors possibly influencing its final results.

However, there are certain downsides when it comes to measuring linguistic distance with the help of computer programs. When applied to Czech, it can be hard to assess the language in its entirety if it can become instead limited by given programs. For example, when a computer program is supposed to evaluate Czech sentences and accepts only certain pre-coded correct answers, it can fail to recognise other correct answers that were not taken into consideration before running the program. When assessing a language with a relatively free word order like Czech, this is particularly relevant and must be considered.

4.2. Assessing linguistic distance through learning acquisition

Certain linguistic characteristics and attributes can become more prevalent during the process of learning new language, especially when it concerns L2 learning or higher.[98] When an individual with already certain language proficiency starts learning new language, the process can be either easier or more difficult depending on the proximity of his native language to this new language.

This is prevalent especially when an individual starts learning new language that belongs to the same language family as their native tongue. Similar bonds and relationships between languages become more obvious in the process of acquiring new language which results in easier and faster understanding of the studied language. However, this applies the other way as well. If L2 language is further from one's native

[98] Some studies that are focused on this topic include e.g., Son, M. (2020), *Cross-Linguistic Syntactic Priming in Korean Learners of English* or Cargnelutti, E. et al. (2022), *Effects of Linguistic Distance on Second Language Brain Activations in Bilinguals: An Exploratory Coordinate-Based Meta-Analysis*.

language, the process of learning the language becomes more demanding because there are no previous familiar connections in the brain. For example, Korean native speakers encounter multiple struggles when they first start learning English language.[99] There are various differentiating aspects between the two that can make the process of learning the new linguistic system difficult.

When learning a foreign language, certain abstract representations are formed in the minds of the learners. One study[100] examines whether L2 learners can develop and share these abstract syntactic representations between L1 and L2, only with different word orders. If that were true, it is worth considering whether distinctive syntactic features of one language can have an influence on the shared representation. It is sure that there is definitely a great potential in approaching assessment of linguistic distance through observing learning processes during acquisition of a new language.

[99] As has been mentioned in the chapter about assessing linguistic distance between Korean and English.
[100] Son 2020.

Summary

Linguistic distance is a difficult concept to capture into some specific measurements. Since it serves as an expression of mutual relationships of all languages, I decided to first examine the nature of language and what it consists of. To better understand the complexity of languages, I divided this thesis into several chapters, each dedicated to a certain aspect that shapes language, and therefore has an influence on the linguistic distance between them.

The aim was to approximate the linguistic distance between Czech and Korean. There are not many sources that cover this topic, so I faced some challenges when searching for relevant material. Because this language pair has not been researched that much, I decided to implement also English that has significantly more studies done on this topic. From that point, I was observing mutual connections and similarities between all three languages, using English as a stepping stone in obtaining information I would otherwise not have access to.

In the beginning of the thesis, the focus was put on cultural influences that shape language and its variation. We could observe how different cultural environments directly affect the accent or vocabulary. For example, this was prevalent with Korean language that has a considerably different variation depending on the country where it is used.

Next, respective language families of Czech, Korean and English were examined. Thanks to the classification of the three languages into larger groups, we could observe some shared similarities and patterns that are unique to a specific language family. This chapter also provided more information about historical background of each language and how it in the end helped in forming the given language to its modern variation. Also, I decided to compare Czech, Korean and English with another language from the same language family. Thanks to this pairing the most obvious similarities between languages disappear and it becomes easier to focus on smaller details that distinguish them.

The following chapter examined the languages in different linguistic dimensions. Then, the relationships between these languages established. I decided to assess the distance/similarity between the pairs based on phonology. In this part, I compared the results of a study that detailed common pronunciation errors which helped me to better identify phonemes that are only present in each of the three languages. In the last part, I briefly mentioned other approaches that seem to be also effective in assessing linguistic distance.

Taking in account theoretical knowledge about possible assessments of linguistic distance, it is evident that there is a bigger linguistic distance between Korean and Czech than there is between Korean and English. English has been becoming a globally spoken language and has influenced also Korean. Due to this influence, Korean keeps adding originally English words with changes in pronunciation and morpheme composition to its vocabulary and accepting these loanwords as its own. Hence, English can be acting as a bridge between the Korean and Czech language to a certain degree, but it can also possibly make the process of assessing linguistic distance more challenging.

Záver

Táto práca skúmala jazykovú vzdialenosť medzi češtinou a kórejčinou. Vzhľadom k tomu, že meranie jazykovej vzdialenosti je pomerne náročný proces, práca sa sústredila predovšetkým na všeobecné začlenenie jazykov do určitých skupín, v rámci ktorých potom boli určité podobnosti/odlišnosti medzi jazykovými pármi zjavnejšie. V úvode sa rozoberala podstata jazyka, jazykového systému a jeho väzieb na ostatné jazyky, s ktorými príde do kontaktu. Ďalej sa skúmali jazykové dimenzie pre každý z posudzovaných jazykov a následne boli v rámci jednej z nich, fonológie, kde sa skúmali ďalšie aspekty jazykovej vzdialenosti pre vytvorené jazykové páry. Na záver bolo potvrdené, že kórejčina je pomerne dosť jazykovo vzdialená od češtiny, ale napriek tomu sa podarilo objaviť aj viaceré podobnosti medzi jazykmi.

Sources:

Bičan, Aleš (2014), "Word Phonology in Czech." *Czech Language News,* Spring 2014, no. 40: 2–4. https://www.academia.edu/97275134/Word_Phonology_in_Czech

Bowern, Claire, and Terry Crowley (2010), *An Introduction to Historical Linguistics*. Oxford: Oxford University Press.

Chipere, Ngoni (2003). *Saussure's Theory of Language.* In: Understanding Complex Sentences: 134–156, 211–224. Palgrave Macmillan, London. https://doi.org/10.1057/9780230005884_7

Chiswick, Barry R., and Paul W. Miller (2004), *Linguistic Distance: A Quantitative Measure of the Distance Between English and Other Languages.* Institute for the Study of Labor. https://docs.iza.org/dp1246.pdf

Cho, Junmo and Hae-Kyeong Park (2006), "A Comparative Analysis of Korean-English Phonological Structures and Processes for Pronunciation Pedagogy in Interpretation Training." *Meta* 51, no. 2: 229–246. DOI:10.7202/013253ar

Dankovičová, Jana (1999), "Czech." *Handbook of the International Phonetic Association* (9th edn): 70–73. Cambridge University Press. ISBN 978-0-521-63751-0

Gamallo, Pablo (2017), *From language identification to language distance.* Physica A: Statistical Mechanics and its Applications, https://www.academia.edu/82492474/From_language_identification_to_language_distance

Heeringa, Wilbert, et al. (2023), "Comparing Germanic, Romance and Slavic: Relationships among linguistic distances." *Lingua* 287, May 2023, 103512. https://doi.org/10.1016/j.lingua.2023.103512

Hutchinson, W.K. (2005), "'Linguistic Distance' as a Determinant of Bilateral Trade." *Southern Economic Journal* 72: 1–15. https://doi.org/10.1002/j.2325-8012.2005.tb00685.x

Jeong, Hyeonjeong, et al. (2007), "Bilingualism: Language and Cognition." *Neurocognitive Approaches to Bilingualism: Asian Languages* 10, no. 2: 175–187. https://doi.org/10.1017/S1366728907002921

Klein, Jared at al. (2018), *Handbook of Comparative and Historical Indo-European Linguistics* 3: p. 1607. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110542431

Lee Amuzie, Grace, and Patti Spinner (2013), "Korean EFL Learners' Indefinite Article Use with Four Types of Abstract Nouns." *Applied Linguistics* 34, no. 4: 415–434, https://doi.org/10.1093/applin/ams065

Matthews, P. H. (2013), "Language families." *Linguistics: A Very Short Introduction.* Oxford, 2003; online edn: p. 45. Oxford Academic. https://doi.org/10.1093/actrade/9780192801487.003.0004

McPhail, Sean A. (2018), "South Korea's Linguistic Tangle: English vs. Korean vs. Konglish: A Study of the Relative Status of Konglish and Its Parent Languages in South Korea." *English Today* 34, no. 1: 45–51. https://doi.org/10.1017/S0266078417000244.

Nam, Hyunjeong (2010), "Konglish, Korean L2 learners' unique interlanguage: Its definition, categories and lexical entries." *Korean Journal of Applied Linguistics* 26, no. 4: 275–308. https://profnam.weebly.com/uploads/6/7/6/1/6761872/0-konglish.pdf

Naughton, James (2005), *Czech: An Essential Grammar*. London: Routledge Press, p. 3–15, 74. ISBN 978-0-415-28785-2

Nerbonne, John, and Erhard Hinrichs (2006), "Linguistic Distances." Proceedings of the Workshop on Linguistic Distances - LD '06. Association for Computational Linguistics. DOI:10.3115/1641976.1641977

Park, Minhyung (2019), "Linguistic Differences in Headlines: Comparison Between Korean and American Newspapers." In: Lewandowska-Tomaszczyk, B. (eds) *Contacts and Contrasts in Cultures and Languages. Second Language Learning and Teaching*. Springer, Cham. https://doi.org/10.1007/978-3-030-04981-2_13

Roach, Peter (2009), *English Phonetics and Phonology: A Practical Course*. Cambridge University Press, p. 26–28, 44–46, 64–66. ISBN 9780521717403

Schirmer, Andreas (2021), "The Drama in the Sentence: Sequence as a Crucial Challenge for Literary Translation from and to Korean." *European Journal of Korean Studies* 21, no. 1: 1–41. DOI: 10.33526/EJKS.20212101.1

Šimáčková, Šárka, et al. (2012), "Czech spoken in Bohemia and Moravia." *Journal of the International Phonetic Association* 42, no. 2: 225–232. DOI:10.1017/S0025100312000102

Sohn, Ho-Min (1999), "Dialects." *The Korean Language.* Cambridge: Cambridge University Press, p. 77–79. ISBN 978-0-521-36123-1

Son, Myeongeun (2020), “Cross-Linguistic Syntactic Priming in Korean Learners of English.” *Applied Psycholinguistics* 41, no. 5: 1223–1247. https://doi.org/10.1017/S0142716420000545.

Starostin, George (2016), *Altaic Languages.* Oxford Research Encyclopedia of Linguistics. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-35

Vovin, Alexander (2009), “Koreo-Japonica: A Re-evaluation of a Common Genetic Origin.” Honolulu, HI: University of Hawaii Press. In: *Hawaii Scholarship Online,* online edn (2016): 3–7. https://doi.org/10.21313/hawaii/9780824832780.003.0001

Vovin, Alexander (2017), *Origins of the Japanese Language.* Oxford Research Encyclopedia of Linguistics. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-277

Yu Cho, Young-mee (2016), *Korean Phonetics and Phonology.* Oxford Research Encyclopedia of Linguistics. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-176

Cargnelutti, Elisa et al. (2022), “Effects of Linguistic Distance on Second Language Brain Activations in Bilinguals: An Exploratory Coordinate-Based Meta-Analysis.” *Frontiers in Human Neuroscience* 15: 744489. DOI:10.3389/fnhum.2021.744489

Internet sources:

101Languages.net (2005), *Learn Czech – Dialects.* https://www.101languages.net/czech/dialects.html

5 Minute English (2024), *Impact of English on Global Languages: A Deep Dive.* https://5minuteenglish.com/impact-of-english-on-global-languages-a-deep-dive/#:~:text=Specific%20Examples%20of%20English%20Influence%20on%20Individual%20Languages,3%20Changes%20in%20African%20and%20Indian%20Languages%20

90 Day Korean (2024), *Konglish – Complete list of Korean loanwords for 2024.* https://www.90daykorean.com/konglish/

Brown, James, "Can All Linguistic Relationships be Explained through Language Family Trees?" University of Aberdeen. https://www.academia.edu/33593319/Can_All_Linguistic_Relationships_be_Explained_through_Language_Family_Trees_Essay?email_work_card=title

Buccini, Anthony F. and William G. Moulton (2024), *Germanic languages.* Encyclopedia Britannica. https://www.britannica.com/topic/Germanic-languages

Czech Time (2024), *Stress and Intonation.* https://www.czechtime.cz/pronunciation/stress-and-intonation/

CzechClass101.com Blog (2020), *Czech Word Order: Loose and Bendy.* https://www.czechclass101.com/blog/2020/08/07/czech-word-order/

Day Translations Team (2018), *Korean Language in North and South Korea: The Differences*. https://www.daytranslations.com/blog/korean-languages/

Everywhere English (2020), *Schwa Pronunciation: The Secret Sound You Need to Speak like a Native.* https://everywhereenglish.eu/free-english-lesson/schwa-pronunciation-the-secret-sound-you-need-to-speak-like-a-native/#:~:text=Words%20can%20be%20very%20similar%20in%20English%20so,this%20right%20can%20affect%20your%20overall%20speaking%20skills.

FluentVista (2023), *The Most Similar Languages: Factors, Families, And Examples.* https://fluentvista.com/which-languages-are-the-most-similar/#:~:text=Language%20Families%20with%20High%20Similarity%201%20Romance%20Languages,known%20as%20Proto-Germanic.%20...%204%20Scandinavian%20Languages%20

Ivanov, V. Vsevolodovich, and Wayles Browne (2021), *Slavic languages*. Encyclopedia Britannica. https://www.britannica.com/topic/Slavic-languageshttps://www.britannica.com/topic/Slavic-languages

Jang, Aejee (2020), *Korean Speech Levels and How To Use Them Properly.* Lingo Deer. https://blog.lingodeer.com/korean-speech-levels/#:~:text=Korean%20formal%20speech%20levels%20%28%EA%B2%A9%EC%8B%9D%EC%B2%B4%29%201%201%29%20Hasoseo-che,one%20of%20the%20most%20commonly%20used%20forms.%20

Kang, Yoonjung (2010), *The emergence of phonological adaptation from phonetic adaptation: English loanwords in Korean*. Phonology, 27(2), 225–253. http://www.jstor.org/stable/40782984

Kim, Hyuna and Laura Smith (2013), *Perception and Production of /l/-/r/ Sounds by Korean Learners of English and German.* Journal of Undergraduate Research, online. Brigham Young University. http://jur.byu.edu/?p=823

Language Magazine (2015), *Bridging the North and South Korean Language Divide.* https://www.languagemagazine.com/2015/03/30/bridging-the-north-and-south-korean-language-divide/#:~:text=After%20over%2060%20years%20of%20separation%2C%20the%20Korean,if%20any%20other%20languages%20are%20included%20at%20all.

Lee, Jin Sook, and Sarah J. Shin (2008), *Korean Heritage Language Education in the United States: The Current State, Opportunities, and Possibilities*. Asia Pacific Center. 20However%2C%20not,the%20predicate%20itself%2C%20can%20be%20dropped%20in%20Korean.

Library.Illinois.edu (2024), *McCune-Reischauer Romanization System for Korean Language.* University of Illinois Urbana-Champaign. https://www.library.illinois.edu/ias/koreancollection/KoreanRomanizationTable/

Martin, Samuel E. (2024), *Korean language.* Encyclopedia Britannica. https://www.britannica.com/topic/Korean-language

McKenna, Linda (2023), *Korean Particles 은/는 vs. 이/가 Explained*. Fluent in Korean. https://fluentinkorean.com/korean-particles-2/

McLaughlin, Scott (2014), “Linguistic Differences Between Korean and English.” Saint Joseph’s University.

https://www.academia.edu/7718593/Linguistic_Differences_Between_Korean_and_English

Michal Miesler (2022), *Bohemians and Moravians: Two bickering nations under one roof.* Kafkadesk Prague Office. https://kafkadesk.org/2022/03/16/bohemians-and-moravians-two-bickering-nations-under-one-roof/

Miller, Jeffrey (2003), "A Word by Any Other Meaning: Konglish." *MED Magazine* (5) March 2003. https://macmillandictionaries.com/MED-Magazine/March2003/05-korean-english.htm

Potter, Simeon and David Crystal (2024), *English language.* Encyclopedia Britannica. https://www.britannica.com/topic/English-language

Racoma, Bernadine (2019), *German Interpreting: Key Differences Between German and English.* Day Translations Blog. https://www.daytranslations.com/blog/german-english-key-differences/

Seoul Korean Language Centre (2022), *Korean words derived from English – Electronics*. https://www.seoulkorean.sg/korean-words-derived-from-english/

Stiles, Matt (2018), *Koreas' unified women's hockey team has exposed a key difference between South and North — their language.* Los Angeles Times. https://www.latimes.com/world/asia/la-fg-koreas-hockey-language-2018-story.html

Temelkova, Kristina (2021), *Slavic languages – Definition, List, Origin, History, Similarities*. Milestone Localization. https://www.milestoneloc.com/slavic-languages/#:~:text=The%20Slavic%20languages%20form%20a%20group%20of%20languages,between%20the%20Dnipro%20River%20and%20the%20Oder%20River.

The Sound of English (2021), *Schwa Sound /ə/*. https://thesoundofenglish.org/schwa/

Ústav pro jazyk český AV ČR (2020), *Phonological Corpus of Czech.* https://ujc.avcr.cz/phword/

Wikipedia, the free encyclopedia (2022), *s.v. Czech conjugations, s.v. Czech declension, s.v. Korean pronouns, s.v. Culture of North Korea, s.v. Juche, s.v. North–South differences in the Korean language, s.v. Czech language, s.v. Language family, s.v. Germanic languages, s.v. History of English, s.v. German language, s.v. Koreanic languages, s.v. Outline of the Korean language, s.v. Altaic languages, s.v. Jeju language, s.v. Computational linguistics, s.v. Czech phonology.* https://en.wikipedia.org

Wilson, Sophie (2016), *10 English Pronunciation Errors by German Speakers.* Pronunciation Studio. https://pronunciationstudio.com/10-english-pronunciation-errors-german-speakers/#:~:text=Like%20many%20languages%2C%20the%20sounds%20%2F%CE%B8%2F%20and%20%2F%C3%B0%2F,is%2C%20th%20at%2C%20th%20ese%20and%20th%20ose.

List of figures: