

Univerzita Karlova

Filozofická fakulta

ÚSTAV ASIJSKÝCH STUDIÍ
ASIJSKÁ STUDIA: KOREANISTIKA

Bakalářská práce

Kristýna Trmalová

Koronavirová krize v tisku KR a KLDR: korpusová analýza lexika

The Coronavirus Crisis in ROK and DPRK's Print: A Corpus-based Lexical
Analysis

Praha 2022

Vedoucí práce: Mgr. Vladislava Mazaná

Poděkování:

Na tomto místě bych chtěla poděkovat Mgr. Vladislavě Mazaně za důvěru v mé schopnosti, nekonečnou trpělivost a ochotu a velmi cenné rady při vedení mé práce. Také děkuji všem blízkým, kteří mi pomáhali a podporovali mě během sbírání dat i samotného psaní práce a vždy mě byli ochotni vyslechnout.

Prohlášení:

Prohlašuji, že jsem bakalářskou práci vypracoval/a samostatně, že jsem řádně citoval/a všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, dne 22. srpna 2022

.....

Podpis

Abstrakt (česky)

Tato práce se zaměřuje na téma koronavirové krize v tisku Korejské republiky a Korejské lidově demokratické republiky. Předmětem analýzy jsou vybrané články z jihokorejského deníku Čosŏn ilbo a severokorejského deníku Rodong sinmun z roku 2020, které mapují počátky koronavirové krize a její vývoj v tisku obou zemí. Využiji metodu korpusové analýzy, konkrétně vygenerování slov s vysokou frekvencí výskytu a klíčových slov. Cílem tohoto výzkumu je identifikace lexikálních odlišností s ohledem na užívání cizího lexika, rozdíly v užívání domácích a sinokorejských výrazů, popis užívání nové specifické slovní zásoby a vliv purifikačního hnutí na jazyk užitý v tisku obou států.

V úvodu práce nastiňuji divergenci korejského jazyka v KR a KLDR a její historický kontext. Dále čtenáře seznámím s korejským lexikem, a především jeho dělením dle původu slov. Následně popisuji konkrétně praktickou stránku a pozadí analýzy, použité nástroje a metody a jejich úskalí.

Hlavní část práce je poté věnována samotné komparativní analýze lexika, kdy nejprve hodnotím rozdíly v lexiku užitém v celém roce 2020 a podíl domácích, sinokorejských a přejatých výrazů a poté se věnuji vybraným klíčovým slovům a jimi tvořeným n-gramům.

V závěru se vyjádřím k problematickým bodům souvisejícím s nejednotnou slovní zásobou korejštiny v KR a KLDR.

Klíčová slova (česky)

Korejština, slovní zásoba, covid, noviny, Korejská republika, KLDR

Abstract (in English):

This thesis focuses on the topic of the coronavirus crisis in the Republic of Korea and the Democratic People's Republic of Korea press. The subject of the analysis are chosen articles published in the South Korean daily newspaper The Chosun Ilbo and the North Korean newspaper Rodong Sinmun in 2020, which map the beginning of the coronavirus crisis and its development in the press of both countries. I will be using the method of corpus analysis, namely generating frequency lists and lists of keywords. The aim of this research is to identify lexical differences with regard to the use of foreign lexicon, differences in the use of native Korean and Sino-Korean expressions, description of the use of some new specific vocabulary and the influence of the purification movement on the language used in the press of both countries.

In the introduction, I outline the divergence of the Korean language in the Republic of Korea and the DPRK and its historical context. Next, I will introduce the reader to the Korean lexicon, especially the classification according to the word origin. Then, I will describe specifically the practical aspects and background of the analysis, the tools and methods used and their difficulties.

The primary part of the work is then devoted to the comparative analysis of the lexicon itself, where I first evaluate the differences in the lexicon used throughout 2020 and the ratio of native, Sino-Korean and borrowed expressions, and then focus on selected keywords and the n-grams they form.

Finally, I will comment on problematic points related to the non-uniform vocabulary of Korean in the ROK and the DPRK.

Klíčová slova (anglicky):

Korean language, vocabulary, covid, newspaper, Republic of Korea, DPRK

OBSAH

1	ÚVOD	7
2	METODOLOGIE	9
2.1	KORPUSOVÁ LINGVISTIKA A VYMEZENÍ TERMÍNŮ.....	9
2.2	MATERIÁL	11
2.3	METODA	12
3	OBEČNÁ CHARAKTERISTIKA KOREJŠTINY	14
4	KOREJSKÁ SLOVNÍ ZÁSoba A JEJÍ DĚLENÍ DLE PŮVODU	16
4.1	SLOVA KOREJSKÉHO PŮVODU (고유어 [KOJUŮ])	16
4.2	SLOVA SINOKOREJSKÁ (한자어 [HANČČAŮ]).....	17
4.3	SLOVA CIZÍHO PŮVODU (외래어, [ŮRÄŮ])	19
4.4	PODÍL UŽITÍ KOREJSKÝCH, SINOKOREJSKÝCH A CIZÍCH SLOV	21
5	PURIFIKAČNÍ SNAHY V KOREJI	23
5.1	JAZYKOVÁ POLITIKA A PURISMUS V KOREJI DO KONCE 2. SV. VÁLKY	23
5.2	KOREJSKÁ LIDOVĚ DEMOKRATICKÁ REPUBLIKA	24
5.3	KOREJSKÁ REPUBLIKA.....	25
6	KOMPARATIVNÍ ANALÝZA LEXIKA	27
6.1	ČOSŮN ILBO.....	27
6.1.1	<i>Slova s nejvyšší frekvencí</i>	27
6.1.2	<i>Klíčová slova</i>	31
6.1.3	<i>Shluky vybraných slov z ČosŇn ilbo</i>	33
6.2	RODONG SINMUN.....	37
6.2.1	<i>Slova s nejvyšší frekvencí</i>	37
6.2.2	<i>Klíčová slova</i>	39
6.2.3	<i>Shluky vybraných slov z Rodong sinmun</i>	42
7	ZÄVĚR	45
8	BIBLIOGRAFIE	47
9	SEZNAM TABULEK A GRAFŮ	49

1 Úvod

Státní útvary na Korejském poloostrově po většinu své historie udržovaly úzký kontakt s čínskými státy a přirozeně docházelo k silnému přejímání mimo jiné kulturních vlivů. Tímto intenzivním kontaktem bylo ovlivněno i písmo a korejský jazyk. Jazykem psané komunikace byla klasická čínština neboli 한문 [hanmun] i poté, co vzniklo původní korejské písmo 한글 [hangŭl]. *Hanmun* se psal pomocí čínských znaků, které se v Koreji přizpůsobovaly fonémickému systému korejštiny. Vzniklo tedy korejské čtení čínských znaků a s ním velmi početná část korejské slovní zásoby, kterou označujeme jako sinokorejskou. S otevíráním Koreje vnějšmu světu přišla korejština do styku s dalšími jazyky. Mimo jiné byla upozaděna japonštinou během okupace v první polovině 20. st. Když došlo po válce k rozdělení Koreje na dva státy – Korejskou republiku a Korejskou lidově demokratickou republiku, obě země cítily potřebu očistit korejštinu od velkého nánosu přejatých slov. Protože ale vlády neměly dostatečně dobré vztahy, aby se dalo uvažovat o společném řešení problému, vznikla situace, která trvá dodnes. Korejšťina má v současnosti dvě standardní formy, jež se v mnohém liší. Dá se tedy předpokládat, že i v oblasti slovní zásoby narazíme na rozdíly v novinách vydávaných v Korejské republice a KLR.

Koronavirová pandemie je nedávné téma, jež se stalo téměř výhradním obsahem tisku a médií po celém světě. Tento fenomén trval od počátku šíření viru na konci roku 2019 po minimálně následující dva roky než proočkovanosť obyvatel, uklidnění situace, a především naléhavější záležitosti COVID-19 z titulních stran vytlačily. Korejský tisk nebyl výjimkou a na téma koronaviru bylo napsáno velké množství článků. S potřebou mluvit o nezvyklé a nové situaci si mluvčí každého jazyka uvykli užívat výrazy, které se s postupující pandemií zautomatizovaly a zažily. Právě zjištění a popis, nakolik se jazyková divergence odrazí ve frekvenci těchto výrazů a klíčových slovech v tisku je hlavním předmětem této práce.

V první kapitole se věnuji podrobnějšímu popisu metodologie, definuji důležité pojmy z oblasti korpusové lingvistiky, přibližuji užitý materiál a postup při výzkumu. Po obecně lingvistickém úvodu upřesňuji rozsah výchozího i referenčního korpusu pro tisk obou Korejí, poté se věnuji konkrétním metodám pro generování slov s vysokou frekvencí a klíčových slov.

V další kapitole seznamuji čtenáře s korejským jazykem a jeho charakteristikou, velmi stručnou genealogií a typologickým zařazením, což považuji za nutnosť pro pohodlnou orientaci v dalších částech práce.

Následně se zabývám korejskou slovní zásobou, a především jejím dělením podle původu a zastoupením lexika v jazyce. Jednotlivé podkapitoly se podrobněji věnují charakteristice domácích korejských slov, sinokorejským výrazům a přejatým slovům z jiných jazyků. Rovněž představuji současnou situaci v korejštině a podíl zastoupení jednotlivých skupin lexika, na než se budu odkazovat dále v praktické části.

Pro pochopení kontextu jazykové divergence a motivace pro zvolení tématu práce je důležitá kapitola zaměřená na purifikační tendence v předválečné Koreji a jazykovou politiku implementovanou separátně v Korejské republice a Korejské lidově demokratické republice. Historický kontext jazykového purismu a následně divergence rozebírá kapitola 5.

Konečně stěžejní část práce obsahuje samotnou korpusovou analýzu doplněnou o tabulky frekventovaných a klíčových slov a grafy. Jednak stavím do kontrastu články z obou zemí za celý rok a také

ukazují podíl tří vrstev slovní zásoby dle původu a jejich vazeb, v nichž se vyskytují. Dále demonstrují užívání klíčových slov na vybraných shlucích a příkladových větách a frázích. Zároveň hodnotím i možné známky purifikačních snah z minulého století.

V závěru shrnuji výsledky analýzy lexika a hodnotím je v kontextu současné situace korejského jazyka s možnými vyhlídkami do budoucna.

Transkripce

V práci užívám pro přepis korejštiny českou vědeckou transkripci. Vzhledem k lingvistickému zaměření práce jsem se rozhodla ponechat korejské písmo přímo v textu i přes možné typografické nedostatky. Severokorejské noviny Rodong sinmun označuji v souladu s oficiálním názvem a se severokorejskou pravopisnou normou, tedy s iniciálou ≡ (na rozdíl od ⊔ v Korejské republice). Rovněž analyzované lexikum a frekventovaná a klíčová slova v tabulkách jsou uvedena v severokorejské podobě. Jihokorejské názvy uvádím dle normy Korejské republiky. Obsah tabulek vztahujících se k Čosön ilbo je taktéž v originální podobě podle pravopisu na Jihu. Pro přepis slov a názvů z čínštiny používám z praktických důvodů pinyin bez tónových značek. Japonština je dle potřeby přepisována českou transkripcí a případně doplněna o původní zápis v hiraganě.

2 Metodologie

2.1 Korpusová lingvistika a vymezení termínů

Korpusová lingvistika pracuje s jazykovým korpusem a zkoumá jej z pohledů jednotlivých jazykových disciplín – syntax, morfologie, lexikologie atd.

„Jazykový korpus je strukturovaný, unifikovaný a často i označovaný (tagovaný)¹ velmi rozsáhlý soubor jazykových dat, elektronicky uložený a zpracováváný (obv. v podobě úhrnu jednotlivých textů), jejichž výběr chce být vzhledem k vytčenému cíli reprezentativní...“². Často ve spojení s jazykovým korpusem mluvíme o pokusu mapovat celý jazyk. Lze ovšem sestavit korpus i z jinak vymezeného vzorku, což je případ této práce. Korpus (resp. dva korpusy), s nímž pracuji, reprezentuje pouze specifický zlomek korejského tisku a všechny výsledky se tedy dají aplikovat pouze na tuto oblast.

„Frekvence se nejjednodušeji a nejčastěji měří v korpusu prostým spočtením všech výskytů jednotlivého slova, formy nebo i slov všech.“³ Zjednodušeně řečeno tedy frekvence znamená kolikrát se slovo vyskytovalo v analyzovaném textu – korpusu.

Seznam slov (word-list) je seznam slov, které se vyskytují v daném textu. Pokud zkoumáme frekvenci jednotlivých slov, pracujeme s typy a tokeny. Typem označujeme jednotlivý slovní tvar, který představuje jednu položku v seznamu, zatímco se v samotném textu může vyskytovat jednou či vícekrát. Token je každý konkrétní výskyt jednotlivého slova (=typu) v textu.⁴ Seznam slov tak uvádí za sebou všechny typy a u nich frekvenci výskytu, tedy celkový počet tokenů každého typu.⁵

Klíčové slovo je slovo, které nemusí mít zvláštní důležitost v daném jazyce, ale je důležité (klíčové) pro daný text, se kterým pracujeme. Pro toto musí splňovat dvě podmínky – vyskytovat se v textu často a hrát v něm signifikantní roli. Míru, nakolik slovo tyto podmínky splňuje, nazýváme klíčovost (keyness).⁶

Referenční korpus slouží jako jakýsi filtr. Slova, která se vyskytují často ve výchozím korpusu splňují jednu podmínku pro to být klíčovými. Pokud se ovšem vyskytují stejně často v korpusu použitým pro

¹ Tagováním rozumíme anotaci korpusu, kdy jednotlivým tokenům přiřazujeme značky/kódy pro lepší orientaci a následnou práci s korpusem, typicky na základě morfologické vlastnosti, příslušnosti k jednomu lexému apod. (Čermák, 2011)

² ČERMÁK, František. *Jazyk a jazykověda*. 2011.s.108

³ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017. s.102

⁴ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017. s.267

⁵ SCOTT, Mike a Chris TRIBBLE. *Textual patterns: key words and corpus analysis in language education*. Amsterdam; Philadelphia: John Benjamins Pub, 2006, s.13

⁶ SCOTT, Mike a Chris TRIBBLE. *Textual patterns: key words and corpus analysis in language education*. Amsterdam; Philadelphia: John Benjamins Pub, 2006, s.55-58

referenci, program je nevyhodnotí jako výrazy, které z výchozího textu nějak zvlášť vyvstávají a nezařadí je do seznamu klíčových slov.⁷

Před vygenerováním seznamu slov je třeba si uvědomit několik věcí. Dle druhu a obsahu textu, který tvoří výchozí korpus, buď zahrneme nebo vyloučíme slova jako čísla a vlastní jména. Jejich vyloučení by vyžadovalo složité konfigurace softwaru, v němž pracujeme (pokud vůbec umožňuje exkluzi takového množství slov a pokud předem víme, jaká vlastní jména se v textu vyskytují), nebo manuální odebírání z textu, což je v praxi náročné a časově velmi neekonomické řešení.⁸

Taktéž v rámci korpusové anotace a konfigurace softwaru musíme rozhodnout, jak důsledně si přejeme dělit složená slova, zda pomlčku považujeme za hranici dvou samostatných slov, zda kapitalizace uvnitř slova znamená jeho rozdělení na části atd. U jazyků, které píší jiným písmem než latinkou, může těchto „šedých“ oblastí v typografii být více. Otázka mezer a dělení slov je komplikovanější třeba u čínského logografického písma, kde se mezi znaky v rámci věty nemusí psát mezery nebo u korejského hláskového písma *hangŭl*, kde je někdy obtížné určit, do jaké míry jsou morfémy volné a lze mezi nimi mezeru napsat.⁹

Další komplikací při práci se seznamem slov je otázka lemmatizace slov. Pokud pracujeme se syntetickým jazykem¹⁰, v textu se bude vyskytovat mnoho výrazů, u kterých je na subjektivním uvážení, zda je považujeme za různá slova nebo je pro účely výzkumu sloučíme do skupiny pod jeden reprezentativní tvar lexému (=lemma).¹¹ K tomuto problému konkrétně pro případ korejštiny se znovu vyjadřuji níže v rámci metody.

S hotovými seznamy lze dále pracovat. Typicky nás zajímá kontext užití daného slova, neboť pouhé slovo samo o sobě nám nepodá další informace kromě frekvence. Soupis výskytů vybraného slova s celým jeho okolím se nazývá konkordance.¹²

⁷ SCOTT, Mike a Chris TRIBBLE. *Textual patterns: key words and corpus analysis in language education*. Amsterdam; Philadelphia: John Benjamins Pub, 2006, s.58-59

⁸ SCOTT, Mike a Chris TRIBBLE. *Textual patterns: key words and corpus analysis in language education*. Amsterdam; Philadelphia: John Benjamins Pub, 2006, s.14

⁹ Psaní mezer je sice upraveno pravopisnou normou, ale v praxi nemusí být tato norma vždy dodržována tak důsledně jako v latině, jak uvidíme v analytické části práce.

¹⁰ Jazyky, které se na základě morfologické typologie řadí mezi flektivní, aglutinační, polysyntetické nebo introflektivní a pro vyjádření gramatických kategorií ohýbají slova a užívají neizolované morfémy. Viz. kapitola 3

¹¹ SCOTT, Mike a Chris TRIBBLE. *Textual patterns: key words and corpus analysis in language education*. Amsterdam; Philadelphia: John Benjamins Pub, 2006, s.14

¹² ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017. s.51

Hit	KWIC
3055	코로나 바이러스 검사에서 음성인 나와 격리가 해제됐고 1명이 신종 코로나 바이러스 감염 여부를 검사 중이라고 했다. 하지만 이 발표 0
3056	원이 성사될지는 불투명하다. 이와 관련, 아사히신문은 한국 정부가 코로나 바이러스 감염 여부를 조사하는 PCR 검사 키트 등을 일본에
3057	과대학과 영국 보건·열대병 대학원(LSHTM)은 개를 훈련시켜 냄새로 코로나 바이러스 감염 여부를 파악할 수 있게 하는 프로젝트에 공동
3058	정26일 서울 은평구 보건소 앞에 설치된 선별진료소를 찾은 시민이 코로나 바이러스 감염 여부를 확인하기 위한 검체 채취를 받기에 앞
3059	. 26일 서울 은평구 보건소 앞에 설치된 선별진료소를 찾은 시민이 코로나 바이러스 감염 여부를 확인하기 위한 검체 채취를 받기에 앞
3060	재명<사진> 경기지사는 2일 오후 신천지 이만희(89) 총회장이 우한 코로나 바이러스 감염 여부 에 대한 검사를 다시 받아야 한다며 직접
3061	재명<사진> 경기지사는 2일 오후 신천지 이만희(89) 총회장이 우한 코로나 바이러스 감염 여부 에 대한 검사를 다시 받아야 한다며 직접
3062	의무 착용하고 있다. 국립현대미술관은 서울관과 과천관 로비에서 코로나 바이러스 감염 예방수칙 을 담은 영상을 상영하고, 확장실과
3063	"고 말했다. 코로나 속 이스라엘 총선 투표 2일(현지 시각) 우한 코로나 바이러스 감염 우려 때문에 격리된 유권자를 위해 마련된 이
3064	"고 말했다. 코로나 속 이스라엘 총선 투표 2일(현지 시각) 우한 코로나 바이러스 감염 우려 때문에 격리된 유권자를 위해 마련된 이
3065	히스패닉 유권자가 바이든에게 몰린 것으로 분석됐다. 또 17일 신종 코로나 바이러스 감염 우려 로 전반적 투표율이 낮아진 가운데 샌더:
3066	히스패닉 유권자가 바이든에게 몰린 것으로 분석됐다. 또 17일 신종 코로나 바이러스 감염 우려 로 전반적 투표율이 낮아진 가운데 샌더:
3067	들에게 약품을 전달하고 침대 시트를 수거했다. 항저우의 한 호텔은 코로나 바이러스 감염 우려 로 호텔에 격리된 승객에게 로봇을 이용하
3068	아니라 이미 미국에 정착한 교민들도 두려움을 느끼고 있다. 단순히 코로나 바이러스 감염 우려 뿐만 아니라, 미국의 사회 안전망이 무너
3069	벚꽃으로 유명한 진해 일원 주요 명소 출입을 전면 통제하기로 했다. 코로나 바이러스 감염 우려 에 국내 최대 벚꽃 축제인 진해군항제를
3070	이다. 수험생이 많이 모이는 온라인 커뮤니티에선 '가정학습 신청 시 코로나 바이러스 감염 의심 증상이 있어 쉽다고 해야 확실하다' 등 '
3071	것이다. 방지환 중앙감염병병원 운영센터장은 "독일 논문에 따르면 코로나 바이러스 감염 6일 에서 10일이면 항체가 생기고, 회복 후 장:
3072	과장이 잘못된 것도 있다. 그는 확진 판정을 받기 전인 지난 8일부터 코로나 바이러스 감염 증상 이 나타났던 것으로 조사됐다. 하지만 곧

Obrazek 1: Ukázka konkordance na slově 코로나 v Čosŏn ilbo

Když zkoumáme vazby daného slova v jednom nebo obou směrech textu (vpravo a vlevo od slova), dostaneme tzv. shluky nebo n-gramy.¹³ Ty mohou představovat kolokace (slova, která se spolu vyskytují s větší pravděpodobností než s jinými a jejich spojení je smysluplné) nebo koligace, tedy spojení slov spíše na základě gramatických funkcí než lexikálních.¹⁴

2.2 Materiál

V této práci podrobují elektronické korpusové analýze vybrané články z korejského tisku z období od 1.1.2020 do 31.12.2020, a především mě zajímají lingvistické rozdíly mezi texty z KLDK a z KR, tudíž je výchozí i referenční korpus tvořen výhradně jimi. Články jsem vybírala z online archivů 조선 일보 *Čosŏn ilbo* a 로동 신문 *Rodong sinmun* dostupných na jejich oficiálních webových stránkách. Rodong sinmun je z dostupných médií v KLDK nejkonzistentnější ve formátu článků, vydává je pravidelně a v době sběru materiálů a psaní práce byl snadno přístupný online, byl proto nejlepším pramenem v dané situaci. Čosŏn ilbo je nejčtenějším deníkem v Korejské republice¹⁵ a rovněž má snadno přístupný archiv článků online i možnost vyhledávání. Jelikož vznik pandemie datujeme do přelomu roku 2019 a 2020 a nejvíce slovní zásoby pravděpodobně vznikalo v jejích počátcích v reakci na potřebu o šíření viru efektivně informovat, je podle mého názoru rok 2020 klíčový pro mé bádání.

Pro selekci článků jsem zvolila hledání klíčového slova „코로나“ [kchorona], které je užíváno v obou státech. Oficiální název nemoci stanovila WHO v únoru 2020 jako „COVID-19“ a virus nazvala

¹³ SCOTT, Mike a Chris TRIBBLE. *Textual patterns: key words and corpus analysis in language education*. Amsterdam; Philadelphia: John Benjamins Pub, 2006, s.19

¹⁴ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017. s.57-59

¹⁵ Chosun Ilbo Has Korea's Biggest Circulation. Dostupné z: http://english.chosun.com/site/data/html_dir/2010/11/30/2010113001011.html.

„SARS-CoV-2“.¹⁶ Pro nemoc se také užívá v angličtině „coronavirus disease“, ve většině jazyků nejčastěji oficiální texty pracují s názvem „COVID-19“, zkráceným „covid“ či variantou slova „corona“, „korona“ apod. Výběrem jednoho konkrétního označení pro koronavirus, které se ujalo v průběhu pandemie, mohlo dojít k vynechání některých článků, které o krizi informovaly bez jejího současného názvu, především v prvních měsících roku 2020. Počet potenciálně nezahrnutých článků však považuji za zanedbatelný, vzhledem k hojnosti materiálu, jež se pohybuje v tisících článků (viz Tabulka 1).

Tabulka 1: Rozsah korpusů užitých v práci

	Články	Typy	Tokeny
Rodong sinmun	2 139	47 003	330 893
Čosŏn ilbo	4 259	198 937	1 829 552

U deníku Čosŏn ilbo byly vybrány pouze články z rubrik zabývajících se společností, světem, politikou a ekonomikou a z několika speciálů věnovaných koronaviru či souvisejícímu tématu. Dohromady obsahují 1 829 552 tokenů.

Ze severokorejského deníku Rodong sinmun byly použity všechny články zahrnující hledané slovo, celkem tvoří korpus o 330 893 tokenech.

Nepoměr v rozsahu dvou korpusů je způsoben nedostatkem materiálu v případě KLDK, která sice stejně pravidelně jako Jižní Korea informovala o počtu nakažených v zahraničí, ale ve srovnání s ní se v mnohem menší míře vyjadřovala k domácí situaci, nenabádala lidi k odpovídající prevenci atd. V důsledku tohoto rozdílu se potom liší i celkový počet vygenerovaných klíčových slov (viz. kapitola 6.1.2).

Do korpusu byl vždy zahrnut text článku a název. Před analýzou jsem z textových souborů odebrala jména autorů článku a dodatečné informace jako místo či datum zveřejnění článku, aby došlo k co nejmenšímu zkreslení výsledků irelevantním kontextem.

2.3 Metoda

Pro tuto analýzu jsem využila volně dostupný software AntConc 3.5.8 od profesora Laurence Anthonyho, který slouží pro práci s korpusem, a především generování seznamů slov na základě různých kritérií, klíčových slov, shluků atd. Pro tuto práci byly užity především funkce Word List, skrze níž jsem získala seznam nejčastěji se objevujících slov, a funkce Keyword List, jež generuje klíčová slova. Ve všech seznamech nebyl brán zřetel na velká a malá písmena, neboť v *hangŭlu* tato opozice neexistuje a v případě nepatrného množství slov v latině, které se v textech vyskytovalo, neměla jejich distinkce výpovědní hodnotu pro účely mé práce.

¹⁶ Naming the coronavirus disease (COVID-19) and the virus that causes it. Dostupné z: [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it).

Frekvence výskytu slov je jednou ze stěžejních informací pro tuto práci a slouží zároveň jako cesta, jak pracovat pouze s relevantním vzorkem, proto je seznam slov řazen sestupně dle jejich frekvence¹⁷ a v analýze se zaměřuji na prvních 100 slov, jejichž frekvence je nejvyšší. Lexikum rozdělují do skupin podle původu slov a hodnotím poměr jejich zastoupení v každém z korpusů.

Ve vygenerovaném seznamu nejfrekventovanějších slov z programu AntConc se často pro jeden lexém setkáme s mnoha morfologickými formami (tzn. tatáž slovesa s různými koncovkami a jmenné výrazy s různými partikulemi), což je dáno aglutinační povahou korejštiny. Vzhledem k rozsahu korpusu jsem se rozhodla ponechat tvary slov v seznamu jako samostatné typy, přestože se tímto může jeden lexém opakovat vícekrát i v rámci vybraných 100 položek. Dalším faktorem, který ovlivňuje výsledky prezentované v seznamu slov, je nekonzistentní užívání mezer mezi slovy, což je problém, jenž byl výrazněji vidět v případě článků z Rodong sinmun.

Pro generátor klíčových slov byla zvolena metoda logaritmu věrohodnosti (log-likelihood), neboť je považována za spolehlivější, nežli chi-kvadrát (chi-square), která při nižší frekvenci dat může vést ke zkresleným výsledkům.¹⁸

K získání klíčových slov slouží jako referenční korpus vždy text článků druhého deníku. Tudíž pro klíčová slova z Čosön ilbo je referenčním korpusem celek tvořený články z Rodong sinmun a naopak. Preferovanou možností je referenční korpus o větším rozsahu nežli analyzovaný text, ale, jak uvádí Scott a Tribble, v případě, kdy to povaha textu a podmínky neumožňují, lze využít i text o méně typech a tokenech.¹⁹ Při generování klíčových slov z textů Rodong sinmun jej software porovnává s několikanásobně větším korpusem z Čosön ilbo a podmínka je tedy dodržena. Ve druhém případě, kdy chceme získat klíčová slova pro Čosön ilbo, musí pro referenci stačit menší korpus.

Předpokládám, že komparací dvou korpusů stejného typu (tedy publicistických textů na podobné téma) budou ze souboru klíčových slov vyřazena slova charakteristická pro publicistický styl, která by se měla překrývat v obou zemích, a snáze vyniknou rozdíly, na něž se tato práce soustředí.

I v případě seznamu klíčových slov pracuji se 100 výsledky s nejvyšší klíčovostí a podrobuji je nejprve jednotlivé analýze a poté opět porovnám do jaké míry jsou zastoupeny skupiny slovní zásoby v KLDR a v KR a jakým způsobem byly tvořeny případné neologismy.

¹⁷ Jiným způsobem práce se seznamem slov by bylo řazení dle abecedy nebo obrácené abecední pořadí, které zohledňuje poslední písmena slov a je užitečné např. při zkoumání tvarosloví a mapování frekvence sufixů. Rovněž lze seznam řadit podle velkých a malých písmen na začátku, počtu slabik atd.

¹⁸ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, Karolinum, 2017, s.102.

¹⁹ SCOTT, Mike a Chris TRIBBLE. *Textual patterns: key words and corpus analysis in language education*. Amsterdam; Philadelphia: John Benjamins Pub, 2006, s. 58-59

3 Obecná charakteristika korejštiny

Jazyky můžeme na základě jejich morfologie a syntaxe třídit do skupin. Tato typologie pracuje s převažujícími rysy v jazyce na základě toho, jaké využívá gramatické prostředky. Žádný z jazyků neodpovídá čistě jednomu typu a může v průběhu svého vývoje měnit dominantní rysy. Vyčleňujeme jazyky flektivní, izolační (=analytické), aglutinační, introflektivní a polysyntetické.²⁰

Flektivní jazyky ohýbají slova pomocí koncovek připojovaným ke kořeni nebo kmeni slova. Tyto koncovky mohou mít zároveň více gramatických funkcí a typicky nedochází ke kumulaci více koncovek za sebou. Uspořádání větných členů je volnější. Mezi flektivní jazyky se řadí mnoho indoevropských jazyků jako čeština nebo francouzština.

Příkladem izolačního neboli analytického jazyka by byla čínština a angličtina. Jde o jazyky, které mají striktní slovosled a namísto skloňování a časování využívají pomocné volné (=izolované) gramatické morfémy. Slova v nich zřídka mění svůj tvar.²¹

V introflektivních jazycích podléhá změnám kořen slova, v němž dochází k alternaci hlásek. Tímto způsobem vyjadřuje gramatické funkce např. arabština.

Polysyntetické jsou jazyky s velkým množstvím plnovýznamových morfémů připadajících na slovo, kde je často obtížné jasně rozlišit jejich hranice. Mezi ně bychom řadili např. inuitské nebo aztécké jazyky.²²

Korejština je typologicky jazyk s převážně aglutinačními rysy – 교착어 [kjočchagö]. To znamená, že v ní dochází k vyjadřování gramatických funkcí pomocí morfémů, které se připojují před či za kořen nebo kmen slova. Tyto v případě korejštiny početné afixy se řadí za sebe podle daných pravidel a většinou platí, že jeden morfém odpovídá právě jednomu gramatickému významu. Výrazněji jsou zastoupeny sufixy. Ke jmenným výrazům se pojí partikule a k verbálním výrazům koncovky.²³

V praxi tedy například z lexému student 학생 [haksäng] můžeme vytvořit postupným přidáváním gramatických funkcí slovo o mnohem více slabikách: 학생들 [haksängdül] – studenti (přidán sufix -들 [tül] vyjadřující plurál), 학생들에게 [haksängdurege] – studentům (pádová partikule dativu -에게 [ege]). Zatímco třeba v češtině, tedy jazyce, který převážně využívá flexi, vyjádříme plurál i pád jedinou koncovkou -ům.

Jako důsledek aglutinace lemma student 학생 [haksäng] reprezentuje mnoho typů: 학생이 [haksängi] (nominativní partikule), 학생은 [haksängün] (tematická partikule), 학생들 [haksängdül] (plurál), 학생들이 [haksängduri] (plurál + nominativ), 학생들에게서는 [haksängduregesönün] (plurál + genitiv + téma).

Z hlediska genealogického zařazení je korejština (prozatím) považována za izolovaný jazyk, neboť nám není známé prokazatelné množství podobností v nejstarší korejštině, které by poukazovaly na příbuznost s jinými jazykovými rodinami. Stejně je na tom japonština, s níž je korejština řazena do jedné jazykové skupiny. Jedna z vědeckých teorií pracuje s podobností slovní zásoby mezi korejštinou a čínštinou, která je

²⁰ ČERMÁK, František. *Jazyk a jazykověda*. 2011. s.231

²¹ Ibid. s.229-230

²² Ibid.

²³ PUCEK, Vladimír. *Gramatika korejského jazyka*. Praha: Karolinum, 2005. s.60-61

však způsobena bohatou vrstvou sinokorejského lexika v korejštině (viz. 4.2) a nevysvětluje původ originálních korejských slov. Když porovnáme tyto dva jazyky také z pohledu gramatiky, uvidíme, že jde o naprosto odlišný systém, kdy proti aglutinační korejštině stojí čínština, která je jazykem silně analytickým. Druhá teorie právě na základě typologie korejštinu přidružuje k altajské jazykové rodině, kam patří např. mongolština, turečtina, finština nebo maďarština. Od této rodiny by se však korejština musela oddělit velmi brzy a vyvíjet se samostatně, abychom vysvětlili výrazně odlišné lexikum.²⁴

Rovněž existuje teorie, která korejštinu pojí s japonštinou, jelikož je geograficky blízká a strukturou velmi podobná. V minulosti byly pokusy o dokázání podobnosti v původním lexiku²⁵, ani tato teorie však zatím nebyla přijata jako spolehlivá.

²⁴ PUCEK, Vladimír. *Gramatika korejského jazyka*. Praha: Karolinum, 2005. s.13-15.

²⁵ YI, Ki-mun a S. Robert RAMSEY. *A history of the Korean language*. Cambridge; New York: Cambridge University Press, 2011. s.26-30

4 Korejská slovní zásoba a její dělení dle původu

Korejská slovní zásoba uvedená v otevřeném internetovém slovníku 우리말샘 [urimalsäm] čítá přes 1 145 000 slov. Zahrnuty jsou obě varianty spisovné korejštiny, tedy standardní jazyk 표준어 [pchjodžunö] v Korejské republice i kulturní jazyk 문화어²⁶ [munhwaö] v Korejské lidové demokratické republice, dialekty provincií obou států a archaismy.²⁷

Obohacování slovní zásoby v korejštině probíhá typickými procesy. Tedy: skládáním, odvozováním pomocí prefixů a sufixů, zkracováním, změnou slovního druhu a přejímáním.

Pro účely jazykovědného výzkumu můžeme slovní zásobu jazyka rozdělovat podle několika kritérií na jednotlivé vrstvy. Můžeme slova dělit na spisovná a nespisovná, citově zabarvená a neutrální, podle slohových příznaků můžeme vyčlenit slova hovorová, knižní apod. a také podle původu můžeme rozlišovat slova původní a slova přejatá. Právě původ slov a zastoupení v současném tisku považují za oblast, ve které budeme pozorovat největší známky jazykové divergence způsobené nejednotnou jazykovou politikou obou států.

Pokud budeme dělit korejské lexikum dle jeho původu, standardně se klasifikuje do tří skupin. Jedná se o původní korejská slova (고유어 [kojuö]), slova, která vychází z čínských znaků – tzv. sinokorejská (한자어 [hanččaö]) a slova přejatá z cizích jazyků, ale zdomácnělá natolik, že jsou považována za součást korejského lexika a podléhají pravidlům korejštiny (외래어 [öräö]). Každá skupina má svá specifika, dominuje v jiném slohovém stylu a věcný význam výrazů často přímo odkazuje na jinou oblast (tj. liší se ve slohové příznakovosti, šíři významu a denotaci).

4.1 Slova korejského původu (고유어 [kojuö])

Původní čistě korejská slova, která se vyznačují tím, že prakticky nesdílí podobnost s jinými jazyky z geograficky blízkých oblastí. To je jeden z důvodů, proč je problematické korejštinu genealogicky zařadit.

Ve Slovníku standardní korejštiny jsou definována takto:

„해당 언어에 본디부터 있던 말이나 그것에 기초하여 새로 만들어진 말. 국어에서는 ‘아버지’, ‘어머니’, ‘하늘’, ‘땅’ 따위가 있다.“²⁸

„Slova, která byla původně v daném jazyce nebo slova tvořená na jejich základě. V korejštině jsou to slova jako 아버지 [abödži] – otec, 어머니 [ömöni] – matka, 하늘 [hanül] – nebe, 땅 [ttang] – země.“

²⁶ Ke vzniku dvou variant spisovného jazyka došlo po rozdělení Korejského poloostrova na dva státy. Termín 문화어 byl poprvé použit Kim Ilsöngem v 60. letech (viz. kapitola 5.2).

²⁷ Sadžön tchonggje 사전 통계 Statistika slovníku. Dostupné z: <https://opendict.korean.go.kr/service/dicStat>

²⁸ Pchjodžungugö täsadžön

<https://stdict.korean.go.kr/search/searchView.do?searchKeyword=%EA%B3%A0%EC%9C%A0%EC%96%B4>

Představují tudíž nejstarší složku lexika a zahrnují mimo jiné dva typy slov, které jsou pro korejštinu velmi typické.

Prvním jsou gramatické morfémy, jež se v korejštině používají v hojném počtu.²⁹ Patří mezi ně jmenné partikule (tematická partikule ~은/는 [ŭn/nŭn], podmětová partikule ~이/가 [i/ka], partikule akuzativu ~을/를 [ŭl/rŭl], partikule instrumentálu ~로/으로 [ro/ŭro] a další), konverba neboli koncovky neukončeného přísudku (slučovací ~고 [ko], odporovací poměr ~지만 [čiman], souběžnost dějů ~면서 [mjŏnsŏ] apod.), a koncovky ukončeného přísudku (~버니다/습니다 [mnida/sŭmnida], ~아/어/여요 [a/ŏ/jŏjo], ~다/ㄴ다/는다 [ta/nda/nŭnda] apod.).

Druhým zajímavým typem jsou onomatopoeia, která mají v korejštině unikátní zastoupení a rovněž tvoří velmi početnou skupinu. Sem bychom řadili výrazy jako 반짝반짝 [panččak panččak] – zářivě, třpytivě nebo 두근두근 [tugŭn tugŭn] – zvuk tlukoucího srdce.

Korejského původu je jedna sada číslovek 1-99³⁰, základní dějová a kvalitativní slovesa, zájmena, názvy jídel, přírodních jevů, tradičních oděvů atd.³¹

Právě původní korejská slovní zásoba oslabovala kontaktem s dalšími jazyky v průběhu historie a její velká část se přestala aktivně používat. Částečně se tento fenomén snažili zvrátit korejští vlastenci v rámci národního obrození, kdy původní korejská slova posilovala pouto k vlastnímu jazyku a národu.³² Další příznivé období pro původní lexikum představuje purifikační hnutí, které začíná v 60. letech 20. století v KLRD a v 70. letech v Korejské republice a podrobněji se jim věnuje následující kapitola.

4.2 Slova sinokorejská (한자어 [hanččaŏ])

V průběhu historie docházelo mezi korejskými a čínskými státy k intenzivnímu kontaktu. Díky němu se postupně spolu s jinými kulturními vlivy dostalo do Koreje čínské znakové písmo a v návaznosti na něj pronikala do korejské slovní zásoby čínská slova, která postupně asimilovala a zdomácněla.³³ Zcela podléhají korejské gramatice a mají schopnost na sebe vázat korejská gramatická slova, ale i tvořit kompozita s původními korejskými morfémy³⁴, proto panuje konsenzus pracovat s čínskými výpůjčkami jako se zvláštní kategorií.

²⁹ Toto je dáno především již zmíněnou aglutinační povahou korejštiny. V kontrastu může stát opět čínština, která jakožto analytický jazyk obsahuje podstatně méně gramatických slov. Dává proto smysl, že většina těchto morfémů v korejštině nebude sinokorejského původu.

³⁰ Korejšťina používá dva druhy číslovek – jednu sadu menších čísel domácího korejského původu a druhou sinokorejskou pokračující do mnohacíferných čísel. Zpravidla se korejské číslovky pojí s korejskými morfémy a sinokorejské číslovky se sinokorejskými morfémy. Korejská gramatika předepisuje kontext použití pro oba druhy číslovek.

³¹ PUCEK, Vladimír. *Lexikologie korejštiny*. Praha: Karolinum, 2013, s. 22.

³² YI, Ki-mun a S. Robert RAMSEY. *A history of the Korean language*. Cambridge; New York: Cambridge University Press, 2011. s.302

³³ Ibid. s.235, 282,

³⁴ PUCEK, Vladimír. *Gramatika korejského jazyka*. Praha: Karolinum, 2005. s.15, 60

Slovník je definuje jednoduše jako „한자에 기초하여 만들어진 말“ tedy „slova vzniklá na základě čínských znaků.“

Čínština byla již ve starověku jazykem asociovaným se vzděláním, a to především díky písemnictví, jež do Koreje proudilo. Následně díky vzestupu konfucianismu a potřebě studovat kanonické knihy v klasické čínštině³⁵ začala být sinokorejská vrstva lexika vnímána jako „vyšší“ a preferovaná především v kruzích učenců. Do současnosti převládá v publicistice a odborných textech, kde by v indoevropských jazycích paralelně dominovaly termíny z latiny nebo řečtiny.³⁶

Sinokorejská slova jsou složena z morfémů, které odpovídají jednomu čínskému znaku. V čínštině má každý znak kodifikovány dvě složky – sémantickou a zvukovou neboli co znak znamená a jak se čte. Toto platí i pro morfémy, které převzala korejština, akorát s rozdílem, že se jejich zápis přizpůbil hláskovému písmu *hangül*. V textu nebude znak, ale přepis jeho zvukové složky. Významovou složku si uvědomí pouze člověk, který zná příslušný znak nebo alespoň jeho čtení a význam, pokud ne grafickou podobu. Velkou výhodou sinokorejské slovní zásoby je ovšem snadné porozumění novým, pro mluvčího dosud neznámým slovům plynoucí ze znalosti významu jednotlivých morfémů.

Jelikož je čtení znaků omezeno fonologickým systémem čínštiny, vzniká i v samotné čínštině mnoho stejných homonymních slabik. Takto se morfémy dostaly stále jako homonyma i do korejštiny, přestože jejich zvuk zpravidla není shodný se čtením v čínštině. V případě čínštiny částečně tento problém zmírňuje užívání čtyř tónů, které však současná korejština nepoužívá, tudíž je odkázána na poznání významu z kontextu, či doplnění znaku pro upřesnění. Pokud se znak 江 znamenající řeka a čtený korejsky *kang* zapíše pouze jako 강, bude neodlišitelný od znaku 強 se stejným čtením a významem silný.

V korejštině buď tyto morfémy fungují jako plnovýznamová jednoslabičná slova (například závislé substantivum 후(後, [hu]) – později/po, numerativ 병(瓶, [pjöng]) – lahev), která však často nemohou stát ve větě samy o sobě, nebo se spojují do víceslabičných slov (např. 공부(工夫, [kongbu]) – studium, 내일(來日, [näil]) – zítra, 중국(中國, [čungguk]) – Říše středu/Čína).

Jak bylo zmíněno výše, zajímavostí je, že, ač obecně platí, že sinokorejské morfémy se pojí s jinými sinokorejskými morfémy, jejich integrace do korejského jazyka funguje do té míry, že ani slova tvořená spojením původního morfému se sinokorejským nejsou ojedinělým jevem. (Jako v případě slova váza – 꽃병 [kkotppjöng], kde se pojí korejský morfém 꽃 [kkot] – květina se sinokorejským 병 [pjöng] – lahev. Zde existuje však i varianta čistě sinokorejská 화병 [hwabjöng]. Ze substantiv sinokorejského původu lze také často utvořit sloveso spojením s korejským verbem 하다 [hada] – dělat. Např.: 공부하다 [kongbuhada] – studovat, „dělat studium“.) Tato jednoduchá slova a kompozita se buď řadí pod jednu ze skupin, nebo se vyčleňují zvlášť jako vlastní kategorie. V této práci jsou pro detailnější výsledky uváděna samostatně a je s nimi zacházeno jako s další skupinou slovní zásoby.

³⁵ Studium a ovládnutí konfuciánských klasik bylo nutným předpokladem pro složení úřednických zkoušek 과거, které jedinci umožňovaly získat uznávanou pozici ve státní správě. Nejen proto bylo (a ostatně stále dodnes je) vzdělání velmi váženou kvalitou.

³⁶ YI, Ki-mun a S. Robert RAMSEY. *A history of the Korean language*. Cambridge; New York: Cambridge University Press, 2011. s.98, 235-236

Kromě slov z čínštiny se sem řadí také slova vzniklá korejským čtením znaků, jimiž se zapisovala např. japonská slova, došlo tedy k přejímání nepřímě. Také neplatí, že spojení sinokorejských morfémů ve slova má vždy stejný znakový zápis v čínštině. Některé složeniny jsou čistě korejskou záležitostí a v čínštině se pro denotát používá zcela jiný výraz.

Sinokorejská slovní zásoba představuje druhou sadu číslovek, vědecké termíny, názvy moderních vynálezů ad. V mnoha případech existuje synonymní výraz v domácí vrstvě lexika a liší se buď konotací nebo pouze schopností fungovat jako samostatný větný člen, a tudíž možnostmi užití. Jako příklad mohou sloužit dvě slova, která obě znamenají „den“: korejské 날 [nal] (např. 좋은 날 [čohŭn nal] – dobrý/pěkný den) vs. sinokorejské 일 [il] (5 일 [o il] – pátý den v měsíci, pátého). V těchto případech slova nelze zaměňovat, přestože se jedná o synonyma.

I když jsou sinokorejské výrazy v jazyce již pevně zakořeněné a u mnoha z nich si ani Korejec neuvědomuje jejich etymologii, probíhaly a stále probíhají snahy upřednostňovat původní slovní zásobu, kde je to možné.

4.3 Slova cizího původu (외래어, [öräö])

„...외국에서 들어온 말로 국어에서 널리 쓰이는 단어. 버스, 컴퓨터, 피아노 따위가 있다.“

외래어 [öräö] jsou „slova, která pochází z cizího jazyka a jsou široce používána v korejštině.“³⁷

Slova cizího původu jsou slova přejímaná z jiných cizích jazyků s částečnou výjimkou čínštiny, která je, jak bylo zmíněno výše, vyčleňována jako vlastní kategorie.³⁸ Zvláštní místo zde nepochybně zastávají výpůjčky z japonštiny, především kvůli nucenému přebírání v době okupace Korejského poloostrova Japonskem v letech 1910-45. Této početné skupiny byla korejština systematicky zbavována po osvobození země, takže v dnešní době je její vliv méně patrný než některých dalších jazyků. Přesto se v korejštině stále vyskytují původně japonská slova, která nejsou spisovnou variantou, ale používají se denně, zatímco existuje korejský ekvivalent. Například pro označení rybího koláčku, jenž je inspirován Japonskem, bývá použito buď slova 오뎡 [odeng], nebo korejského výrazu 어묵³⁹ [ömuk], který je také správnou variantou dle Slovníku standardní korejštiny (표준국어대사전 [pchjodžungugö täsadžön]) vydaného Národním institutem korejského jazyka (국립국어원 [kungnip kugöwön])⁴⁰. V KLR se setkáme s čistě korejským označením 물고기떡 [mulkkogittök].

³⁷ Pchjodžungugö täsadžön

<https://stdict.korean.go.kr/search/searchView.do?pageSize=10&searchKeyword=%EC%99%B8%EB%9E%98%EC%96%B4>

³⁸ Najdeme slova cizího původu i z čínštiny. Typicky však patří mezi sinokorejská.

³⁹ Výpůjčka z japonštiny pochází ze slova *oden* (おでん), do korejštiny foneticky přepsáno jako 오뎡 [odeng]. Tento název pro rybí koláček na špejli vznikl nesprávným přenesením významu z původního *oden*, které označuje jiný druh japonského pokrmu. 어묵 [ömuk] je spojením sinokorejského morfému 魚 어 [ö], jenž znamená „ryba“ a korejského morfému 묵 [muk].

⁴⁰ Pchjodžungugö täsadžön <https://stdict.korean.go.kr/>

Starší výpůjčky přišly do korejštiny především z mongolštiny a sanskrtu. Mongolština ovlivnila korejštinu v období mongolské nadvlády nad Korejí ve 13. století. V současné korejštině není přejímání slov z mongolštiny běžnou praxí a z historie zůstalo pouze skromné množství výrazů ze specializovaných oblastí jako je chov koní a sokolnictví.⁴¹

Sanskrt sloužil jako výchozí jazyk pro většinu výrazů spjatých s buddhismem. Stejně jako samotné sútry a výklady učení i slovní zásobu z velké části čerpali Korejci spíše z Číny než z Indie. Jelikož fonologický fond čínštiny je velmi omezující, při zápisu cizích slov čínskými znaky dochází k odchýlení od původního znění. Korejským čtením čínských znaků se zvuková podoba slov posouvá ještě více. Mezi tuto slovní zásobu patří slova jako Buddha – 부처 [pučchö], Maitréja – 미륵 [mirük] atd.

V bližší historii od otevření Koreje, tedy přelomu 19. a 20. století, se korejština skrze cestovatele a misionáře střetávala s geograficky vzdálenějšími evropskými jazyky.⁴² Jednalo se konkrétně o francouzštinu, němčinu, italštinu a španělštinu.

V této práci a v článcích, kterými se zabývá, budou pravděpodobně nejvýznamněji zastoupena přejatá slova původem z angličtiny a ruštiny. Oba tyto jazyky byly v úzkém kontaktu s korejštinou hlavně od druhé poloviny 20. století, což je dáno faktem, že po kapitulaci Japonska a konci druhé světové války se Korea nedobrovolně ocitla pod správou Sovětského svazu a Spojených států amerických.

Ruština je relevantní především v případě KLR, protože SSSR se zhostil vlády nad severní polovinou Korejského poloostrova a docházelo zde k přijímání ideologie, čímž byl jazyk ovlivněn společně s myšlením. Sice se očišťování jazyka a snaha upřednostňovat korejskou vrstvu slovní zásoby týkaly i rusismů, ale přesto, jak uvidíme ve výsledcích analýzy v dalších kapitolách, hraje ruština roli v divergenci současného lexika. Minimálně ovlivnila pravidla pro přepis cizích slov do korejštiny,

Vzhledem k politické situaci v průběhu poválečného období a mezinárodním vztahům Spojených států s Korejskou lidově demokratickou republikou se může zdát, že se zastoupení slov původem z angličtiny bude týkat výhradně Jižní Koreje, která byla vedena vojenskou vládou armády Spojených států v Koreji (USAMGIK). Vliv angličtiny je v Korejské republice pocíťován nepochybně výrazně více než v Severní Koreji, ale k přejímání docházelo v menší míře i tam.

V Jižní Koreji je angličtina jedním z hlavních vyučovaných cizích jazyků v průběhu školní docházky i po ní.⁴³ Také silný příliv americké populární kultury, ekonomická spolupráce a turismus přispívají k výraznému přejímání anglicismů, ať už do standardního spisovného jazyka nebo jako trendy slangové výrazy mladších generací. I pro tato slova existují často korejské či alespoň sinokorejské ekvivalenty, ale stejně jako v případě sinokorejských slov může být rozdíl v šíři významu či kontextu, kdy se slova užívají, a nemusí nutně být libovolně zaměnitelná. Mezi zdomácnělé anglicismy patří například 카드 [kchadü] – karta, 팝 [pchap] – pop (hudební žánr), 커피숍 [kchöpchi sjop] – kavárna, 원피스 [wönpchisü] – šaty. Rovněž mnoho vědeckých termínů a neologismů se odvozuje z angličtiny (viz. slovo 코로나바이러스 [kchoronabairösu] v kapitole 6).

⁴¹ YI, Ki-mun a S. Robert RAMSEY. *A history of the Korean language*. Cambridge; New York: Cambridge University Press, 2011.s.96-97

⁴² ECKERT, Carter J. *Dějiny Koreje*. Praha: Nakladatelství Lidové noviny, 2001.

⁴³ PUCEK, Vladimír. *Lexikologie korejštiny*. Praha: Karolinum, 2013.s.29

Slova cizího původu mají schopnost vázat na sebe prvky korejské gramatiky (tzn. partikule a koncovky) a fungují v rámci věty stejně jako původní korejské i sinokorejské lexikum. Stejně jako sinokorejská slova mohou spojením s korejskými nebo sinokorejskými morfémy tvořit kompozita a lze je verbalizovat pomocným slovesem 하다 [hada].⁴⁴

Pro demonstraci vlivu jednotlivých jazyků mohou sloužit opět data z Urimalsäm o původu jednotlivých slov cizího původu evidovaných ve slovníku.⁴⁵ Zde je uvedeno deset jazyků, z nichž pochází nejvíce výpůjček. Angličtina je zdaleka nejčastějším zdrojem pro přejímání. Patrný je stále i vliv japonštiny a mnoho slov pochází také například ze zmíněné ruštiny a dalších jazyků.

Tabulka 2: Výchozí jazyky pro výpůjčky cizího původu v korejštině

	Původní jazyk	Počet hesel
1	Angličtina	150 984
2	Francouzština	3 221
3	Němčina	2 258
4	Japonština	1 471
5	Italština	1 121
6	Latina	900
7	Řečtina	738
8	Španělština	475
9	Ruština	416
10	Čínština	178

4.4 Podíl užití korejských, sinokorejských a cizích slov

Lingvisté se snažili mnohokrát zmapovat podíl jednotlivých skupin v jazyce buď analýzou hesel oficiálních slovníků, nebo se zaměřením na vymezený materiál (např. odborné texty, mluvený jazyk, krásnou literaturu, publicistický styl atp.). Podíl zastoupení skupin se liší právě v závislosti na oblasti a stylu. V krásné literatuře a každodenních konverzacích můžeme očekávat vyšší výskyt původní domácí slovní zásoby, především pro její „lidovější“ charakter. Naopak v odborných pracích a v publicistice, která tvoří výchozí materiál pro tuto práci, lze předpokládat vyšší frekvenci sinokorejských výrazů, neboť představují formálnější prostředí, jež vyžaduje přesnou terminologii. Zejména lze očekávat tento rozdíl u tisku ze Severní Koreje, kde purifikační hnutí více ovlivnilo i jazyk běžného každodenního života. Pro ukázkou porovnejme například slovo pro (vstupní) dveře v Jižní Koreji – 출입문 [čchurimmun] (出入門), které je složeno ze tří sinokorejských morfémů, zatímco v „severokorejštině“ bychom se setkali s 나들문 [nadŭlmun], kde byly první dva znaky „přeloženy“ – nahrazeny původní korejskou slabikou se stejným významem.

Jak jsem zmínila, určitě spolehlivě a průkazně podíl zastoupení hlavních tří skupin původu slovní zásoby pro celou korejštinu není dobře možné. Pokud opět nahlédneme do slovníku 우리말샘 [urimalsäm],

⁴⁴ PUCEK, Vladimír. *Lexikologie korejštiny*. Praha: Karolinum, 2013, s.35

⁴⁵ *Sadžön tchonggje* 사전 통계 Statistika slovníku. Dostupné z:

získáme alespoň hrubou představu, jak si jednotlivé vrstvy lexika stojí. Toto je upravená tabulka týkající se běžných slov (일상어 [ilsangŏ]) a odborných termínů (전문어 [čönmunŏ]).⁴⁶

Tabulka 3: Zastoupení lexika dle původu ve slovníku Urimalsäm

Původ	Běžná slova		Odborné termíny	
	Počet hesel	Podíl (%)	Počet hesel	Podíl (%)
Korejský	231 702	48,018	22 882	8,47
Sinokorejský	160 619	33,287	160 872	59,545
Cizí	5 828	1,208	45 259	16,752
Sinokorejský + cizí	2 953	0,612	11 461	4,242
Sinokorejský + korejský	79 019	16,376	27 972	10,354
Cizí + korejský	1 157	0,24	1 247	0,462
Sino. + cizí + korejský	1 254	0,26	474	0,175

Data ze slovníku ukazují že v běžném jazyce převládají původní korejská slova⁴⁷, zatímco v rámci odborné terminologie jsou na první příčce sinokorejská slova, která tvoří dokonce více než polovinu slovní zásoby. Slovník rovněž samostatně vyčleňuje kompozita, jež byla vytvořena kombinací morfémů nesterjného původu. Obzvláště kombinace sinokorejského morfému s původním korejským se v obou oblastech ukazuje jako nejproduktivnější. Zastoupení slov cizího původu je výrazně vyšší mezi termíny, zatímco v oblasti běžného jazyka tvoří pouze přes 1 % hesel. Navíc pouhý 8,5% podíl domácí slovní zásoby v odborné terminologii je zajímavým protipólem ke snaze upřednostňovat ji ve veřejném sektoru a naznačuje, že v tisku zaměřeném na téma zdravotnictví, ekonomiky a světového dění skutečně můžeme předpokládat vyšší podíl sinokorejských slov, která ustála purifikační hnutí.

Co nelze z těchto dat určit, je to, jak moc se liší současný podíl korejských, sinokorejských a cizích slov ve standardním jazyce KR a kulturním jazyce KLDR, protože slovník nenabízí statistiku, která by zahrnovala pouze spisovný jazyk, ani data zvlášť pro každou zemi.

⁴⁶ *Sadžön tchonggje* 사전 통계 Statistika slovníku. Dostupné z: <https://opendict.korean.go.kr/service/dicStat>

⁴⁷ Zde je třeba brát v potaz, že slovník obsahuje i archaická slova a dialekty. Část tohoto součtu mohou tvořit původní korejské výrazy, které byly nahrazeny sinokorejským nebo cizím ekvivalentem, a lokální slovní zásoba, jež není aktivně užívána.

5 Purifikační snahy v Koreji

Již několikrát bylo v práci zmíněno purifikační hnutí a jeho vliv na současnou podobu korejštiny v obou korejských státech. Jazykový purismus je jev, se kterým se setkává někdy ve své historii prakticky každý jazyk. Lze jej definovat jako snahu oprostit jazyk od nových cizích vlivů a změn, které puristé vnímají jako negativní, a kodifikaci jedné správné normy, která je viděna jako čistá.⁴⁸ Týká se období, kdy jazyk přichází do kontaktu s jiným, a v důsledku ustupuje používání původní domácí slovní zásoby a gramatických forem a je nahrazováno cizími variantami. V případě češtiny si s purismem často spojujeme období národního obrození. U korejštiny pozorujeme nejvýraznější tendence od cca 2. pol. 19. st., kdy byly součástí nacionalistického osvíceneckého hnutí.

5.1 Jazyková politika a purismus v Koreji do konce 2. sv. války

Korejské království Čosön přicházelo do kontaktu se západními státy a jejich činnost nejen v Číně probouzela v lidech proticizinecké nálady. Zároveň pro posílení vlasteneckého ducha bylo nutné nejprve snížit podíl negramotného obyvatelstva, aby poté mohlo přijímat moderní poznatky. Proto bylo třeba uspořádat korejštinu, která by se mohla vyučovat v reformovaném školském systému. Obrozenci používali vlastní korejské písmo *hangŭl* namísto klasické čínštiny, neboť její osvojení je podstatně rychlejším procesem, a snažili se skrze slovníky lidem přiblížit původní korejské výrazy.⁴⁹ Organizace jako Korejská lingvistická společnost (조선어 학회 [čosönŏ hakhö], dál jako KLS) překládaly knihy z *hanmunu*, vytvářely učebnice, gramatiky a slovníky korejštiny. Noviny a další publikace též vycházely v korejském písmu. Známou ukázkou z tohoto období by byly Noviny nezávislosti (독립신문 [tongnip sinmun]), které vycházely v letech 1896-99 v korejštině a angličtině.⁵⁰

Tyto snahy pokračují v období japonské okupace od roku 1910 do konce druhé světové války. Korejšтина postupně ustupovala japonštině ve veřejném i soukromém prostoru. Později byla japonština stanovena úředním jazykem a vytlačila korejštinu i ze školních osnov.⁵¹ Reakcí na tuto situaci byla opět obrozenecká činnost. Díky korejským vlastencům a jazykovědcům podporovaným Japonskem byla v roce 1912 zavedena Pravidla korejského pravopisu pro obecné školy (보통학교용 언문 철자법 [potchonghakkjojong ŏnmun čchŏlččapŏp]), ve kterých bylo uvedeno, že norma vychází ze sŏulské podoby korejštiny užívané střední vrstvou obyvatelstva, jež byla už dříve obecně přijímána jako standardní forma jazyka.⁵² Na tomtéž základě stojí i dnešní standardní jazyk v Korejské republice – 표준어 [pčhjodžunŏ]. Z čistě korejské iniciativy KLS poté vyšel Návrh jednotného korejského pravopisu (한글 맞춤법 통일안 [hangŭl maččumpŏp tchongiran]) v roce 1933, který sloužil jako vzor pro pozdější jazykové reformy v obou rozdělených státech.

⁴⁸ ČERMÁK, František. *Jazyk a jazykověda*. 2011. s.83-84

⁴⁹ PUCEK, Vladimír a Marta BUŠKOVÁ. *Jazyková politika v Koreji*. Praha: Karolinum, 2000, s.66

⁵⁰ Ibid. s.78

⁵¹ LEE, Sung-nyong. *On the Need of Teaching Chinese Characters*. s.49

⁵² PUCEK, Vladimír a Marta BUŠKOVÁ. *Jazyková politika v Koreji*. Praha: Karolinum, 2000. s.90

Skutečně o purifikaci lze hovořit až po odchodu Japonska z Korejského poloostrova. Severní část kontroloval SSSR a polovina ležící na jih od 38. rovnoběžky padla pod správu USA. V roce 1948 tak vznikly dva samostatné korejské státy potýkající se s velmi vysokým procentem negramotnosti v mateřském jazyce, kterou zapříčinila japonská jazyková politika v uplynulých 40 letech. Řešení tohoto problému se obě administrativy zhostily nezávisle na sobě.

5.2 Korejská lidově demokratická republika

V KILDR žily po rozdělení odhadem 2 miliony negramotných Korejců.⁵³ Jazyk, jakožto důležitý nástroj pro šíření hodnot a politické ideologie, byl pro Komunistickou stranu Koreje důležitou prioritou. Co nejdříve se tedy snažila vzdělat lid a zajistit obecnou školní docházku. Byly založeny speciální školy, které museli navštěvovat všichni negramotní Severokorejci. Pro účely výuky bylo nutné korejštinu přizpůsobit – zjednodušit tak, aby se jí byl každý schopen naučit co možná nejrychleji. Nejprve byly eliminovány četné japanismy. Smíšený způsob zápisu obsahující čínské znaky byl z osnov vynechán a vyučovalo se pouze korejské písmo. Učebnice neobsahovaly žádné čínské znaky. Kim Il-söng ve svém projevu v roce 1946 navíc nabádal k socializaci jazyka. Argumentoval, že znalost znaků byla výsadou malé skupiny lidí, a odkazovala na dřívější feudální období, což bylo pro Severní Koreu v této fázi nežádoucí. V následujících letech bylo zakázáno používání znaků v novinách a literatuře. Výjimku tvořily pouze vysvětlivky v závorkách pro rozlišení mnoha homonym.⁵⁴ Později bylo do školních osnov opět zařazeno omezené množství znaků, v publikacích se však do současnosti nevyskytují.

Po eradikaci znaků zahájila strana v 60. letech hnutí za purifikaci jazyka (말 다듬기 운동 [mal tadümgi undong]) a Kim Il-söng v projevech vyzval lingvisty k vytvoření nového slovníku a analýze sinokorejské terminologie. Úkolem bylo zhodnotit, které výrazy budou ponechány a které by měly být nahrazeny odpovídající novou slovní zásobou. Kromě sinokorejských slov měly být omezeny i sinojaponské nebo čistě japonské výpůjčky a méně zdomácnělá slova přejatá z angličtiny a ruštiny. Význam čistě korejských slov se také posouval tak, aby sloužil ideologii. Slova s nežádoucím významem se nevyučovala a nezahrnovala do slovníků s nadějí, že přestanou být součástí aktivní komunikace.⁵⁵

Tento nový jazyk již neměl být označován jako standardní řeč *pchjodžunö*, která v sobě nesla negativní rysy kapitalismu a buržoazie, stejně jako množství anglicismů a japanismů, ale novým termínem kulturní řeč *munhwaö* [munhwaö]. Kulturní řeč nevychází z dialektu oblasti Söulu, ale z nářečí provincie Pchjöngan, tedy z okolí Pchjöngjangu, severního hlavního města a centra revoluce. Má obsahovat převážně původní či nově vytvořená korejská slova. Jak bylo naznačeno v předchozí kapitole, slovní zásoba tvořená ze sinokorejských morfémů sice má proti původně korejské výhodu v tom, že lidem znalým čínských znaků umožňuje snadno odvodit význam slova právě na základě významu jeho jednotlivých částí, ale Severokorejci, kteří výuku znaků neabsolvovali, si významy morfémů do této míry neuvědomovali.⁵⁶ Také se změnilo vědecké názvosloví, dětem bylo žádoucí dávat originální korejská jména namísto jmen složených ze sinokorejských morfémů.

⁵³ KUMATANI, Akiyasu. *Language Policies in North Korea*. 1990. s.90

⁵⁴ KIM, Chin-Wu. *Linguistics and language policies in North Korea*. 1978. s.166-7

⁵⁵ Ibid., s.168

⁵⁶ LEE, Sung-nyong. *On the Need of Teaching Chinese Characters*. 1972. s.50

Mnoho zdomácňelých slov v jazyce zůstalo (např. 강 (江, [kang]) – řeka, 산 (山 [san]) – hora, 학교 (學校, [hakkjo]) – škola). Ostatní byla buď nahrazena existujícím korejským synonymem, archaickým či nespisovným výrazem, který se stal normou, nebo se posunul význam slov tak, aby odpovídal obsahu sinokorejských. Slovní zásoba byla obohacena procesy typickými pro korejštinu – skládáním, odvozováním pomocí prefixů a sufixů, či změnou slovního druhu. Také přichází jeden z hlavních rysů kulturního jazyka, který spočívá v doslovném „překladu“ sinokorejského morfému korejským slovem, jenž má za následek dlouhá, často krkolomná sousloví).⁵⁷ Nová slovní zásoba byla představována ve speciálních sloupcích v tisku.

Pokud se početná sinokorejská vrstva slovní zásoby zapisuje pouze hangülem, přicházíme o rozlišující sémantickou složku znaků. Eliminace těchto slov tak svým způsobem poměrně efektivně řeší problém vysokého počtu homonym.

5.3 Korejská republika

Korejská republika podrobila jazyk také několika změnám – vymýcení japanismů, protlačování zápisu čistě hangülem a purifikační hnutí (언어 순화 [önö sunhwa])⁵⁸ v 70. letech. Důraz byl kladen nejen na slovní zásobu, ale také „krásnější jazyk“, správné užívání honorifik, výslovnost apod.

I když ani v Korejské republice dnes nejsou čínské znaky uplatňovány v běžném psaní, zůstávají v tisku (jak je patrné i v praktické části práce), odborné literatuře, školních osnovách i v každodenním životě na cedulích, v restauracích apod. Ve vědeckých publikacích se stejně jako v Severní Koreji znaky užívají především pro účely rozlišení homonym a snazší dohledatelnost termínů. Zpravidla jsou do textu v hangülu doplněny v závorkách za daným slovem. Pokud jde o standardní výuku, sada znaků určených ke studiu na školách se v obou zemích liší.

Přestože prostředky jazykové politiky obou států se v mnohém shodují, jedná se o dva vzájemně nezávislé a nekoordinované procesy a výsledná podoba jazyka se hlavně v oblasti slovní zásoby liší, neboť slova určená k odstranění byla vybírána na základě jiných kritérií. V poslední době rovněž dochází k novému přejímání cizích slov, kdy Korejská republika čerpá převážně z angličtiny a japonštiny, zatímco v případě KLDK přišlo mnoho slov spíše z ruštiny.

Pro každý živý jazyk je přirozené, že se vyvíjí, ať už plně nekontrolovaně, nebo za pomoci kodifikovaných úprav. Změna, kterou korejština prodělala v rozpětí pouhých cca 40 let, kdy byly úpravy nejintenzivnější, ale rozhodně není zcela běžnou praxí. Fakt, že jazyková divergence je natolik znatelná, že komplikuje jednání mezi státy, se poprvé výrazněji ukázal v roce 2001 při diplomatických setkáních, kdy vznikla nedorozumění a neshody ohledně textu smlouvy o demilitarizované zóně (DMZ, 비무장지대 [pimudžangdžidä]).⁵⁹ Severní i jižní podoba korejštiny jsou navzájem stále dostatečně srozumitelné,

⁵⁷ KUMATANI, Akiyasu. Language Policies in North Korea. 1990. s.97-99

⁵⁸ Všimněme si, že samotný název zažitý pro purifikační hnutí v KLDK sestává z domácích korejských slov – 말 다듬기 [mal dadŭmgi] – řeč a broušení/kultivace, zatímco v Korejské republice se označuje dvěma sinokorejskými výrazy 언어 순화 [önö sunhwa] – jazyk a očista/purifikace a ve slovnících nalezneme u doporučených slov termín 순화어 [sunhwaö] – čisté slovo.

⁵⁹ ČCHÖ, Hočchöl. Söro tarŭn öhwüi tongčilsöngül wihan pangan 서로 다른 어휘의 동질성을 위한 방안 (Návrh na sjednocení vzájemně odlišného lexika). 2001 s.64

abychom je mohli považovat za jeden jazyk, ale v rámci mezikorejských vztahů, nebo dokonce potenciálního sjednocení obou států, jsou existence dvou spisovných norem a možná lexikální nedorozumění palčivým problémem.

V reakci na diplomatická úskalí se dodnes lingvisté snaží přijít s adekvátním řešením problému. Vznikají korejsko-korejské slovníky či překladové aplikace a uprchlíci ze Severní Koreje, kteří svůj domov naleznou na Jihu, dostávají lekce jazyka v rámci vzdělávání, jenž jim má umožnit začlenění do jihokorejské společnosti a života.

6 Komparativní analýza lexika

Následující kapitola je věnována samotné analýze korpusu. U každého deníku se zaměřím nejprve na nejčastěji se vyskytujících 100 slov a zastoupení původních korejských slov, sinokorejských slov, slov cizího původu a jejich kombinací v kontrastu se statistikou z otevřeného slovníku Urimalsam a hypotézou vycházející z purifikačních snah v obou zemích. Jinými slovy, zda podíl domácí slovní zásoby výrazně převyšuje ostatní skupiny a zda poměr zastoupení skupin odpovídá celkovému stavu slovní zásoby podle online slovníku nebo se od něj odchyluje.

Pro srovnání, jak si počínaly oba státy při tvoření případných neologismů a ustálení lexika souvisejícího s koronavirovou krizí poslouží pravděpodobně lépe klíčová slova, kterým jsou věnovány podkapitoly 6.1.2 a 6.2.2. U vybraných slov následující podkapitoly 6.1.3 a 6.2.3 sledují konkordanci neboli okolí jejich výskytu a shluky, jež nejčastěji tvoří.

Data jsou znázorněna pomocí tabulek a grafů pro lepší přehlednost. Seznamy nevyčleňují plnovýznamová a gramatická slova a vlastní jména, ale třídí typy na základě jejich původu (korejského – K, sinokorejského – S a cizího – C). Slova odvozená morfémy jiného původu a kompozita morfémů nestejného původu jsou zde uváděna zvlášť (v tabulce značeno jako kombinace zkratk, např. S+K pro slovo složené z částí sinokorejského a korejského původu). Homonyma nejsou brána v potaz, pokud se jedná o slova stejného původu. Pokud jsou různého původu, je tak uvedeno v poznámce a počet tokenů každého významu byl zohledněn a započítán do poměrů jednotlivých skupin. V tabulce je z praktických důvodů uveden původ pouze toho slova, které se z homonym objevovalo nejčastěji.

Pokud se nachází v seznamu gramatické slovo (převážně partikule) jako samostatný výsledek, je počítáno jako plnohodnotný typ. Jestliže jde však o spojení plnovýznamového slova a partikule, tuto partikuli nezohledňuji a nepovažuji za součást slova a neuvádím jako morfém domácího původu.

6.1 Čosŏn ilbo

6.1.1 Slova s nejvyšší frekvencí

Články z jihokorejského deníku Čosŏn ilbo tvoří objemnější ze dvou korpusů s několikanásobně vyšším počtem tokenů, tudíž i frekvence u nejčastějších typů se pohybuje ve vyšších číslech než v případě novin z KLDK. Program AntConc pro soubor článků z roku 2020 vygeneroval tento seznam.

Tabulka 4: Nejfrekventovanější slova ve vybraných člancích Čosŏn ilbo

Pořadí	Frekvence	Slovo	Původ	Pořadí	Frekvence	Slovo	Původ
1	20363	일	S	51	2226	차	S
2	17724	있다	K	52	2152	가	K
3	17581	코로나	C	53	2108	정부	S
4	15334	고	K	54	2088	로	K
5	14179	했다	K	55	2063	정부가	S
6	8409	수	K	56	2057	현지	S
7	8086	등	S	57	2012	없다	K
8	7801	이	K	58	2001	대	S
9	7543	지난	K	59	1992	그는	K
10	7214	월	S	60	1986	조	S**

11	6651	한	K*	61	1942	정부는	S
12	5861	년	S	62	1921	등을	S
13	5668	이날	K	63	1900	마스크	C
14	5507	명	S	64	1900	시각	S
15	5445	며	K	65	1891	감염	S
16	4840	있는	K	66	1886	확진	S
17	4769	것으로	K	67	1858	확진자	S
18	4590	이라고	K	68	1857	올해	K
19	4538	것	K	69	1847	때문에	K
20	4032	말했다	K	70	1831	시	S
21	4027	것이다	K	71	1820	것은	K
22	3937	는	K	72	1811	관계자는	S
23	3844	전	S	73	1810	다른	K
24	3809	중	S	74	1778	의	K
25	3561	서울	K	75	1767	하루	K
26	3379	미국	S	76	1756	이번	S+K
27	3364	만	S	77	1755	또	K
28	3341	한다	K	78	1753	라고	K
29	3248	위해	S+K	79	1733	하는	K
30	3230	명이	S	80	1731	오후	S
31	3189	중국	S	81	1719	관련	S
32	2981	를	K	82	1717	국내	S
33	2904	이후	S	83	1710	이상	S
34	2845	방역	S	84	1709	가장	K
35	2829	밝혔다	K	85	1688	있다는	K
36	2826	대한	S+K	86	1661	통해	S+K
37	2776	더	K	87	1658	일부더	S+K
38	2698	을	K	88	1638	나온다	K
39	2682	미	S	89	1623	할	K
40	2591	트럼프	C	90	1612	두	K
41	2584	확진자가	S	91	1605	하고	K
42	2524	에	K	92	1603	때	K
43	2450	최근	S	93	1589	마스크를	C
44	2380	지난달	K	94	1577	가운데	K
45	2353	같은	K	95	1570	백신	C
46	2324	따르면	K	96	1567	약	S
47	2276	대해	S+K	97	1564	안	K
48	2255	하지만	K	98	1552	모두	K
49	2246	바이러스	C	99	1552	위한	S+K
50	2243	우한	C***	100	1544	거리	K

*한 [han] nejčastěji ve významu *jeden* nebo jako minulé participium od slovesa 하다 [hada], 390 tokenů jako sinokorejské 한 (韓 [han]) – *Korea, korejský*

**조 [čo] nejčastěji jako sinokorejská číselná jednotka (兆) *bilion* a příjmení Čo 조 (曹/趙 [čo]), 326x ve jméně prezidenta USA Joea Bidena (cizího původu)

***우한 [uhan] město Wuhan v ČLR. I když jeho název lze psát znaky (a tento zápis se objevil coby typ v seznamu na 4661. místě), jeho původ není jasně určen. Výslovnost 우한 [uhan] totiž neodpovídá správnému čtení znaků 무한 (武漢, [muhan]) a poukazuje na výpůjčku z čínštiny (s původní výslovností) či jiného jazyka.

Doplňný původ všech slov vychází z informací na webových stránkách Slovníku standardní korejštiny (표준국어대사전 [pchjondžungugö täsadžön])⁶⁰, tři slova (červeně zvýrazněná) se ve slovníku nenacházela a musela být vyhledána v Urimalsäm. Příjmení bývalého amerického prezidenta Donalda Trumpa (트럼프 [tchrömpchü]), které se v tabulce objevuje, ve standardním slovníku pochopitelně chybělo. Slovo 코로나 [kchorona] bylo uvedeno v několika významech, ale ne v souvislosti s koronavirem. Ve slovníku Urimalsäm je tato definice na čtvrtém místě po těch, které se shodují se Slovníkem standardní korejštiny.

„코로나 (corona)

코로나 「001」 「명사」 『음악』 악보에서, 음표나 쉼표의 위나 아래에 붙어서 본래 박자보다 두세 배 길게 늘어 연주하라는 기호...⁶¹

„코로나 (corona)

코로나 「002」 「명사」 『천문』 태양 대기(大氣)의 가장 바깥층에 있는 얇은 가스층. 온도는 100만 °C 정도로 매우 높다. 개기 일식 때에는 맨눈으로 볼 수 있으며, 보통 때에는 코로나그래프 따위로 관측할 수 있다.

코로나 「003」 「명사」 『화학』 두 전극 사이에 높은 전압을 가하면 불꽃을 내기 전에 전기장의 강한 부분만이 발광하여 전도성(傳導性)을 갖는 현상. 송전선 상호 간이나 송전선과 대지 사이에서 일어난다.

코로나 (←coronavirus)

코로나 「004」 「명사」 ‘코로나바이러스’ 를 일상적으로 이르는 말.⁶²

Slovo 확진자 [hwakččindža] překvapivě také není obsaženo ve Slovníku standardní korejštiny. Ve slovníku Urimalsäm je definováno jako „člověk, kterému byl s jistotou diagnostikován druh nebo stav nemoci“⁶³, tedy potvrzený případ.⁶⁴ I když slovo existovalo a bylo užíváno již dříve (příklady a-b), z příkladových vět je patrná jeho současná asociace s onemocněním COVID-19 (c-k).

- a. 00 시 보건 당국은 2 차 감염이 우려됐던 이질 **확진자** 가족 25 명의 가검물을 채취해 조사한 결과 감염자가 나타나지 않았다고 밝혔습니다. <<와이티엔 2004 년 5 월>>
- b. 충남 00 시는 저소득층 의료 급여 수급자의 암 조기 발견 및 암 **확진자** 치료비 지원을 위해 11~23 일 시 보건소와 보건 지소에서 무료 암 검진을 실시한다. <<연합뉴스 2005 년 3 월>>
- c. **신종 코로나바이러스 감염증 확진자**가 다녀간 영화관 00 점이 영화 상영을 중단했다. <<조선비즈 2020 년 1 월>>
- d. 전 세계 **확진자** 수는 지난 27 일 60 만 명에 도달한 지 불과 이틀 만에 10 만여 명이 증가하며 폭발적인 증가세를 보이고 있다. <<노컷뉴스 2020 년 3 월>>

⁶⁰ Pchjodžungugö täsadžön <https://stdict.korean.go.kr/>

⁶¹ Urimalsäm

https://opendic.korean.go.kr/search/searchResult?query=%EC%BD%94%EB%A1%9C%EB%82%98&dicType=1&wordMatch=Y&searchType=¤tPage=1&cateCode=&fieldCode=&spCode=&divSearch=search&infoType=confirm&rowsperPage=10&sort=W&side_data=0%7C1476185

⁶² ibid

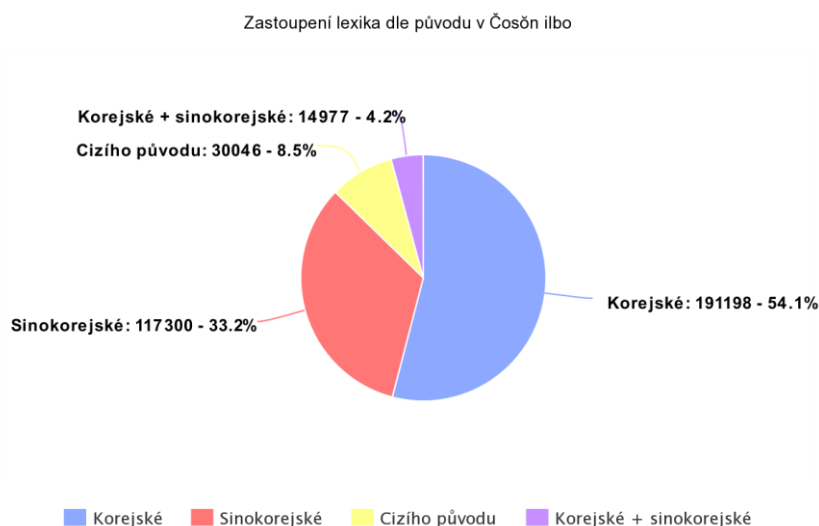
⁶³ 질환의 종류나 상태를 확실하게 진단받은 사람

⁶⁴ Urimalsäm https://opendic.korean.go.kr/dictionary/view?sense_no=649636&viewType=confirm

- e. 이는 중국에 이어 두 번째로 많은 규모로 이탈리아와 중동이 **확진자** 치료 대응에서도 실패하면서 사망자 발생 2·3위를 차지했다.《매일경제 2020년 3월》
- f. 프랑스와 독일, 스페인은 이날 **확진자** 수가 각각 9000~1만 5000명을 넘어 한국을 따돌렸다.《매일경제 2020년 3월》
- g. 서울 00구 00동 00빌딩 콜센터와 관련된 신종 **코로나바이러스 감염증 확진자**가 서울에서만 69명에 이르는 것으로 파악됐다.《연합뉴스 2020년 3월》
- h. 국내 **코로나 19 신규 확진자**가 사흘 연속 20명대를 유지했다.《매일경제 2020년 4월》
- i. 이탈리아를 제치고 미국에 이어 **확진자** 2위를 기록한 스페인도 점차 안정세를 보이고 있다.《매일경제 2020년 4월》
- j. **코로나 19 확진자** 가운데 의료진이 사망한 것은 이번이 처음이다.《매일경제 2020년 4월》
- k. **코로나 19 확진자**가 전 세계적으로 300만 명을 넘어섰다.《매일경제 2020년 4월》⁶⁵

Tento seznam frekventovaných slov nám poskytuje vzorek, na základě něhož můžeme odhadovat zastoupení domácích korejských, sinokorejských a cizích slov ve článcích. Zde je původ typů znázorněn na grafu včetně počtu tokenů a procent zastoupení jednotlivých skupin.

Graf 1: Zastoupení lexika dle původu ve frekventovaných výrazech ze článků Čosŏn ilbo



Stejně jako v běžném jazyce se zde nejpočetněji vyskytuje domácí slovní zásoba, která tvoří více než polovinu sta nejfrekventovanějších slov (54,1 %). V tabulce je zahrnut velký počet gramatických morfémů, závislých substantiv a pomocných sloves korejského původu, což k tomuto číslu přispívá.

Naopak sinokorejská slova jsou zpravidla plnovýznamová a, naproti mému předpokladu výše, v publicistickém stylu stejně (tedy ne nijak výrazně více) prominentní jako obecně v každodenní korejštině. Tvoří 33,2 % z 35 3521 tokenů analyzovaného výběru. Ve statistice slovníku Urimalsám, která je uvedena v kapitole 4.4, představovala sinokorejská slova naprosto shodně 33,2 % lexika.

O dost méně jsou zastoupena slova cizího původu, konkrétně 8,5 %. Slova vzniklá spojením sinokorejských a původně korejských slov nebo morfémů potom tvoří 4,2 %. Z většiny se jedná o sinokorejské plnovýznamové morfémy či jejich kompozita, které byly verbalizovány pomocným slovesem

⁶⁵ Urimalsám https://opendic.korean.go.kr/dictionary/view?sense_no=649636

하다 [hada]. Kombinace původního a cizího lexika, ani slova tvořená spojením morfémů všech tří původů zároveň se ve 100 typech vůbec nenacházela. Z tohoto vzorku lze usuzovat, že slova smíšeného původu jsou ve člancích méně častými skupinami než slova s „čistým“ původem.

6.1.2 Klíčová slova

Jak bylo vysvětleno výše, klíčová slova by měla odrážet povahu a obsah textu lépe než pouhý seznam slov seřazený podle počtu tokenů. Předpokládala jsem, že postavením korpusu Čosön ilbo proti Rodong sinmun se mi podaří částečně eliminovat irelevantní lexikum. Pokud nahlédneme do tabulky níže, uvidíme, že seznam slov s vysokou frekvencí a klíčová slova se v mnoha případech shodují. To může být způsobeno tím, že se v těchto případech dva korpusy nedostatečně překrývaly.

Tabulka 5: Klíčová slova vybraných článků Čosön ilbo

Klíčovost	Frekvence	Slovo	Původ	Klíčovost	Frekvence	Slovo	Původ
6669.49	17581	코로나	C	494.28	1438	대통령이	S
5238.69	15334	고	K	487.92	2255	하지만	K
4476.48	14179	했다	K	484.81	1731	오후	S
4409.68	17724	있다	K	477.41	1413	문	S
1948.47	5445	며	K	466.97	1236	문재인	S
1645.24	4769	것으로	K	450.36	1857	올해	K
1604.98	5861	년	S	448.82	2088	로	K
1570.95	4590	이라고	K	440.81	2063	정부가	S
1570.87	20363	일	S	437.36	2226	차	S
1400.89	8409	수	K	427.14	1296	경기	S
1387.01	4032	말했다	K	422.71	1202	때문이다	K
1170.37	3809	중	S	419.36	1110	대선	S
1153.91	3844	전	S	418.63	1213	최대	S
1098.13	4538	것	K	413.31	1094	코로나로	C
1082.49	3379	미국	S	409.56	1710	이상	S
1001.8	2845	방역	S	407.39	2524	에	K
979.24	2591	트럼프	C	404.52	1767	하루	K
949.51	2682	미	S	396.63	6651	한	K*
948.38	4027	것이다	K	391.77	1037	온라인	C
901.2	8086	등	S	390.76	1215	여	S
848.78	2246	바이러스	C	387.37	7214	월	S
847.64	2243	우한	C*	387.24	1025	사태	S
845.94	3189	중국	S	384.84	1753	라고	K
842.9	2324	따르면	K	379.91	1129	오전	S
838.08	3561	서울	K	379.67	1108	전국	S
828.44	2380	지난달	K	377.79	1000	신규	S
790.14	2904	이후	S	375.9	995	만원	S+K
777.87	3937	는	K	372.77	1233	만에	K
763.57	2057	현지	S	372.16	1108	알려졌다	K
693.61	2584	확진자가	S	369.43	1036	됐다	K
688.3	1886	확진	S	364.53	1143	교수는	S
670.89	2108	정부	S	363.59	1104	열린	K
666.13	1900	시각	S	360.17	1331	주	S
650.48	1858	확진자	S	359.27	951	연합뉴스	C
639.37	1986	조	S*	358.35	1224	뒤	K

636.15	1719	관련	S	356.47	976	사태로	S
627.02	2981	를	K	355.12	940	두기	K
624.88	1811	관계자는	S	345.12	971	확산	S
595.12	1638	나온다	K	344.05	1069	김	S
593.21	1570	백신	C	343.37	1847	때문에	K
571.61	1900	마스크	C	340.93	1342	세	S**
550.12	1456	이라며	K	340.76	902	보도했다	S+K
549.74	1455	민주당	S	339.62	899	북한	S
541.45	1717	국내	S	333.71	1526	대통령은	S
533.68	1447	판정을	S	333.58	883	자가	S
533.63	1544	거리	K	323.75	857	청와대	S
524.73	2698	을	K	323.03	887	지원	S
522.91	5668	이날	K	321.49	851	수도권	S
518	1371	억원	S+K	318.26	1577	가운데	K
511.85	2001	대	S	314.77	865	추가	S

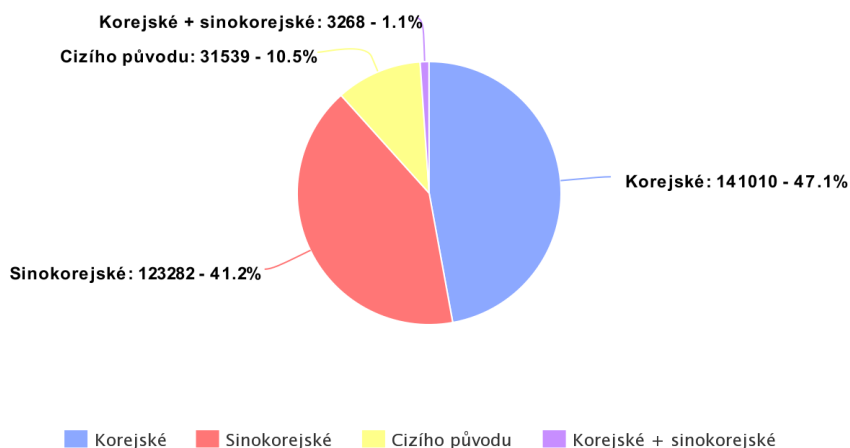
*Problematika typů 한 [han], 조 [čo] a 우한 [uhan] viz. výše Tabulka 4

**세 [se] nejčastěji jako počítací jednotka pro věk (歲), ale 448x jako korejská číslovka tři (셋 [set])

Rozdíl, byť malý, je vidět v poměru korejských, sinokorejských a cizích slov. U klíčových slov je zastoupení korejských a sinokorejských slov téměř vyrovnané. Korejské lexikum nyní představuje pouhých 47,1 % a sinokorejské 41,2 %. Slova cizího původu je také o 2 % více. Kombinací korejských a sinokorejských morfémů je zanedbatelné množství a ostatní varianty nejsou zastoupeny vůbec.

Graf 2: Zastoupení lexika dle původu ve frekventovaných výrazech ze článků Čosön ilbo

Klíčová slova dle původu v Čosön ilbo



Na prvním místě je slovo 코로나 [kchorona], které sloužilo jako vyhledávací kritérium, je tudíž jisté, že je v každém článku obsaženo alespoň jednou. Frekvence uvedena u typu dokazuje mnohem bohatší zastoupení. Následující konverba 고 [ko], 며 [mjö] a slovesa 있다 [itta], 하다 [hada] (respektive 했다 [hätta]) jsou běžnou součástí psaného korejského projevu a pro účely práce nejsou relevantní. To samé platí pro množství závislých substantiv, partikulí a koncovek. Přesto je v seznamu několik výsledků, které přináší při klusterové analýze zajímavé shluky.

6.1.3 Shluky vybraných slov z Čosŏn ilbo

코로나

Slovo **코로나** [kchorona] tvořilo nejčastěji tyto shluky se sinokorejskými výrazy: **코로나 사태** [kchorona satchä] – vypuknutí korony, **코로나 확산** [kchorona hwaksan] – šíření korony, **코로나 확진** [kchorona hwakččin] – potvrzený případ (diagnóza) korony, **코로나 확진자** [kchorona hwakččindža] – člověk s potvrzenou nákazou koronavirem, **코로나 확진 판정** [kchorona hwakččin pchandžong] – potvrzená diagnóza korony, **코로나 감염** [kchorona kamjöm] – koronavirová infekce, **코로나 방역** [kchorona pangjök] – koronavirová prevence, **코로나 환자** [kchorona hwandža] – pacient s korunou, **코로나 검사** [kchorona komsa] – test na koronavirus.

Mimo ně najdeme i spojení s dalším slovem cizího původu (z angličtiny) nebo shluk o více slovech různého původu jako: **코로나 바이러스** [kchorona pairösü] – koronavirus, **코로나 백신** [kchorona päksin] – očkování proti koruně, **코로나 바이러스 감염증** [kchorona pairösü kamjömdžüng] – koronavirová infekce, **코로나 팬데믹** [kchorona pchändemik] – pandemie koronaviru.

Některé věty ukazují tendenci vysvětlovat nebo asociovat cizí slova se sinokorejským ekvivalentem, např.:

„전 세계 코로나 바이러스 팬데믹(세계적 대유행)이 진정될 기미를 보이지 않는 데다...“

[ččön segje kchorona pairösü pchändemik (segječök täjuhäng)i čindžöngdöl kimirül poidži annün teda...]

„Celosvětová koronavirová pandemie (světová pandemie) nevykazuje známky uklidnění a...“

방역

Sinokorejské slovo **방역** [pangjök] prevence šíření nemoci se pojí s dalšími sinokorejskými slovy: **방역 당국** [pangjök tangguk] – orgány pro prevenci (šíření) nemocí, **방역 수칙** [pangjök sučchik] – nařízení pro prevenci nemocí, **방역 조치** [pangjök čočchi] – kroky pro prevenci, **방역 지침** [pangjök čičchim] – pokyny pro prevenci nemocí, **방역 강화** [pangjök kanghwa] – posílení prevence nemocí, **방역 대책** [pangjök täčchäk] – preventivní opatření proti šíření nemocí, **방역 실패** [pangjök silpchä] – neúspěšná prevence, **방역 체계** [pangjök čhegje] – systém prevence šíření nemocí.

확진, 확진자

Potvrzený případ **확진** [hwakččin] najdeme v sinokorejských spojeních: **확진 판정** [hwakččin pchandžöng] – potvrzená diagnóza, **확진 판정을 받다** [hwakččin pchandžöngül patta] – obdržet jistou (potvrzenou) diagnózu⁶⁶, **확진 환자** [hwakččin hwandža] – potvrzený pacient, **확진 사례** [hwakččin sarje] – případ potvrzené nemoci.

⁶⁶ Sloveso 받다 [patta] – dostat, obdržet je korejského původu.

Již výše vypíchnutý výraz **확진자** [hwakččindža] člověk s potvrzenou nemocí, který vznikl odvozením od **확진** [hwakččin] pomocí sufixu pro názvy osob **자** [ča] byl většinou rozvíjen přívlasky označujícími virus: **코로나 확진자** [kchorona hwakččindža] – potvrzený případ koronaviru, **우한 코로나 확진자** [uhan kchorona hwakččindža] – potvrzený případ koronaviru z Wuhanu.

Ale také při sčítání případů: **신규 확진자** [singju hwakččindža] – nový potvrzený případ, **하루 확진자** [haru hwakččindža] – denní potvrzené případy⁶⁷, **첫 확진자** [čchöt hwakččindža] – první potvrzený případ, **국내 확진자** [kungnä hwakččindža] – domácí (v zemi) potvrzený případ, **코로나 신규 확진자** [kchorona singju hwakččindža] – nový potvrzený případ koronaviru, **하루 신규 확진자** [haru singju hwakččindža] – nové potvrzené případy za den.

바이러스

Anglicismus **바이러스** [pairösü] se kromě spojení se samotným slovem **코로나** [kchorona] vyskytuje také v delších označeních viru: **우한 코로나 바이러스** [uhan kchorona pairösü] – koronavirus z Wuhanu,⁶⁸ **신종 코로나 바이러스 감염증** [sindžong kchorona pairösü kamjömdžung] – infekce koronaviru nového typu.

우한

Především v počátku pandemie byl často COVID-19 označován podle čínského města Wuhan, odkud pochází. Otázku čtení znaků a zda se jedná sinokorejské slovo či o výpůjčku, viz. Tabulka 4. Pro účely analýzy slovo považuji za přejeté z čínštiny, ne sinokorejské. Kromě jiných, zmíněných výše, má koronavirus i tyto názvy: **우한 코로나** [uhan kchorona] – korona z Wuhanu, **우한 코로나 바이러스 감염증** [uhan kchorona pairösü kamjömdžung] – infekce koronavirem z Wuhanu, **우한 폐렴** [uhan pchjerjöm] – wuhanský zápal plic, **우한 코로나 확산** [uhan kchorona hwaksan] – šíření korony z Wuhanu.

Wuhanský zápal plic **우한 폐렴** [uhan pchjerjöm] se vyskytuje výhradně v lednových článcích a nezřídka provázen některým z jiných názvů viru či onemocnění. Viz. např.:

37.5 도 이상 발열이나 기침 등 증상이 없는 상태에서 **우한 폐렴(신종 코로나 바이러스)**에 전염된 사례가 세계 곳곳에서 잇따르고 있다.

[samsipčhilčomodo isang parjörina kičchim tūng čüngsangi omnün sangtchäesö **uhan pchjerjöm (sindžong kchorona pairösü)**je čönjömdön sarjega segje kotkosesö ittarūgo itta]

Ve světě se objevují případy nákazy **wuhanským zápallem plic (koronavirem nového typu)** bez příznaků jako jsou teplota vyšší než 35,7 °C nebo kašel.

Objevoval se i zápis doplněn o znaky:

신종 코로나는 중국 우한(武漢)에서 2019년 12월 초에 발생한 것으로 알려졌다.
[sindžong kchoronanün čungguk uhan (muhan)esö ičchönsipkunjon sibiwöl čhoe palsänghan kösūro alljödžjötta]

⁶⁷ Také ve spojení se **수** [su] ve významu „počet případů za den“

⁶⁸ Zajímavostí je, že všechna tři slova zde jsou cizího původu.

Korona nového typu je známá tím, že vypukla v Číně ve Wuhanu na začátku prosince 2019.

마스크

Další slovo přejaté z angličtiny – rouška 마스크 [masŭkchŭ] je podle příkladových vět ze slovníku Urimalsŏm rovněž v současnosti úzce spojováno s pandemií COVID-19.

- a) 그는 감기를 옮기지 않으려고 마스크를 썼다.
- b) 신종 코로나 바이러스가 유행하며 마스크가 많이 팔렸다.
- c) 약국과 우체국 등 공적 마스크를 판매하는 판매처 위치와 판매 수량 등을 실시간으로 알려 주는 서비스가 시작됐다.《스포츠조선 2020년 3월》⁶⁹

Ve člancích Čosŏn ilbo se vyskytovalo například v těchto slovních spojeních se slovy korejského i sinokorejského původu: 마스크 착용 [masŭkchŭ čhagjong] – nošení roušky, 마스크 쓰기 [masŭkchŭ ssŭgi] – nošení roušky⁷⁰, 마스크 사용 [masŭkchŭ sajong] – používání roušky, 마스크 대란 [masŭkchŭ tŕaran] – roušková krize, 마스크 부족 [masŭkchŭ pudžok] – nedostatek roušek, 마스크 공급 [masŭkchŭ konggŭp] – zásobení rouškami, dodávka roušek, 마스크 등 [masŭkchŭ tŭng] – roušky atp., 마스크 없이 [masŭkchŭ ősŭ] – bez roušky, 마스크 품귀 [masŭkchŭ pchumgwŭ] – nedostatek roušek (nouze), 마스크 착용 의무화 [masŭkchŭ čhagjong ŭimuhwa] – zavedení povinnosti nosit roušku, 마스크 수급 [masŭkchŭ sugŭp] – nabídka a poptávka po rouškách, 마스크 생산 [masŭkchŭ sängsan] – produkce roušek

감염

Infekce, sinokorejským slovem 감염 [kamjŏm], tvořila také několik opakujících se spojení: 감염 경로 [kamjŏm kjŏngno] – trasa nákazy, 감염 사례 [kamjŏm sarje] – případ nákazy, 감염 확산 [kamjŏm hwaksan] – šíření infekce, 감염 여부 [kamjŏm jŏbu] – stav infekce, 감염 위험 [kamjŏm ŭhŏm] – riziko nákazy, 감염 예방 [kamjŏm jebang] – prevence nákazy, 감염 가능성 [kamjŏm kanŭngsŏng] – možnost nákazy, 감염 의심 [kamjŏm ŭisim] – podezření infekce.

Mimo ně se slovo vyskytuje v odvozeném 감염증 [kamjŏmdžŭng] (viz. distribuce výše ve spojení s dalšími klíčovými slovy) a 감염되다 [kamjŏmdŏda] být infikován a 감염자 [kamjŏmdža] nakažená osoba. Tato dvě slova nejsou ve člancích tak častá a ve frekvenci je jasně vidět preference slov příbuzných s 학진 [hwakččŭin] potvrzený případ.

백신

Jak postupoval rok 2020, stále častěji se objevovalo slovo vakcína 백신 [pŕksin]. Nejprve v souvislosti s vývojem, poté distribucí. Opět nejde o žádný neologismus spojený pouze s koronavirem, ale v současnosti je s ním slovo asociováno, jak demonstrují příkladové věty z Urimalsŏm:

⁶⁹ Urimalsŏm https://opendic.korean.go.kr/dictionary/view?sense_no=39931

⁷⁰ Dvě slovesa pro nošení jsou v tomto případě synonyma.

- a) 각종 전염병에 대한 예방 접종은 병원체나 독소를 약화시킨 **백신**을 인체에 투여해 면역을 유지하도록 하는 조치로 접종으로 인한 발열, 두통, 오한 등의 부작용은 피할 수 없는 실정이다.《연합뉴스 1991년 2월》
- b) 신종 코로나바이러스 감염증의 팬데믹을 멈출 수 있는 **백신**이 이르면 올해 안에 나올 수 있다는 관측이다.《헤럴드경제 2020년 4월》
- c) 한국뿐 아니라 세계 각국은 예산 지원 외에도 임상 평가를 빠르게 진행할 수 있는 체도를 마련하는 등 치료제와 **백신** 개발에 박차를 가하고 있다.《뉴스 1 2020년 6월》⁷¹

Spojení s cizím slovem **백신** [päksin]: 백신 개발 [päksin käbal] – vývoj vakcíny, 백신 접종 [päksin čöpčong] – vakcinace, 백신 후보 [päksin hubo] – adept na vakcínu⁷², 백신 확보 [päksin hwakppo] – zajištění vakcín, 백신 도입 [päksin toip] – představení vakcíny, 백신 공급 [päksinkonggüp] – dodávka vakcíny, zásobování vakcínou, 백신 무료 [päksin murjo] – vakcína zdarma, 백신 생산 [päksin sängsan] – výroba vakcíny.

사회적 거리 두기

Za použití sinokorejského a korejských slov 사회적 [sahödžök] společenský, 거리 [kõri] prostor, vzdálenost a 두기 [tugi] nechat vzniklo sousloví 사회적 거리 두기 [sahödžök kõri tugi] – udržování rozestupů, společenská izolace (social distancing), které se vyskytovalo v těchto spojeních: 사회적 거리 두기 수준 [sahödžök kõri tugi sudžun] – stupeň odstupů, izolace, 거리 두기 조치 [sahödžök kõri tugi čočchi] – kroky pro udržování odstupů, 거리 두기 3 단계 [kõri tugi sam tangje] – třetí stupeň odstupů, izolace.

자가 격리

Druhým souslovím, jež se srazilo ve spojení s koronavirovou pandemií, je 자가 격리 [čaga kjõngni] – izolace, vlastní karanténa (self-quarantine), tedy karanténa, kterou člověk absolvuje sám doma mimo vládní a jiné instituce. Také ve spojeních: 자가 격리 조치 [čaga kjõngni čočchi] – opatření pro izolaci, 자가 격리 기간 [čaga kjõngni kigan] – trvání karantény, 자가 격리 앱 [čaga kjõngni äp] – aplikace k izolaci, 자가 격리 위반 [čaga kjõngni üban] – porušení karantény, 자가 격리 의무 [čaga kjõngni üimu] – povinnost izolovat se, 자가 격리 중(이다) [čaga kjõngni čung(ida)] – (být) v karanténě, 자가 격리 통보 [čaga kjõngni čočchi] – oznámení o izolaci.

Od slova 자가 stojí za zmínku rovněž sousloví 자가 진단 – sebediagnóza.

⁷¹ *Urimalsäm* https://opendic.korean.go.kr/dictionary/view?sense_no=10372

⁷² Mezi konkurujícími firmami, které vyvíjely vakcinační látky souběžně

Vidíme tedy, že ve slovních spojeních dochází ke kombinaci slov nezávisle na jejich původu a zastoupena jsou původní korejská, sinokorejská i cizí slova. Nezdá se, že by byla přítomna výrazná snaha vyhýbat se výpůjčkám či sinokorejským slovům a za každou cenu upřednostňovat korejské lexikum.

6.2 Rodong sinmun

6.2.1 Slova s nejvyšší frekvencí

Ze druhého korpusu tvořeného texty tisku Rodong sinmun program AntConc opět vygeneroval jak slova s největší frekvencí výskytu, tak klíčová slova, pro což byl jako reference užit korpus Čoson ilbo. I zde jsou výsledky následně tříděny podle původu slov.

Tabulka 6: Nejfrekventovanější slova ve vybraných článcích Rodong sinmun

Pořadí	Frekvence	Slovo	Původ	Pořadí	Frekvence	Slovo	Původ
1	13766	만	S	51	682	우리	K
2	6736	명이	S	52	671	없었다	K
3	6341	명	S	53	666	하면서	K
4	4854	명의	S	54	642	및	K
5	3874	위한	S+K	55	630	비해	S+K
6	2597	한다	K	56	621	한편	S+K
7	2423	명에	S	57	601	났다	K
8	2244	이	K	58	599	신형코로나비루스감염으로	S+C
9	2201	사망자가	S	59	580	그는	K
10	2174	사망자수는	S	60	576	또한	K
11	2158	감염자와	S	61	575	등록되었다	S+K
12	2060	명으로	S	62	575	사망자	S
13	1838	대한	S+K	63	563	신형코로나비루스감염자수가	S+C
14	1752	감염되고	S+K	64	560	감염자는	S
15	1730	일	S	65	557	더	K
16	1676	그에	K	66	547	있는	K
17	1615	그	K	67	546	등	S
18	1588	대처하기	S+K	68	539	하고있다	K
19	1552	신형코로나비루스감염증피해	S+C+K	69	524	많은	K
20	1532	사망하였다	S+K	70	522	신형코로나비루스감염증을	S+C
21	1515	감염자수는	S	71	521	비롯한	K
22	1479	확대되고있는	S+K	72	515	전날에	S+K
23	1415	하루동안에	K	73	511	나라에서	K
24	1269	모든	K	74	508	전국적인	S
25	1222	철저히	S+K	75	507	감염자로	S
26	1162	감염자수가	S	76	504	세계적으로	S
27	1144	새로	K	77	501	나라	K
28	1122	감염자가	S	78	495	사업을	S
29	1084	각각	S	79	475	관련하여	S+K
30	1076	이르렀다	K	80	470	전국적으로	S
31	1049	늘어났다	K	81	459	이것은	K
32	1029	달하였다	S+K	82	458	나라의	K
33	997	명이다	S	83	457	위해	S+K

34	979	지난	K	84	455	시현재	S
35	964	일현재	S	85	448	늘어나	K
36	960	막기	K	86	443	이날	K
37	951	계속	S	87	443	전파를	S
38	941	여명	S	88	436	한다고	K
39	920	밝혔다	K	89	432	감염자	S
40	909	증가하였다	S+K	90	428	함께	K
41	890	사망자는	S	91	426	여명이	S
42	866	그중	S+K	92	424	된다	K
43	818	신형코로나비루스에	S+C	93	414	맞게	K
44	812	더욱	K	94	402	발생하여	S+K
45	809	발생하였다	S+K	95	401	약	S
46	754	수는	K	96	393	여명의	S
47	737	목숨을	K	97	388	대해	S+K
48	735	월	S	98	383	전염성폐염환자로	S
49	728	의하면	S+K	99	382	확진되었다	S+K
50	712	인한	S+K	100	377	현재	S

Prvním rozdílem, kterého si lze všimnout v porovnání s tabulkami 4 a 5, je délka některých nalezených typů, jež by bylo možné rozdělit a stále dostat smysluplná nezávislá slova (viz. níže). Tento jev je způsoben jinou pravopisnou normou v Korejské republice a Korejské lidově demokratické republice. V souladu s munhwaŏ lze psát dohromady i celé fráze.

Z významu několika prvních frekventovaných slov je lépe než u Čoson ilbo patrné nejčastější téma článků. Zdaleka nejvíce poskytoval Rodong sinmun reporty počtu případů lidí nakažených koronavirem a lidí, co na jeho následky zemřeli. To se odráží v četnosti slov 만 [man] 10000, 명 [mjöng], 명이 [mjöngi] – osoba (numerativ), 명의 [mjöngüi] – osob, 여명 [jömjöng] – přibližně osob, 사망자/사망자수 [samangdža/samangdžasu] – zemřelý/počet zemřelých, 감염자/감염자수 [kamjödža/kamjödžasu] – nakažený/počet nakažených, 확진자 [hwakččindža] – potvrzený případ⁷³, tuto domněnku potvrzuje následující část věnovaná shlukům obsahujícím vybraná slova. Téměř všechna okolí jejich výskytu jsou podobná jako to, jež můžeme vidět zde ve zvolené větě z červencového článku.

„이밖에 에스빠냐와 도이쉴란드에서 1 0 0 0 명이 훨씬 넘는 감염자들이 등록되었다.“

[ipakke esüppanja wa toičhüllandüesö čchön mjöngi hwölssin nömnün kamjödžadüri tünngnokdöjötta]

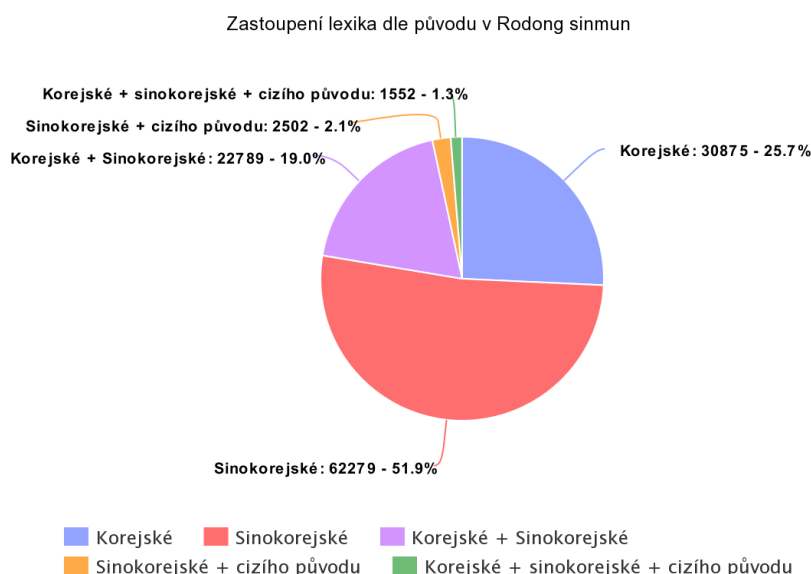
„Mimo to Španělsko a Německo zaregistrovalo zdaleka přes 1000 nakažených osob.“

I v hlášení o šíření nákazy ve světě se ukazují drobné rozdíly mezi Jižní a Severní Korejí. Zatímco zde vidíme 에스빠냐 [esüppanja] – Španělsko (ze španělského España) a 도이쉴란드 [toičhüllandü] – Německo (z. něm. Deutschland), tedy názvy států podle jejich oficiálních jmen v původním jazyce,

⁷³ I když není tak časté, vyskytuje se toto slovo i v Rodong sinmun. V seznamu slov podle frekvence je 416. To vzhledem k malému korpusu znamená, že se ve člancích naskytlo přesně 90x.

v Korejské republice bychom se dle Slovníku standardní korejštiny setkali se 스페인 [süpchein] zřejmě z anglického Spain⁷⁴ a sinokorejské 독일 [togil] – Německo.

Graf 3: Zastoupení lexika dle původu ve frekventovaných výrazech ze článků Rodong sinmun



Zdaleka nejvíce zastoupená jsou sinokorejská slova, což je překvapivé, vzhledem k cílům vytyčeným během purifikačního hnutí a snahám o upřednostňování původně korejského lexika namísto sinokorejského. Zatímco v Jižní Koreji tvořilo více než polovinu původní lexikum, zde se ukázalo být téměř 52% zastoupení sinokorejských slov. Až na druhém místě je domácí slovní zásoba s pouze čtvrtinovým podílem (25,7 %), poté kombinace korejských a sinokorejských slov s 19 %. Zvláštní se také může zdát 0% zastoupení slov čistě cizího původu, jelikož kombinace původních a cizích morfémů a všech tří se zde vyskytuje, ač v jednotkách procent. Je otázka, nakolik byly výsledky zkresleny několika slovními spojeními bez mezer, které byly tím pádem počítány jako jeden typ smíšeného původu.

6.2.2 Klíčová slova

Tabulka 7: Klíčová slova vybraných článků Rodong sinmun

Klíčovost	Frekvence	Slovo	Původ	Klíčovost	Frekvence	Slovo	Původ
36772.28	13766	만	S	1715.48	507	감염자로	S
14170.61	6736	명이	S	1690.65	560	감염자는	S
13926.24	4854	명의	S	1684.94	1269	모든	K
9574.48	6341	명	S	1648.1	511	나라에서	K
8734.33	3874	위한	S+K	1643.7	666	하면서	K
8268.5	2174	사망자수 는	S	1641.96	508	전국적인	S
7951.08	2158	감염자와	S	1623.33	1838	대한	S+K
6556.04	2423	명에	S	1602.54	459	이것은	K
6477.22	2201	사망자가	S	1585.89	951	계속	S
6440.29	1752	감염되고	S+K	1502.54	402	발생하여	S+K

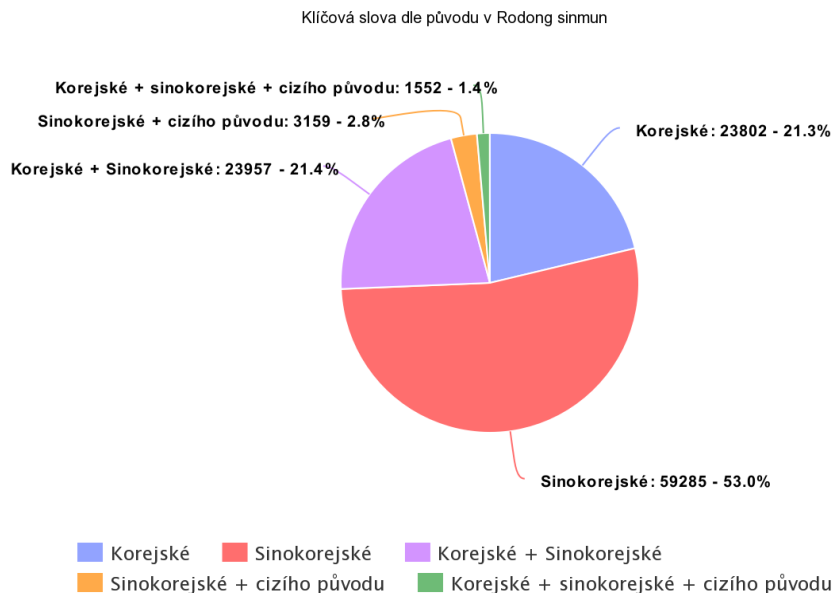
⁷⁴ Slovník však neuvádí původ z angličtiny, pouze původní slovo „Spain“ určující, že jde o 외래어 [öraö] – slovo cizího původu

5900.32	1552	신형코로나비루스 감염증파 해	S+C+K	1472.2	576	또한	S+K
5890.49	1588	대처하기	S+K	1454.92	383	전염성폐 염환자로	S
5824.21	1532	사망하였 다	S+K	1451.12	382	확진되었 다	S+K
5819.67	1676	그에	K	1409.32	371	받고있다	K
5759.51	1515	감염자수 는	S	1379.33	504	세계적으 로	S
5622.51	1479	확대되고 있는	S+K	1375.12	362	하루사이 에	K
5378.98	1415	하루동안 에	K	1371.33	361	강조하였 다	S+K
4416.47	1162	감염자수 가	S	1367.53	360	일군들과	K
3910.62	1029	달하였다	S+K	1364.5	458	나라의	K
3879.62	1222	철저히	S+K	1338.61	443	전파를	S
3873.42	1076	이르렀다	K	1329.53	350	신형코로나비루스 감염증의	S+C
3867.17	2060	명으로	S	1319.27	754	수는	K
3663.43	964	일현재	S	1311.82	365	하였다	K
3454.29	909	증가하였 다	S+K	1296.21	448	늘어나	K
3200.67	1144	새로	K	1277.76	630	비해	S+K
3108.29	818	신형코로나비루스 에	S+C	1275.08	359	인민의	S
3074.07	809	발생하였 다	S+K	1222.42	521	비롯한	K
2884.19	1049	늘어났다	K	1214.27	470	전국적으 로	S
2817.27	1122	감염자가	S	1188.83	324	나라에서 는	K
2806.9	941	여명	S	1186.84	501	나라	K
2791.54	2597	한다	K	1180.52	347	엄격히	
2663.48	728	의하면	S+K	1174.51	414	맞게	K
2647.14	866	그중	S+K	1173.75	309	말하였다	K
2536.17	997	명이다	S	1166.15	307	신형코로나비루스 감염자가	S+C
2342.43	737	목숨을	K	1161.83	393	여명의	S
2300.18	1084	각각	S	1109.17	712	인한	S+K
2275.78	599	신형코로나비루스 감염으로	S+C	1087.7	495	사업을	S
2217.54	671	잃었다	K	1071.17	282	일군들은	K
2184.56	575	등록되었 다	S+K	1063.57	280	감염되어	S+K
2172.49	1615	그	K	1050.67	621	한편	S+K

2172.37	890	사망자는	S	1025.36	575	사망자	S
2138.95	563	신형코로나 나비루스 감염자수 가	S+C	1020.29	272	힘있게	K
2113	960	막기	K	1017.98	268	발표하였 다	S+K
1983.64	539	하고있다	K	1005.12	268	방역사업 을	S
1983.13	522	신형코로나 나비루스 감염증을	S+C	1003.89	426	여명이	S
1865.42	812	더욱	K	992.25	276	확진되고	S+K
1832.45	601	났다	K	979.99	258	실속있게	S+K
1818.13	515	전날에	S+K	977.58	268	아시아에 서	C
1779.27	475	관련하여	S+K	962.51	264	하여	K
1728.51	455	시현재	S	961	253	책임적으 로	S

V případě klíčových slov z korpusu z Rodong sinmun se stejně jako u jihokorejského ekvivalentu změnilo od seznamu slov spíše méně než více výsledků. Stále spolehlivě na prvních místech vyšly numerativy a číslo. Tuto skutečnost adresuji níže v rámci zkoumání n-gramů. Nedostatečný filtr pro výběr klíčových slov ze seznamu slov s vysokou frekvencí poukazuje na nepříliš šťastnou volbu referenčního korpusu.

Graf 4: Zastoupení lexika dle původu ve klíčových slovech ze článků Rodong sinmun



Podíl sinokorejských slov je u klíčových slov ještě o procento vyšší než v prvním seznamu slov. Původních korejských slov ubylo, tvoří zhruba pětinu z vybraných (21,3 %) tokenů. Spojení sinokorejských slov se slovy cizího původu a kombinace všech tří původů se stále vyskytuje ve velmi malých procentech, navíc je při pohledu do tabulky poměrně jasné, že by se jejich zastoupení nemělo počítat, protože jediné

spojení 신형코로나비루스감염증피해 [sinhjǒngkchoronabirusŭkamjǒmdžūngpchihä] obsahuje části všech tří původů. Tento výraz ovšem nelze považovat za jedno slovo, jak rozebereme v následující části.

6.2.3 Shluky vybraných slov z Rodong sinmun

Nejvýrazněji ze všech shluků vystává u Rodong sinmun fráze:

„계속 확대되고 있는 신형 코로나 비루스 감염증 피해, 그에 대처하기 위한 노력“
[kjesok hwaktädōgo innūn sinhjǒng kchorona pirusŭ kamjǒmdžūng pchihä, kŭe täčchōhagi ūhan norjōk]
„Stále se zvětšující snaha vyhnout se a vypořádat se s infekcí koronaviru nového typu“

a její varianty.⁷⁵ Jde o jedno z opakujících se hesel, které slouží v podstatě jako titul či podtitul článků věnovaných koronavirové tematice a v průběhu roku zůstává prakticky neměnné. Jedinou změnou by byla absence slova stále 계속 [kjesok] nebo alternativní pořadí některých členů:

„신형 코로나 비루스 감염증 피해 계속 확대, 그에 대처하기 위한 노력“
[sinhjǒng kchorona pirusŭ kamjǒmdžūng pchihä kjesok hwaktä, kŭe täčchōhagi ūhan norjōk]

Sloveso 확대되다 [hwaktädōda] rozšiřovat se, expandovat se téměř vždy vyskytlo právě ve zmíněném hesle. Mimo něj bylo užito například v této větě:

정부는 신형코로나비루스감염증대책본부회의를 열고 감염이 **확대되고있는** 미국에서 들어오는 입국자들에 대해 지정된 장소들에서 2주일동안 격리시키기로 하였다.

[čōngbunūn sinhjǒngkchoronabirusŭkamjǒmdžūngtäčhäkbunbuhōürül jōlgo kamjōmi **hwaktädōgoinnūn** migugesō tūrōonūn ipgukččadūre tähä čidžōngdōn čangsodūresō i čuil Tongan kjōngni sikchigiro hajōtta]

Vláda uspořádala schůzi ústředí pro kontrolu infekce koronaviru nového typu a rozhodla, že lidé přilétající z USA, kde se nákaza **šíří**, budou podrobování dvoutýdenní karanténě v určených místech.

Nejde o slovo exkluzivní pro Severní Koreu. Našli bychom i ve druhém korpusu, ale objevuje se v zanedbatelném množství.

Pokud porovnáme jednotlivá slova hesla s tabulkou nejvíce frekventovaných slov, zjistíme, že většina z nich se v seznamu vyskytuje ve velkém počtu. Je možné, že tato fráze nevhodně ovlivnila výsledky, a bylo by proto zajímavé vygenerovat nový seznam frekventovaných slov po odebrání tohoto hesla a srovnat jej se stávajícím.

V tabulkách 6 a 7 svou délkou poněkud „svítí“ již zmíněný typ 신형코로나비루스감염증피해 [sinhjǒngkchoronabirusŭkamjǒmdžūngpchihä] a varianty jako 신형코로나비루스감염증 [sinhjǒngkchoronabirusŭkamjǒmdžūng]. Pokud jej však rozdělíme na jednotlivá slova, dostaneme velmi podobné označení pro koronavirus jako to, se kterým jsme se setkali v deníku Čosōn ilbo.

⁷⁵ Zde jsem pro přehlednost slova oddělila mezerami i tam, kde se v původním textu nevyskytovaly.

신형 [新型, sinhjöng] je zde vlastně ekvivalentem pro 신종 [sindžong]. První slovo, použito zde, je složeno ze znaku 新 (새 신, [sä, sin], nový) a 型 (모형 형, [mohjöng, hjöng], *model, typ, druh*). Slovo, které pro nazývání COVID-19 zvolili v Čosön ilbo, má stejný první znak s významem nový, ale druhý morfém je 種 (씨 종, [ssi, džong], *semeno*, v použití pak *druh, typ, odrůda*). Oba vyjadřují novou formu koronaviru.

Ve spojení 코로나 바이러스 [kchorona pirusü] je změna pouze v pravopise a výslovnosti slova virus, jelikož KLDK zvolila zachovat podobu v latině, zatímco v Korejské republice je předepsáno používat slovo tak, jak bylo převzato z angličtiny.

감염 [kamjöm] či 감염증 [kamjömdžüng] nákaza nebo infekce byla stejně užívána i v Jižní Koreji.

V uvedeném hesle se navíc název koronaviru v zásadě automaticky pojí se slovem 피해 [pchihä] vyhýbání, únik, vyvarování se.

Nepřilíš jinak je tomu u odvozeného slova 신형코로나바이러스감염자 [sinhjöngkchoronabirusükamjömdža] – člověk nakažený koronavirem nového typu, kde opět oddělením mezerami získáme slova 신형 [sinhjöng] – nová podoba, typ, 코로나 바이러스 [kchorona pirusü], případně 코로나바이러스 [kchoronabirusü] – koronavirus a 감염자 [kamjömdža] – nakažený, jež by už jako samostatné typy v seznamu nepůsobily nijak rušivě ani neobvykle.

Další shluky (ať už původně odděleny mezerou či nikoli) byly typicky tvořeny 2-3 slovy.

감염자

Místo „potvrzených případů“ bylo v případě Rodong sinmun v pravidelných statistikách častěji užíváno slovo 감염자 [kamjömdža] – nakažený, infikovaný. Ze shluků pak můžeme zmínit například: 감염자 발생 [kamjömdža palsäng] – výskyt nakažených, 감염자 증가 [kamjömdža čüngga] – nárůst nakažených. Opět se v žádném případě nejedná o slovo, které by nebylo použito i v tisku v Jižní Koreji, na seznamu z korpusu Čosön ilbo jej najdeme na 587. místě. Toto a dalších několik slov nám tedy naznačují, že část rozdílů v lexiku je především otázkou preference, ne nutně neexistencí stejného slova ve spisovném jazyce druhého státu.

사망자

Kromě počtu nakažených byly často uváděny i počty obětí koronaviru. 사망자 [samangdža] – zemřelý bylo v textu v těchto spojeních: 사망자가 났다 [samangdžaga natta], 사망자가 발생하였다 [samangdžaga palsänghajötta], 사망자가 등록되었다 [samangdžaga tünngnokdöjötta].

Typicky se obě tato klíčová slova vyskytovala v úzké blízkosti:

„...감염자는 5 5 만 7 1 3 5 명, 사망자는 1 만 5 5 5 8 명이다.“

[kamjömdžanün osiboman čchilčchönpäksamsibo mjöng, samandžanün man očchönopäkosippchal mnjöngida]

„...nakažených případů je 557 135, zemřelo 15 558 lidí.“

„일본에서 1 만 7 7 9 0 명이 **감염되고** 그중 9 2 3 명이 **사망하였다**.“
[ilboneso man čchilčchönčchilpökkusip mjöngi **kamjödögo** küdzung kupäkisipsam mjöngi **samanghajötta**]
„V Japonsku se **nakazilo** 17 790 lidí, z toho 923 **zemřelo**.“

막기

Slovo **막기** [makki] – zamezení, prevence se přirozeně pojilo se substantivy s významem něčeho nežádoucího. V případě tématu koronaviru jde o slova označující samotný koronavirus nebo nemoc jím způsobenou: 신형코로나비루스감염증을 막기 [sinhjöngkchoronabirusükamjödžüngül makki] – prevence proti nákaze koronavirem nového typu, 전파를 막기 [čönpcharül makki] – zamezit šíření, 신형코로나비루스감염증을 철저히 막기 [sinhjöngkchoronabirusükamjödžüngül čchöldžohi makki] – striktně zamezit infekci koronavirem nového typu, 신형코로나비루스감염증의 전파를 막기 [sinhjöngkchoronabirusükamjödžüngüi čönpcharül makki], 전염병을 막기 – zamezit šíření nákazy koronavirem nového typu, 류입을 막기 [rjuibül makki] – zamezit vlně, přílivu (nákazy), 악성비루스의 전파를 막기 [aksöngbirusüüi čönpcharül makki] – zamezit šíření zákeřného viru.

전파

„Šíření“, které se objevilo již ve spojení výše, bylo poměrně produktivní, co se pestrosti shluků týče: 전파를 막기 위해 [čchönpcharül makki ühä] – za účelem zamezení šíření, 전파를 억제하기 위해 [čönpcharül ökčehagi ühä] – za účelem potlačení šíření, 전파를 막기 위한 방역조치들 [čönpcharül makki ühan pangjökčočchidül] – preventivní kroky pro zamezení šíření

일꾼들

일꾼들 [ilkkundül] – dělník je jedno z mnoha slov, které se liší v obou státech i svým pravopisem. Ve standardní korejštině v Korejské republice je psáno 일꾼 [ilkkun] v souladu se skutečnou výslovností a v novinách se až na velmi specifickou výjimku nenalézalo. V Rodong sinmun tvořilo tato spojení: 일꾼들과 근로자들이 [ilkkundülgwa küllodžadüri] – dělníci a pracující, 일꾼들과 당원들과 근로자들이 [ilkkundülgwa tangwöndülgwa küllodžadüri] – dělníci, členové strany a pracující, 모든 일꾼들 [modün ilkkundülgwa] – všichni dělníci.

7 Závěr

Sbírání dat pro sestavení korpusu bylo časově náročnější, než jsem očekávala. Články bylo třeba kopírovat do textového souboru manuálně a po jednom. Poté jsem zbavila články doplňujících informací, které by svým opakováním zkreslily výsledky analýzy. Při výběru článků pro korpus, který bude prezentovat tisk Korejské lidově demokratické republiky, jsem narazila na omezený přístup k severokorejským zdrojům, s čímž jsem počítala. Nečekaným problémem ovšem byl fakt, že webové stránky, na nichž se nachází archiv starších článků Rodong sinmun, byly velmi nestabilní a náchylné na výpadky a další problémy. Toto činilo manuální shromažďování textu značně vyčerpávajícím.

Pokud jde o deník Čosön ilbo, jenž tvořil korpus z novin Jižní Koreje, zde bylo naopak problémem velké množství článků, z nichž jsem musela vybrat jen ty, které budou pro práci relevantní. Nakonec jsem zvolila jako rozhodující kritérium rubriku, do které článek spadá. Tento krok sice pomohl vyselektovat pouhý zlomek výsledků, které web nabídl po hledání klíčového slova 코로나 [kchorona], ale přesto byl vzorek nepoměrně větší než soubor článků z Rodong sinmun. Vzhledem k seznamu klíčových slov by možná porovnání Čosön ilbo i Rodong sinmun vůči jinému korpusu zcela odlišného typu místo mezi sebou přineslo uspokojivější výsledky.

Nicméně seznamy výsledků vygenerované softwarem AntConc pro Čosön ilbo vyšly víceméně k mé spokojenosti a s pomocí klíčových slov a shluků se s korpusem pracovalo výborně. Různorodá tematika článků odrážela celosvětový trend a to, jak se některé výrazy během roku vyvíjely, byly nahrazovány jinými, upřesněny atp.

Opak by se dal říci o korpusu sestávajícího ze článků deníku Rodong sinmun. Až na výjimky zůstávaly téma i forma článků po celý rok 2020 stejné nebo velmi podobné a nenabízely vhodný ekvivalent k mnoha spojením, která vešla v užívání na Jihu. Celou analýzu navíc komplikovala odlišná norma v otázce psaní mezer mezi slovy a z toho plynoucí nepraktická dlouhá slovní spojení, která ve vygenerovaných seznamech však figurovala jako jeden typ. Přes tato úskalí lze pozorovat větší či drobnější rozdíly v lexiku preferovaném v Korejské republice a v Korejské lidově demokratické republice.

Konkrétně klíčová slova z deníku Čosön ilbo tvořila rozmanitější a mnohem větší množství shluků a nabízela mnoho druhů vět, které ukazovaly jejich užití v různém kontextu. To odpovídá rozmanitosti samotných článků, které tvoří daný korpus. Naopak pro Rodong sinmun bylo typičtější velmi husté zastoupení vybraných klíčových slov. Ve většině příkladových vět se tak vyskytovalo hned několik shluků. Navíc se velmi často opakovala variace na jednu frázi sestávající rovněž hned z několika klíčových slov.

Přesto, že použití druhého tisku coby referenčního korpusu nebylo ideální, je nakonec ze seznamu klíčových slov i rozboru shluků znát pravděpodobné zaměření článků a jejich obsah. Rodong sinmun hlavně informoval o počtu případů ve světě, dopadu koronaviru na lidi a povzbuzoval k prevenci především obecnými hesly. Na druhé straně z výsledků Čosön ilbo vidíme i postupný vývoj od opatrnosti z neznámého a varování, přes nedostatek ochranných pomůcek, čekání na vývoj spolehlivé vakcíny a opatření pro její efektivní distribuci až ke konkrétním krokům v případě nákazy. Dalo by se říci, že se vyjadřuje více prakticky a méně se snaží budit sounáležitost.

Lexikum s nejvyšší frekvencí i klíčová slova sice byly rozdílné, míra divergence ale vůbec neodpovídala hypotéze a mým očekáváním uvedeným v úvodních kapitolách. Rozdíly jsou především

otázkou preference jednoho ze synonymních výrazů, které ale nutně nemusí být ani jiného původu. Pokud bylo užíváno jiné slovo stejného významu, bylo zastoupeno i ve druhém korpusu, jen na mnohem nižší pozici co se týče frekvence. Purifikační snahy se tedy pravděpodobně dotýkají výrazněji jiné tematiky nebo oblasti jazyka, než je tisk na téma pandemie neznámého viru.

Jeden z rozdílů byl v samotném názvu koronavirové pandemie, kde se liší výchozí cizí jazyk, z něhož je název přejímán. Korejská republika vychází z anglického slova *virus*, které foneticky přepisuje jako 바이러스 [pairösü], od čehož pochází název 코로나 바이러스 [kchorona pairösü], zatímco v Severní Koreji je upřednostněn latinský původ a od něj se odvíjí pravopis. Virus se zde píše 비루스 [pirusü] a koronavirus je 코로나 비루스 [kchorona pirusü].

V označení „koronavirus nového typu“ se také liší 신형 코로나 비루스 [sinhjŏng kchorona pirusü] v Rodong sinmun a 신종 코로나 바이러스 [sindžong kchorona pairösü] v jihokorejském Čosŏn ilbo. Jde o slova velmi podobného významu a je zajímavé vidět jasnou preferenci jiné varianty v každé ze zemí. Další rozdíly poté byly například ve slovech 확산 [hwaksan] a 전파 [čönpcha] – šíření.

Pro přesnější výsledky by nejspíš bylo třeba zvolit jiný referenční korpus a věnovat čas lemmatizaci typů.

Snahy o sjednocení jazyka ať už pro účely diplomacie nebo případné budoucí unifikace Korejí nejsou novou záležitostí, a přestože tato analýza neukázala tak výrazné rozdíly, dvě pravopisné normy pro jeden jazyk jsou problematickou záležitostí pro mezikorejskou komunikaci, přestože běžný rozhovor mezi Jihokorejcem a Severokorejcem by připomínal divergenci spíše na úrovni nářečí. V současnosti je běžně dostupný online slovník kulturního jazyka KLCDR a porovnávací nástroj, který nabízí jihokorejské ekvivalenty.⁷⁶

⁷⁶ Slovník dostupný z: <https://nkinfo.unikorea.go.kr/nkp/term/termDicaryList.do>

Porovnání: <https://nkinfo.unikorea.go.kr/nkp/term/skNkLangCompare.do>

8 Bibliografie

Prameny

SCOTT, Mike a Chris TRIBBLE. *Textual patterns: key words and corpus analysis in language education*. Amsterdam; Philadelphia: John Benjamins Pub, 2006. Studies in corpus linguistics v. 22. ISBN 978-90-272-2293-0.

Čosŏn Ilbo 조선일보 [online]. 2020 [cit. 2022-06-15]. Dostupné z: <https://www.chosun.com/>.

Pchjodžungugŏ täsadžŏn 표준국어대사전. Kungnip kugŏwŏn 국립국어원 [online]. [cit. 20.08.2022]. Dostupné z: <https://stdict.korean.go.kr/>

Rodong Sinmun 로동신문 [online]. 2020 [cit. 2022-06-14]. Dostupné z: <http://www.rodong.rep.kp/>.

Uri malsäm 우리말샘. Kungnip kugŏwŏn 국립국어원 [online]. [cit. 2022-08-20]. Dostupné z: <https://opendict.korean.go.kr/>

ANTHONY, Laurence *AntConc*, version 3.5.8 (Windows) [software]. 2019 [cit. 2021-12-07]. Dostupné z: <https://www.laurenceanthony.net/software/antconc/>.

Sekundární zdroje

Publikace v češtině

ECKERT, Carter J. *Dějiny Koreje*. Praha: Nakladatelství Lidové noviny, 2001. Dějiny států. ISBN 80-7106-411-4.

ČERMÁK, František, 2011. *Jazyk a jazykověda: přehled a slovníky* [online]. Vydání 4., v Karolinu 2., doplněné. Praha: Univerzita Karlova, nakladatelství Karolinum [cit. 2022-08-15]. ISBN 978-80-246-2360-3. Dostupné z: <https://ebookcentral.proquest.com/lib/mlpe-ebooks/detail.action?docID=1996747>

ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017. ISBN 978-80-246-3710-5.

PUCEK, Vladimír. *Gramatika korejského jazyka*. Praha: Karolinum, 2005. ISBN 978-80-246-1004-7.

PUCEK, Vladimír. *Jazyková politika v Koreji*. Praha: Karolinum, 2000. ISBN 80-246-0066-8.

PUCEK, Vladimír. *Lexikologie korejštiny*. Praha: Karolinum, 2013. ISBN 978-80-246-2204-0.

Publikace v angličtině

KUMATANI, Akiyasu. Language Policies in North Korea. *International Journal of the Sociology of Language* [online]. 1990, roč. 82 [cit. 2020-04-16]. ISSN 0165-2516. DOI: <https://doi.org/10.1515/ijsl.1990.82.87>.

LEE, Sung-nyong. On the Need of Teaching Chinese Characters. In: *Korea Journal*, April 1972, 12(4), p. 49-52.

YI, Ki-mun a S. Robert RAMSEY. *A history of the Korean language*. Cambridge ; New York: Cambridge University Press, 2011. ISBN 978-0-521-66189-8.

Naming the coronavirus disease (COVID-19) and the virus that causes it, 2020. *World Health Organisation* [online]. Geneva: World Health Organization [cit. 2022-08-14]. Dostupné z: [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it)

Chosun Ilbo Has Korea's Biggest Circulation. *Chosun Ilbo* [online]. Seoul: Chosun Media, 2010, November 30, 2010 12:49 [cit. 2022-08-18]. Dostupné z: http://english.chosun.com/site/data/html_dir/2010/11/30/2010113001011.html

Publikace v korejštině

ČCHÖ, Hočchöl 최호철. Sŏro tarŭn ōhwiŭi tongčilsŏngŭl wihan pangan 서로 다른 어휘의 동질성을 위한 방안 (Návrh na sjednocení vzájemně odlišného lexika). *Sägugŏsānghwal 새국어생활* [online]. Kungnip kugŏwŏn 국립국어원, 2001, 11(1), s.61-72 [cit. 2022-08-19]. Dostupné z: https://www.korean.go.kr/nkview/nklife/2000_4.htm

9 Seznam tabulek a grafů

Tabulky

Tabulka 1: Rozsah korpusů užitých v práci	12
Tabulka 2: Výchozí jazyky pro výpůjčky cizího původu v korejštině	21
Tabulka 3: Zastoupení lexika dle původu ve slovníku Urimalsäm	22
Tabulka 4: Nejfrekventovanější slova ve vybraných článcích Čosön ilbo	27
Tabulka 5: Klíčová slova vybraných článků Čosön ilbo	31
Tabulka 6: Nejfrekventovanější slova ve vybraných článcích Rodong sinmun	37
Tabulka 7: Klíčová slova vybraných článků Rodong sinmun	39

Grafy

Graf 1: Zastoupení lexika dle původu ve frekventovaných výrazech ze článků Čosön ilbo	30
Graf 2: Zastoupení lexika dle původu ve frekventovaných výrazech ze článků Čosön ilbo	32
Graf 3: Zastoupení lexika dle původu ve frekventovaných výrazech ze článků Rodong sinmun	39
Graf 4: Zastoupení lexika dle původu ve klíčových slovech ze článků Rodong sinmun	41